Skript: Lineare Algebra, II

A. Kresch

Frühling 2017

Inhaltsverzeichnis

1 Elementare Modultheorie mit Anwendungen

Analog zu endlichdimensionalen Vektorräumen über einem Körper sind endlich erzeugte freie Moduln über einem nichttrivialen kommutativen Ring mit Eins. (Ein kommutativer Ring R mit Einselement $1 \in R$ gilt als nichttrivial, wenn $1 \neq 0$ in R, d.h., wenn $R \neq \{0\}$.) Die Kardinalität einer Basis heisst dann *Rang*, nach wie vor werden lineare Abbildung zwischen endlich erzeugten freien Moduln durch Matrizen dargestellt.

Ein Grund für das Interesse liegt an der folgenden Bemerkung.

Sei K ein Körper mit Polynomring $K[T]$. Dann wird ein K-Vektorraum V mit Endomorphismus $f\colon V \to V$ eindeutig durch die Struktur eines $K[T]$-Moduls bestimmt.

Denn ein $K[T]$-Modul ist eine abelsche Gruppe V mit Skalarmultiplikation $K[T] \times V \to V$. Die abelsche Grupe V mit eingeschränkter Skalarmultiplikation $K \times V \to V$ ist ein K-Vektorraum. Zudem bestimmt Skalarmultiplikation mit T einen Endomorphismus $f\colon V \to V$. Die Eigenschaften bestimmen eindeutig die Skalarmultiplikation mit einem beliebigen Polynom.

1.1 Endlich erzeugte Moduln über euklidischen Ringen

Betrachten wir ein kommutativer Ring R mit Einselement als R-Modul über sich selbst, so heisst ein Untermodul von R ein **Ideal** von R. Das Argument, wonach das Minimalpolynom einer quadratischen Matrix über einem Körper wohldefiniert ist, können wir so fassen:

Sei R ein euklidischer Ring, $f\colon R \setminus \{0\} \to \mathbb{N}$ ein euklidischer Betrag. Ist $I \subset R$ ein von Null verschiedenes Ideal und $0 \neq a \in I$ ein Element mit der Eigenschaft $f(a) \leq f(b)$ für alle $0 \neq b \in I$, so gilt $I = aR$.

Solches a teilt alle Elemente von I. Denn ein nicht durch a teilbares Element von R kann als $qa + r$ geschrieben mit q, $r \in R$, $r \neq 0$ und $f(r) < f(a)$ und deshalb kann nicht zum Ideal I gehören.

Proposition 1.1. *Ist R ein euklidischer Ring und M ein endlich erzeugter freier R-Modul, so sind alle Untermoduln von M endlich erzeugt, frei und vom Rang $\leq \operatorname{rk} M$.*

Beweis. Sei $n := \operatorname{rk} M$. Die Aussage wird durch Induktion nach n bewiesen, mit Fällen $n = 0$ (trivial) und $n = 1$ (obige Bemerkung) als Induktionsanfang. Sei $n \geq 2$. Es genügt, den Fall $M = R^n$ zu behandeln. Sei $M' \subset R^n$ ein Untermodul. Sei $\mathrm{pr}_n\colon R^n \to R$ die $n$te Projektionsabbildung $(a_1, \dots, a_n) \mapsto a_n$. Dann ist $\mathrm{pr}_n(M')$ ein Ideal von R. Ist $\mathrm{pr}_n(M') = 0$, so gilt $M' \subset R^{n-1}$ und M' ist endlich erzeugt, frei und vom Rang $\leq n-1$ nach der Induktionshypothese. Sonst gibt es ein $0 \neq a' \in R$, so dass $\mathrm{pr}_n(M') = a'R$. Es gibt also ein Element $(a'_1, \dots, a'_{n-1}, a') \in M'$, während ein allgemeines Element von M' die Form $(a_1, \dots, a_{n-1}, a'b)$ hat. Durch

$$(a_1, \dots, a_{n-1}, a'b) = (a_1 - a'_1 b, \dots, a_{n-1} - a'_{n-1} b, 0) + b(a'_1, \dots, a'_{n-1}, a')$$

haben wir

$$M' \cong (M' \cap R^{n-1}) \oplus R.$$

Nach der Induktionshypothese ist $M' \cap R^{n-1}$ endlich erzeugt, frei und vom Rang $\leq n-1$, also ist M' endlich erzeugt, frei und vom Rang $\leq n$. $\square$

Theorem 1.2. *Sei R ein euklidischer Ring und $A \in \mathrm{Mat}(m \times n, R)$.*
(i) *Es gibt ein $r \in \mathbb{N}$ sowie $S \in GL_m(R)$ und $T \in GL_n(R)$, jeweils ein endliches Produkt von Elementarmatrizen, so dass*

$$SAT = \begin{pmatrix} a_1 & & 0 & 0 & \cdots & 0 \\ & \ddots & & & & \\ 0 & & a_r & 0 & \cdots & 0 \\ 0 & & \cdots & & & 0 \\ 0 & & \cdots & & & 0 \end{pmatrix},$$

wobei $a_1, \ldots, a_r \in R$ von Null verschieden sind, mit $a_i \mid a_{i+1}$ für alle i.
(ii) *Sei $S' \in GL_m(R)$, $T' \in GL_n(R)$, so dass $S'AT'$ eine Blockstruktur wie oben hat, mit $r' \times r'$-Diagonalblock und von Null verschiedenen Diagonaleinträgen $a'_1, \ldots, a'_{r'} \in R$ mit $a'_i \mid a'_{i+1}$ für alle i. Dann gilt $r' = r$ und es gibt Einheiten $u_1, \ldots, u_r$, so dass $a'_i = u_i a_i$ für alle i.*

Für den Beweis brauchen wir das folgende Lemma.

Lemma 1.3. *Sei R ein euklidischer Ring. Eine von Null verschiedene Matrix $A = (a_{ij}) \in \mathrm{Mat}(m \times n, R)$ kann stets mit elementaren Zeilen- und Spaltenumformungen in eine Form*

$$\begin{pmatrix} d & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & * & \\ 0 & & & \end{pmatrix}$$

für ein von Null verschiedenes Element $d \in R$ gebracht werden, wobei der mit $*$ gezeichnete $(m-1) \times (n-1)$-Block durch d teilbar ist.

Beweis. Sei $f\colon R \smallsetminus \{0\} \to \mathbb{N}$ ein euklidischer Betrag. Wir verwenden Induktion nach

$$k := \min\{f(a_{ij}) \,|\, 1 \le i \le m,\ 1 \le j \le n,\ a_{ij} \ne 0\}.$$

Da zwei Zeilen mit elementaren Zeilenumformungen bzw. zwei Spalten mit elementaren Spaltenumformungen vertauscht werden können, kann angenommen werden, dass $a_{11} \ne 0$ mit $f(a_{11}) = k$.

Zuerst betrachten wir den Fall, $a_{11} \mid a_{ij}$ für alle i und j. Dann kann für alle $i \ge 2$ mit einer elementaren Zeilenumformung der Eintrag an der $(i,1)$-Stelle annulliert werden. Ebenso kann mit einer elementaren Spaltenumformung der Eintrag an der $(1,j)$-Stelle annulliert werden für alle $j \ge 2$. Dadurch kommt die Matrix in die gewünschte Form.

Falls $k = 0$ gilt $a_{11} \mid a_{ij}$ für alle i und j, also sind wir im obigen Fall. Falls es i und j mit $a_{11} \nmid a_{ij}$ gibt, dann gilt $k > 0$ und wir kommen so zu einer Situation, in der wir die Induktionshypothese anwenden können. Gilt $a_{11} \nmid a_{i1}$ für ein $i \geq 2$, so schreiben wir

$$a_{i1} = qa_{11} + r$$

mit q, $r \in R$, $r \neq 0$ und $f(r) < k$ und kommen durch Addition des $(-q)$-Vielfachen der ersten Zeile zur $i$ten zu einer Matrix mit kleinerem Wert von k. Gilt $a_{11} \nmid a_{1j}$ für ein $j \geq 2$, so gehen wir ähnlich mit einer elementaren Spaltenumformung vor. Gilt $a_{11} \mid a_{i1}$ für alle i und $a_{11} \mid a_{1j}$ für alle j, so können wir wie oben die sonstigen Einträge in der ersten Zeile und Spalte annullieren, dann addieren wir für ein i und j mit $a_{11} \nmid a_{ij}$ die $j$te Spalte zur ersten, damit der so entstehende $(i, 1)$-Eintrag nicht durch a_{11} teilbar wird. So kommen wir wie oben durch Addition eines Vielfachen der ersten Zeile zur $i$ten zu einer Matrix mit kleinerem Wert von k. In jedem Fall haben wir nach der Induktionshypothese das gewünschte Ergebnis. □

Wir bemerken, dass d in Lemma 1.3 die gebildete Matrix teilt, deshalb ist auch die ursprüngliche Matrix A durch d teilbar. Tatsächlich ist d in Lemma 1.3 ggT der Einträge von A.

Wir bezeichnen mit $s(A)$, für eine Matrix $A \in \mathrm{Mat}(m \times n, R)$, das grösstmögliche $s \in \mathbb{N}$, $s \leq \min(m, n)$, so dass:

- Der linke obere $s \times s$-Block von A ist eine Diagonalmatrix, die Diagonaleinträge $a_1, \ldots, a_s$ sind jeweils von Null verschieden und erfüllen $a_i \mid a_{i+1}$ für alle i.
- Der linke untere $(m - s) \times s$-Block und rechte obere $s \times (n - s)$-Block sind Null.
- Ist $s > 0$, so teilt a_s den rechten unteren $(m - s) \times (n - s)$-Block von A.

Beweis von Theorem 1.2(i). Für den Beweis verwenden wir fallende Induktion nach $s := s(A)$. Die Matrix A hat die gewünschte Form falls schon der rechte untere $(m - s) \times (n - s)$-Block von A Null ist. Falls nicht, dann wenden wir Lemma 1.3 auf diesen Block an. Ist $s > 0$, so haben wir $a_s \mid d$, weil a_s den Block teilt und d aus Lemma 1.3 ggT der Einträge aus dem Block ist. Also gilt für die Matrix A', die aus A entsteht, wenn wir die entsprechenden Zeilen- und Spaltenumformungen zum Block von A durchführen:

$$s(A') > s.$$

Nach der Induktionshypothese kann A' in die gewünschte Form gebracht werden, deshalb kann A in die gewünschte Form gebracht werden. □

Bemerkung 1.4. Implizit im Beweis von Theorem 1.2(i) ist ein iteratives Verfahren, eine Matrix A in die gewünschte Form zu bringen. Diese Form heisst *Smith-Normalform* und das iterative Verfahren kann in Fällen, wo ein euklidischer Betrag berechenbar ist und ein euklidisches Algorithmus existiert, als Algorithmus verstanden werden. So haben wir den *Smith-Algorithmus*, zum Beispiel, über dem Ring $\mathbb{Z}$ sowie über dem Ring $K[T]$ für einen Körper K, in dem wir berechnen können.

Der Beweis von Theorem 1.2(ii) ist abhängig von einem Struktursatz über endlich erzeugte Moduln über einem euklidischen Ring. Mindestens können wir die Aussage $r' = r$ aus (ii) belegen: es gilt $\ker(A) \cong R^{n-r} \cong R^{n-r'}$, deshalb $n - r = n - r'$.

Korollar 1.5. *Sei R ein euklidischer Ring und $n \in \mathbb{N}$. Dann kann jede Matrix aus $GL_n(R)$ als endliches Produkt von Elementarmatrizen geschrieben werden.*

Beweis. Sei $A \in GL_n(R)$. Nach Theorem 1.2(i) gibt es S, $T \in GL_n(R)$, jeweils ein endliches Produkt von Elementarmatrizen, so dass

$$A = S \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix} T,$$

mit $a_1, \ldots, a_n \in R^\times$ (weil A invertierbar ist). Die Diagonalmatrix ist Produkt von n Elementarmatrizen. □

Theorem 1.6 (Struktursatz für endlich erzeugte R-Moduln)**.** *Sei R ein euklidischer Ring und M ein endlich erzeugter R-Modul. Dann gibt es eindeutige natürliche Zahlen r und s sowie von Null verschiedene Nichteinheiten $a_1, \ldots, a_r$ mit der Eigenschaft $a_i \mid a_{i+1}$ für alle i, so dass*

$$M \cong R/a_1R \oplus \cdots \oplus R/a_rR \oplus R^s.$$

Die Elemente $a_1, \ldots, a_r$ sind eindeutig bis auf Multiplikation mit Einheiten.

Für den Beweis brauchen wir einige Definitionen sowie ein Lemma.

Definition. Sei R ein Integritätsbereich. Für ein R-Modul M definieren wir den **Torsionsuntermodul** $M_{\text{tors}} \subset M$ als

$$M_{\text{tors}} := \{m \in M \mid \exists\, a \in R,\, a \neq 0 : am = 0\}$$

sowie für $0 \neq a \in R$ den a-Torsionsuntermodul

$$M[a] := \{m \in M \mid am = 0\}$$

und a-primären Torsionsuntermodul

$$M(a) := \bigcup_{n \in \mathbb{N}_{>0}} M[a^n].$$

Ist $a \in R$ prim mit Restklassenkörper $k := R/aR$, so verleiht die Struktur von M als R-Modul eine Struktur von $M[a]$ als k-Vektorraum und allgemeiner für $\ell \in \mathbb{N}_{>0}$ eine Struktur von $M[a^\ell]/M[a^{\ell-1}]$ als k-Vektorraum.

Lemma 1.7 (Chinesischer Restsatz)**.** *Sei R ein euklidischer Ring, seien a, $b \in R$ teilerfremde von Null verschiedene Elemente. Dann haben wir mit $\bar{c} \mapsto (\bar{c}, \bar{c})$ einen Isomorphismus*

$$R/abR \to R/aR \oplus R/bR.$$

Beweis. Weil 1 ggT von a und b ist, gibt es r, $s \in R$ mit $1 = ra + sb$. Wir definieren $R/aR \oplus R/bR \to R/abR$ durch $(\bar{u}, \bar{v}) \mapsto sb\bar{u} + ra\bar{v}$. Die beiden Abbildungen sind inverse Homomorphismen. □

Beweis von Theorem 1.6. Sei $P \subset R$ eine Menge von Primelementen, in der genau ein Repräsentant aus jeder Menge $\{up \,|\, u \in R^\times\}$ mit p prim enthalten ist. Es gibt also für $0 \neq a \in R$ eine eindeutige Einheit u sowie für die Primfaktoren $p_1, \ldots, p_t$ von a aus P eindeutige positive ganze Zahlen $c_1, \ldots, c_t$, so dass

$$a = up_1^{c_1} \ldots p_t^{c_t}.$$

Mit dem iterierten chinesischen Restsatz haben wir dann

$$R/aR \cong R/p_1^{c_1}R \oplus \cdots \oplus R/p_t^{c_t}R.$$

Sei $m_1, \ldots, m_n \in M$ ein Erzeugendensystem von M mit n so klein wie möglich. Entsprechend ist ein surjektiver R-Modulhomomorphismus

$$R^n \to M.$$

Der Kern ist nach Proposition 1.1 isomorph zu R^r für ein $r \in \mathbb{N}$, $r \leq n$. Wir wenden Theorem 1.2(i) an: nachdem wir $m_1, \ldots, m_n$ durch ein anderes n-elementiges Erzeugendensystem ersetzen und eine geeignete Basis des Kerns auswählen wird M dargestellt als Kokern einer $(n \times r)$-Matrix

$$\begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_r \\ 0 & \cdots & 0 \\ 0 & \cdots & 0 \end{pmatrix}$$

mit von Null verschiedenen Einträgen $a_1, \ldots, a_r \in R$ und $a_i \mid a_{i+1}$ für alle i. Es gilt $a_i \notin R^\times$ für alle i, denn eine Einheit als Eintrag würde zu einem Erzeugendensystem mit weniger als n Elementen führen. Wir haben also

$$M \cong R/a_1R \oplus \cdots \oplus R/a_rR \oplus R^{n-r}. \tag{1}$$

Insbesondere gilt $M_{\text{tors}} \cong R/a_1R \oplus \cdots \oplus R/a_rR$ und $M/M_{\text{tors}} \cong R^{n-r}$.

Seien $p_1, \ldots, p_t$ die Primfaktoren von a_r aus P. Wir schreiben

$$\begin{aligned} a_1 &= u_1 p_1^{c_{11}} \cdots p_t^{c_{1t}}, \\ &\cdots \\ a_r &= u_r p_1^{c_{r1}} \cdots p_t^{c_{rt}}, \end{aligned} \tag{2}$$

mit $u_1, \ldots, u_r \in R^\times$ und $c_{ij} \in \mathbb{N}$ für alle $1 \le i \le r$ und $1 \le j \le t$. Weil $a_i \mid a_{i+1}$ für alle i gilt

$$c_{1j} \le c_{2j} \le \cdots \le c_{rj}$$

für alle j. Aus (1) und

$$R/a_iR \cong R/p_1^{c_{i1}}R \oplus \cdots \oplus R/p_t^{c_{it}}R$$

folgt

$$M(p_j) \cong R/p_j^{c_{1j}}R \oplus \cdots \oplus R/p_j^{c_{rj}}R \tag{3}$$

fü alle j, mit $M(p) = 0$ für $p \in P \smallsetminus \{p_1, \ldots, p_t\}$.

Sei $k_j := R/p_jR$ für alle j. Mit der Bemerkung,

$$\dim_{k_j}(M[p_j^\ell]/M[p_j^{\ell-1}]) = |\{i \mid c_{ij} \ge \ell\}|$$

für alle $1 \le j \le t$ und $\ell \in \mathbb{N}_{>0}$, sehen wir, dass das Isomorphietyp von M die Menge $\{p_1, \ldots, p_t\}$ und für jedes p_j die Folge $c_{1j} \le \cdots \le c_{rj}$ in (3) eindeutig bestimmt.

Weil $r = \max\{\dim_{k_1}(M[p_1]), \ldots, \dim_{k_t}(M[p_t])\}$ und s der Rang vom endlich erzeugten freien R-Modul M/M_{tors} ist, sind r und s eindeutig bestimmt. Die Eindeutigkeit von $\{p_1, \ldots, p_t\}$ und von $(c_{ij})_{i=1,\ldots,r}$ für jedes p_j impliziert, dass $a_1, \ldots, a_r$ in (2) bis auf die Einheiten $u_1, \ldots, u_r$ eindeutig bestimmt sind. $\square$

Beweis von Theorem 1.2(ii). Wie schon erwähnt haben wir $r' = r$. Es gilt

$$\begin{aligned} \operatorname{coker}(A) &\cong R/a_1R \oplus \cdots \oplus R/a_rR \oplus R^{m-r} \\ &\cong R/a_1'R \oplus \cdots \oplus R/a_r'R \oplus R^{m-r}. \end{aligned}$$

Die Eindeutigkeit der Anzahl von Quotientenmoduln in Theorem 1.6 impliziert, dass gleich viele von $a_1, \ldots, a_r$ Einheiten sind, wie von $a_1', \ldots, a_r'$. Die Eindeutigkeit bis auf Multiplikation mit Einheiten impliziert, dass es Einheiten $u_1, \ldots, u_r$ gibt, mit $a_i' = u_ia_i$ für $i = 1, \ldots, r$. $\square$

1.2 Anwendungen

In den wichtigsten Fällen $R = \mathbb{Z}$ und $R = K[T]$ für einen Körper K bekommen wir nennenswerte Sätze.

Korollar 1.8 (Struktursatz für endlich erzeugte abelsche Gruppen)**.** *Sei G eine endlich erzeugte abelsche Gruppe. Dann gibt es eindeutige natürliche Zahlen r und s und eindeutig ganze Zahlen $a_1, \dots, a_r$, jeweils ≥ 2 mit $a_i \mid a_{i+1}$ für alle i, so dass*

$$G \cong \mathbb{Z}/a_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/a_r\mathbb{Z} \oplus \mathbb{Z}^s.$$

Der Beweis benutzt den Fakt, dass eine abelsche Gruppe G eine eindeutige Struktur als $\mathbb{Z}$-Modul besitzt.

Beweis. Wir betrachten G als $\mathbb{Z}$-Modul und wenden den Struktursatz für endlich erzeugte $\mathbb{Z}$-Moduln an. □

Korollar 1.9. *Sei $f\colon N' \to N$ ein Homomorphismus von freien endlich erzeugten $\mathbb{Z}$-Moduln desselben Rangs. Dann ist $|\det(f)|$ wohldefiniert und, falls positiv, die Ordnung der abelschen Gruppe* $\operatorname{coker}(f)$*, die im sonstigen Fall unendlich ist.*

Die *Ordnung* einer Gruppe G ist die Anzahl Elemente, falls G endlich ist.

Beweis. Weil $\mathbb{Z}^\times = \{\pm 1\}$, ist $|\det(f)|$ wohldefiniert. Wir wählen Basen von N' und N nach Theorem 1.2 aus, d.h., so dass f durch eine Diagonalmatrix dargestellt wird mit r von Null verschiedenen Diagonaleinträgen $a_1, \dots, a_r \in \mathbb{Z}$ mit $a_i \mid a_{i+1}$ für alle i. Es gilt $r = n$ genau dann, wenn $|\det(f)| > 0$. Dann ist

$$\operatorname{coker}(f) \cong \mathbb{Z}/a_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/a_n\mathbb{Z},$$

ein endliche Gruppe der Ordnung

$$|a_1 \cdots a_n| = |\det(f)|.$$

Sonst ist

$$\operatorname{coker}(f) \cong \mathbb{Z}/a_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/a_r\mathbb{Z} \oplus \mathbb{Z}^{n-r}$$

unendlich. □

Die Begleitmatrix eines normierten Polynoms $f \in K[T]$,

$$f = T^n + a_{n-1}T^{n-1} + \cdots + a_1 T + a_0$$

ist die Matrix

$$C(f) := \begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & -a_{n-2} \\ 0 & \dots & 0 & 1 & -a_{n-1} \end{pmatrix}.$$

mit charakteristischem Polynom f.

Korollar 1.10 (Frobenius-Normalform). *Sei $A \in \mathrm{M}_n(K)$ eine quadratische Matrix über einem Körper K. Dann gibt ein eindeutiges $r \in \mathbb{N}$ und eindeutige normierte Polynome $f_1, \ldots, f_r \in K[T]$ positiven Grades mit $f_i \mid f_{i+1}$ für alle i, so dass A ähnlich ist zur Blockdiagonalmatrix*

$$\begin{pmatrix} C(f_1) & & 0 \\ & \ddots & \\ 0 & & C(f_r) \end{pmatrix}$$

Beweis. Die zu A assoziierte lineare Abbildung $K^n \to K^n$ bestimmt eine Struktur von K^n als $K[T]$-Modul. Wir wenden die Struktursatz an:

$$K^n \cong K[T]/f_1K[T] \oplus \cdots \oplus K[T]/f_rK[T].$$

(Der Betrag s aus dem Struktursatz muss Null sein, weil $K[T]$ als K-Vektorraum unendlichdimensional ist.) Von jedem Faktor $K[T]/f_iK[T]$ wählen wir die Basis als K-Vektorraum

$$1, T, \ldots, T^{\deg(f_i)-1}$$

aus. Bezüglich dieser Basis ist die darstellende Matrix die Begleitmatrix von f_i. Auf diese Weise kommen wir zur oben gezeigten Blockdiagonalmatrix. Ist A ähnlich zu einer solchen Blockdiagonalmatrix, so ist die von A bestimmte $K[T]$-Modul K^n isomorph zu $K[T]/f_1K[T] \oplus \cdots \oplus K[T]/f_rK[T]$. Die Eindeutigkeit aus dem Struktursatz belegt die Eindeutigkeit der Frobenius-Normalform. □

Korollar 1.11 (Jordan-Normalform). *Sei $A \in \mathrm{M}_n(K)$ eine quadratische Matrix über einem Körper K. Angenommen, A ist trigonalisierbar. Dann ist A ähnlich zu einer Blockdiagonalmatrix, wobei die Diagonalblöcke sogenannte Jordan-Blöcke sind:*

$$\begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix},$$

mit $\lambda \in K$. Die Jordan-Blöcke sind eindeutig bis auf Permutation.

Wir erinnern uns, trigonalisierbarkeit ist äquivalent zur Faktorisierung des charakteristischen Polynoms in lineare Faktoren.

Beweis. Sei $\lambda_1, \ldots, \lambda_t$ die Eigenwerte von A. Nach der Hypothese gibt es eine Faktorisierung des charakteristischen Polynoms in $K[T]$ mit Faktoren

$$T - \lambda_1, \ldots, T - \lambda_t,$$

jedem Faktor mit entsprechender Vielfachheit.

Sei $f_1, \ldots, f_r \in K[T]$ wie in Korollar 1.10. Weil deren Produkt das charakteristische Polynom von A ist, haben wir

$$f_i = (T-\lambda_1)^{m_{i1}} \cdots (T-\lambda_t)^{m_{it}} \tag{1}$$

für natürliche Zahlen m_{ij}, $1 \le i \le r$, $1 \le j \le t$ mit

$$m_{1j} \le m_{2j} \le \cdots \le m_{rj} \tag{2}$$

für alle j.

Wir folgen dem Beweis von Korollar 1.10 und wenden den chinesischen Restsatz an:

$$\begin{aligned} K^n &\cong K[T]/f_1K[T] \oplus \cdots \oplus K[T]/f_rK[T] \\ &\cong \big(K[T]/(T-\lambda_1)^{m_{11}}K[T] \oplus \cdots \oplus K[T]/(T-\lambda_t)^{m_{1t}}K[T]\big) \\ &\quad \oplus \cdots \oplus \big(K[T]/(T-\lambda_1)^{m_{r1}}K[T] \oplus \cdots \oplus K[T]/(T-\lambda_t)^{m_{rt}}K[T]\big). \end{aligned}$$

Von einem Faktor

$$K[T]/(T-\lambda_j)^{m_{ij}}$$

wählen wir die Basis als K-Vektorraum

$$(T-\lambda_j)^{m_{ij}-1}, \ldots, T-\lambda_j, 1$$

aus, um zu einem Jordan-Block zu kommen.

Die Eindeutigkeit bis auf Permutation ist die Aussauge, dass die Anzahl Jordan-Blöcke jeder Grösse und zu jedem Eigenwert eindeutig bestimmt ist. Diese sind durch (1)–(2) von den Polynomen $f_1, \ldots, f_r$ bestimmt und letztere sind eindeutig. $\square$

1.3 Kommutierende Blockmatrizen

Blockmatrizen bieten Vorteile an. Insbesondere haben Blockdreiecksmatrizen die günstige Eigenschaft, dass die Determinante das Produkt der Determinanten der Diagonalblöcke ist. Wir beschreiben eine Situation, in der auch ohne Dreiecksform die Determinante günstig zu berechnen ist.

Wir betrachten eine quadratische Matrix mit $p \times p$-Blöcken für ein $p \in \mathbb{N}_{>0}$ über einem kommutativen Ring R mit Einselement. Eine solche Matrix hat die Grösse $kp \times kp$ für ein $k \in \mathbb{N}$ und kann betrachtet werden

- als $kp \times kp$-Matrix über R, sowie
- als $k \times k$-Blockmatrix mit $p \times p$-Blöcken.

Für $A \in \mathrm{Mat}(p \times p, R)$ bilden die Elemente $a_dA^d + \cdots + a_1A + a_0E_p$ für $d \in \mathbb{N}$ und $a_0, \ldots, a_d \in R$ ein kommutativer Ring $R[A] \subset \mathrm{Mat}(p \times p, R)$. Ein leichte Verallgemeinerung dieser Konstruktion ist möglich.

Lemma 1.12. *Sei R ein kommuativer Ring mit Einselement, $p \in \mathbb{N}_{>0}$ und $\Omega \subset \mathrm{Mat}(p \times p, R)$ eine Teilmenge mit der Eigenschaft,*

$$AA' = A'A$$

für alle A, $A' \in \Omega$. Dann ist die Menge aller R-Linearkombinationen von Produkten $A_1 A_2 \cdots A_d$ mit $d \in \mathbb{N}$ und $A_1, \ldots, A_d \in \Omega$ ein kommutativer Ring

$$R[\Omega] \subset \mathrm{Mat}(p \times p, R)$$

mit Einselement E_p, wobei das leere Produkt als Einheitsmatrix E_p verstanden wird.

Die Hypothese kann so mit Worten gefasst werden: „die Elemente aus Ω kommutieren paarweise“.

Beweis. Sei $R[\Omega]$ die in der Aussage beschriebene Menge. Es ist klar, dass die Summe zweier Elemente aus $R[\Omega]$ noch zu $R[\Omega]$ gehört. Auch das Produkt zweier Elemente aus $R[\Omega]$ gehört zu $R[\Omega]$. Es gilt $aE_p \in R[\Omega]$ für alle $a \in R$; insbesondere gilt $E_p \in R[\Omega]$. Die Behauptung, dass $R[\Omega]$ kommutativ ist, folgt aus der Gleichung

$$(A_1 \cdots A_d)(A'_1 \cdots A'_{d'}) = (A'_1 \cdots A'_{d'})(A_1 \cdots A_d) \tag{1}$$

für d, $d' \in \mathbb{N}$ und $A_1, \ldots, A_d, A'_1, \ldots, A'_{d'} \in \Omega$. Weil die Elemente aus Ω paarweise kommutieren, haben wir $(A_1 \cdots A_d)A' = A'(A_1 \cdots A_d)$ für alle $A' \in \Omega$ (Induktion nach d). Jetzt folgt (1) durch Induktion nach d'. $\square$

Proposition 1.13. *Sei R ein kommuativer Ring mit Einselement, seien k, $p \in \mathbb{N}_{>0}$, sei $A \in \mathrm{Mat}(kp \times kp, R)$ eine Blockmatrix $A = (A_{ij})_{1 \le i,j \le k}$ mit $A_{ij} \in \mathrm{Mat}(p \times p, R)$ für alle $1 \le i, j \le k$. Angenommen, die Elemente aus*

$$\Omega := \{A_{ij} \,|\, 1 \le i, j \le k\}$$

kommutieren paarweise. Dann gilt

$$\det(A) = \det(\det{}'(A)),$$

wobei wir an der rechten Seite A als Element von

$$\mathrm{Mat}(k \times k, R[\Omega])$$

betrachten und mit $\det'(A)$ die $k \times k$-Determinante

$$\det{}'(A) \in R[\Omega] \subset \mathrm{Mat}(p \times p, R)$$

bezeichnen.

Beweis. Zuerst beweisen wir die Proposition unter der zusätzlichen Annahme, dass es kein $0 \neq a \in R$ gibt, mit $a \det(A_{11}) = 0$. Wir verwenden Induktion nach k. Der Fall $k = 1$ ist klar. Für den Induktionsschritt definieren wir folgende Blockmatrizen:

$$B := \begin{pmatrix} E_p & & & 0 \\ & A_{11} & & \\ & & \ddots & \\ 0 & & & A_{11} \end{pmatrix}$$

und

$$C := \begin{pmatrix} E_p & & & 0 \\ -A_{21} & E_p & & \\ \vdots & & \ddots & \\ -A_{k1} & 0 & & E_p \end{pmatrix}.$$

Wir berechnen

$$CBA = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ 0 & & & \\ \vdots & & * & \\ 0 & & & \end{pmatrix}.$$

Deshalb gilt nach der Induktionshypothese

$$\begin{aligned} \det(A_{11})^{k-1} \det(A) = \det(CBA) &= \det(A_{11}) \det(*) \\ &= \det(A_{11}) \det(\det'(*)) \\ &= \det(A_{11})^{k-1} \det(\det'(A)), \end{aligned}$$

wobei wir für den letzten Schritt

$$A_{11}^{k-1} \det'(A) = \det'(CBA) = A_{11} \det'(*)$$

benutzen.

Für den allgemeinen Fall betrachten wir

$$R \subset R[T].$$

Wir definieren

$$\widetilde{A}_{11} := TE_p + A_{11} \in \mathrm{Mat}(p \times p, R[T])$$

und

$$\widetilde{A}_{ij} := A_{ij}$$

für alle $1 \leq i, j \leq k$ mit $(i, j) \neq (1, 1)$. Die Determinante von $\widetilde{A}_{11}$ ist ein normiertes Polynom und deshalb hat die Eigenschaft, dass das Produkt mit einem beliebigen von Null verschiedenen Polynom von Null verschieden ist. Es gilt also für die entsprechende Blockmatrix $\widetilde{A} \in \mathrm{Mat}(kp \times kp, R[T])$: $\det(\widetilde{A}) = \det(\det'(\widetilde{A}))$. Durch Auswertung an $T = 0$ folgt $\det(A) = \det(\det'(A))$. □

Quelle für Proposition 1.13 ist der Artikel „Determinants of commuting-block matrices“ von I. Kovacs, D. S. Silver und S. G. Williams, *American Mathematical Monthly* Band 106 (1999), S. 950–952.

2 Bilinearformen

Wir bekommen mit einer geeigneten Bilinearform die Struktur auf $\mathbb{R}^n$, die dazu führt, dass ein Vektor $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ eine Länge hat:

$$\|\mathbf{x}\| := \sqrt{x_1^2 + \cdots + x_n^2}.$$

2.1 Metrik und Norm

Wir definieren auf $\mathbb{R}^n$ den Abstand (auch Metrik genannt):

$$d(\mathbf{x}, \mathbf{y}) := \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}.$$

Definition. Sei X eine Menge. Eine **Metrik** auf X ist eine Abbildung

$$d\colon X \times X \to \mathbb{R}_{\geq 0}$$

mit den folgenden Eigenschaften für x, y, $z \in X$:

- $d(x, y) = 0$ genau dann, wenn $x = y$;
- $d(y, x) = d(x, y)$;
- $d(x, z) \leq d(x, y) + d(y, z)$.

Die angegebene Formel für die sogenannte euklidische Metrik auf $\mathbb{R}^n$ ist verbunden mit der Länge eines Vektors.

Definition. Sei V ein reeller Vektorraum. Eine **Norm** auf V ist eine Abbildung

$$\|\cdot\|\colon V \to \mathbb{R}_{\geq 0}$$

mit den folgenden Eigenschaften für v, $w \in V$ und $\lambda \in \mathbb{R}$:

- $\|v\| = 0$ genau dann, wenn $v = 0$;
- $\|\lambda v\| = |\lambda|\,\|v\|$;
- $\|v + w\| \leq \|v\| + \|w\|$.

Es ist klar: *Ist* $\|\cdot\|\colon V \to \mathbb{R}_{\geq 0}$ *eine Norm auf einem reellen Vektorraum* V*, so wird mit*

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

eine Metrik auf V *definiert.*

Die Normen, die die meiste Relevanz geniessen, kommen aus bilinearen Abbildungen $V \times V \to \mathbb{R}$ (vgl. Korollar 2.2, unten).

2.2 Euklidische Vektorräume

Sei V ein Vektorraum über einem Körper K. Eine **Bilinearform** auf V ist eine bilineare Abbildung $s\colon V \times V \to K$. Wir haben schon bilineare und multilineare Abbildungen im Kontext der Determinante gesehen: die Abbildung muss linear in jedem Faktor sein, wenn ein beliebiger Vektor $v \in V$ in den anderen Faktor eingesetzt wird. Für Bilinearformen gelten die Bezeichnungen *symmetrisch*, *schiefsymmetrisch*, *alternierend*: eine Bilinearform $s\colon V \times V \to K$

heisst ...	falls stets gilt
symmetrisch	$s(w, v) = s(v, w)$
schiefsymmetrisch	$s(w, v) = -s(v, w)$
alternierend	$s(v, v) = 0$

Eine alternierende Bilinearform ist stets schiefsymmetrisch. Falls $\mathrm{char}(K) \neq 2$: (i) die Begriffe *schiefsymmetrisch* und *alternierend* sind äquivalent; (ii) für eine symmetrische Bilinearform $s\colon V \times V \to K$ gilt die **Polarisationsformel**

$$s(v, w) = \frac{s(v + w, v + w) - s(v, v) - s(w, w)}{2}.$$

Definition. Sei V ein reeller Vektorraum. Eine symmetrische Bilinearform $s\colon V \times V \to \mathbb{R}$ auf V heisst **positiv definit**, falls $s(v, v) > 0$ für alle $0 \neq v \in V$.

Eine positiv definite symmetrische Bilinearform heisst auch **Skalarprodukt** und wird oft mit $\langle \cdot, \cdot \rangle$ bezeichnet, d.h., $(v, w) \in V \times V$ wird auf $\langle v, w \rangle \in \mathbb{R}$ abgebildet. Für $v \in V$ schreiben wir

$$\|v\| := \sqrt{\langle v, v \rangle}. \tag{1}$$

Ein **euklidischer Vektorraum** ist ein reeller Vektorraum mit Skalarprodukt.

Proposition 2.1 (Cauchy-Schwarzsche Ungleichung)**.** *Ist* V *ein euklidischer Vektorraum, so gilt*

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

für alle v*,* $w \in V$*, wobei das Skalarproduk mit* $\langle \cdot, \cdot \rangle$ *bezeichnet und* $\|v\|$ *wie in* (1) *definiert wird. Gleichheit gilt genau dann, wenn* v *und* w *linear abhängig sind.*

Beweis. Falls v und w linear abhängig sind, gilt $v = \lambda w$ oder $w = \lambda v$ für ein $\lambda \in \mathbb{R}$ und somit $|\langle v, w\rangle| = |\lambda| \, \|w\|^2 = \|v\| \, \|w\|$ bzw. $|\langle v, w\rangle| = |\lambda| \, \|v\|^2 = \|v\| \, \|w\|$.

Angenommen, v und w sind linear unabhängig. Es gilt

$$\begin{aligned}
0 &< \langle v, v\rangle \Big\langle w - \frac{\langle v, w\rangle}{\langle v, v\rangle} v, w - \frac{\langle v, w\rangle}{\langle v, v\rangle} v \Big\rangle \\
&= \langle v, v\rangle\langle w, w\rangle - \frac{\langle v, w\rangle}{\langle v, v\rangle}\langle v, v\rangle\langle w, v\rangle - \frac{\langle v, w\rangle}{\langle v, v\rangle}\langle v, v\rangle\langle v, w\rangle + \frac{\langle v, w\rangle^2}{\langle v, v\rangle^2}\langle v, v\rangle^2 \\
&= \langle v, v\rangle\langle w, w\rangle - \langle v, w\rangle^2
\end{aligned}$$

und es folgt, $|\langle v, w\rangle| < \|v\| \, \|w\|$. $\square$

Korollar 2.2. *Sei V ein euklidischer Vektorraum mit Skalarproduk $\langle \cdot, \cdot \rangle$. Dann ist die Abbildung $\|\cdot\| \colon V \to \mathbb{R}_{\geq 0}$ aus (1) eine Norm auf V.*

Beweis. Die ersten zwei Bedingungen für eine Norm gelten offensichtlich. Nach der Cauchy-Schwarzschen Ungleichung gilt

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

für alle v, $w \in V$ und die dritte Bedingung folgt. $\square$

Definition. Das **Standardskalarprodukt** auf $\mathbb{R}^n$ ist

$$\langle \mathbf{x}, \mathbf{y}\rangle := x_1 y_1 + \cdots + x_n y_n.$$

Die zum Standardskalarprodukt assoziierte Norm bzw. Metrik heisst **euklidische Norm** bzw. **euklidische Metrik**.

Wichtig ist die Formel:

$$\langle \mathbf{x}, \mathbf{y}\rangle = {}^t\mathbf{x}\,\mathbf{y},$$

wobei wir Vektoren aus $\mathbb{R}^n$ als Spaltenvektoren betrachten. Rechterseits erscheint das Produkt eines Zeilenvektors ($1 \times n$-Matrix) mit einem Spaltenvektor ($n \times 1$-Matrix), das eine 1×1-Matrix, d.h., eine reelle Zahl ergibt. Deshalb, für $A \in \mathrm{Mat}(n \times n, \mathbb{R})$:

$$\langle A\mathbf{x}, \mathbf{y}\rangle = {}^t\mathbf{x}\,{}^tA\,\mathbf{y} \qquad \text{und} \qquad \langle \mathbf{x}, A\mathbf{y}\rangle = {}^t\mathbf{x}\,A\mathbf{y}.$$

Mit der Cauchy-Schwarzsche Ungleichung können wir den Winkel zwischen zwei Vektoren definieren. Sei V ein euklidischer Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und seien v und w von Null verschidene Vektoren aus V. Es gilt also

$$-1 \leq \frac{\langle v, w\rangle}{\|v\| \, \|w\|} \leq 1$$

und wir definieren

$$\angle(v,w) := \cos^{-1} \frac{\langle v,w\rangle}{\|v\|\,\|w\|}.$$

Es gilt $0 \le \angle(v,w) \le \pi$.

Proposition 2.3 (Kosinussatz)**.** *Mit der obigen Notation gilt*

$$\|v-w\|^2 = \|v\|^2 + \|w\|^2 - 2\|v\|\,\|w\| \cos\angle(v,w).$$

Beweis. Es gilt

$$\begin{aligned} \|v-w\|^2 &= \langle v-w, v-w\rangle \\ &= \langle v,v\rangle - 2\langle v,w\rangle + \langle w,w\rangle \end{aligned}$$

und Letzteres ist $\|v\|^2 + \|w\|^2 - 2\|v\|\,\|w\| \cos\angle(v,w)$. □

Mit dem Begriff **orthogonal** wird die Bedingung $\langle v,w\rangle = 0$ angedeutet, geschrieben: $v \perp w$. Ferner definieren wir **orthogonales Komplement** eines Vektors $w^\perp := \{v \in V \mid v \perp w\}$ sowie einer Teilmenge $\Omega^\perp := \{v \in V \mid v \perp w \text{ für alle } w \in \Omega\}$ für $\Omega \subset V$. Eine Familie $(v_i)_{i\in I}$ von Vektoren aus V heisst orthogonal, falls $v_i \perp v_j$ für alle i, $j \in I$ mit $i \ne j$ und **orthonormal** falls zusätzlich $\|v_i\| = 1$ für alle $i \in I$ gilt. Wir bemerken, dass eine orthogonale Familie $(v_i)_{i\in I}$ mit $v_i \ne 0$ für alle i stets linear unabhängig ist. Denn aus $\sum_{i\in I} a_i v_i = 0$ (mit $a_i \in \mathbb{R}$ für alle $i \in I$ und $a_i \ne 0$ für nur endlich viele i) folgt durch $\langle \cdot, v_i\rangle$, dass $a_i = 0$ für alle i.

Proposition 2.4 (Gram-Schmidtsches Orthogonalisierungsverfahren)**.** *Sei V ein endlichdimensionaler euklidischer Vektorraum und $v_1, \ldots, v_m$ eine orthonormale Familie von Vektoren aus V. Dann gibt es eine Erweiterung zu einer Orthonormalbasis $v_1, \ldots, v_n$ von V.*

Beweis. Sei $n := \dim(V)$. Nachdem, was wir oben bemerkt haben, ist $v_1, \ldots, v_m$ eine linear unabhängige Familie. Also haben wir $m \le n$. Sei $W := \operatorname{span}(v_1, \ldots, v_m)$. Wir beweisen die Aussage durch Induktion nach $n-m$. Der Fall $n-m = 0$ ist trivial. Ist $n-m > 0$, d.h., $W \subsetneq V$, so gibt es ein $w \in V \setminus W$. Sei

$$w' := \langle w, v_1\rangle v_1 + \cdots + \langle w, v_m\rangle v_m.$$

und $w'' := w - w'$. So haben wir

$$w' \in W, \qquad w'' \in W^\perp, \qquad w = w' + w''$$

und w' und w'' sind eindeutig mit diesen Eigenschaften. Es gibt also einen von Null verschiedenen Vektor $w'' \in W^\perp$. Sei $\lambda := \|w''\|$, dann gilt $\lambda^{-1} w'' \in W^\perp$ mit $\|\lambda^{-1} w''\| = 1$. Sei

$$v_{m+1} := \lambda^{-1} w''.$$

Dann ist $v_1, \ldots, v_{m+1}$ eine orthonormale Familie und es gibt dann nach der Induktionshypothese eine Erweiterung zu einer Orthonormalbasis von V. □

Korollar 2.5. *Sei V ein endlichdimensionaler euklidischer Vektorraum und $W \subset V$ ein Untervektorraum. Dann gilt*

$$V = W \oplus W^\perp.$$

Beweis. Sei $n := \dim(V)$ und $m := \dim(W)$. Die Einschränkung von $\langle\cdot,\cdot\rangle$ auf W ist ein Skalarprodukt auf W. Es gibt nach dem Gram-Schmidtschen Orthogonalisierungsverfahren eine Orthonormalbasis $v_1, \dots, v_m$ von W. Wie im Beweis von Proposition 2.4 gibt es für alle $w \in V$ eindeutige Vektoren $w' \in W$ und $W'' \in W^\perp$ mit $w = w' + w''$. Also ist V direkte Summe von W und $W^\perp$. □

Als Allgemeinerung der Cauchy-Schwarzschen Ungleichung definieren wir die **Gramsche Determinante** von Vektoren $v_1, \dots, v_m$ aus einem euklidischen Vektorraum

$$G(v_1, \dots, v_m) := \det(\langle v_i, v_j\rangle)_{1\le i,j\le m}$$

mit den folgenden Eigenschaften.

Proposition 2.6. *Für Vektoren $v_1, \dots, v_m$ aus einem euklidischen Vektorraum V gilt stets*

$$G(v_1, \dots, v_m) \ge 0,$$

mit Gleichheit genau dann, wenn $v_1, \dots, v_m$ linear abhängig sind.

Beweis. Falls $v_1, \dots, v_m$ linear abhängig sind, ist es klar, dass die Gramsche Matrix (die Matrix aus der Definition der Gramschen Determinante) linear abhängige Zeilen hat. Deshalb ist die Gram-Determinante in diesem Fall Null.

Angenommen, $v_1, \dots, v_m$ sind linear unabhängig. Wir zeigen $G(v_1, \dots, v_m) > 0$ durch fallende Induktion nach dem grössten $k \in \mathbb{N}_{>0}$, so dass $v_1, \dots, v_k$ orthogonal sind. Als Induktionsanfang gilt der Fall $k = m$, so dass die Gramsche Matrix eine Diagonalmatrix ist mit positiven Diagonaleinträgen. Angenommen, $k < m$. Sei $W := \operatorname{span}(v_1, \dots, v_k)$ und

$$v_{k+1} = v'_{k+1} + v''_{k+1}$$

mit $v'_{k+1} \in W$ und $v''_{k+1} \in W^\perp$ (Korollar 2.5). Dann sind $v_1, \dots, v_k, v''_{k+1}$ orthogonal. Die Gramsche Matrix von

$$v_1, \dots, v_k, v''_{k+1}, v_{k+2}, \dots, v_m$$

entsteht aus derjenigen von $v_1, \dots, v_m$ durch eine Folge von elementaren Zeilenumformungen, gefolgt von einer Folge von elementaren Spaltenumformungen, jeweils des Typs „ein Vielfaches einer Zeile/Spalte zu einer anderen addieren“. Die Gramsche Determinante bleibt also unverändert und ist nach der Induktionshypothese positiv. □

Proposition 2.7 (Hadamardsche Ungleichung). *Seien $v_1, \ldots, v_m$ von Null verschiedene Vektoren aus einem euklidischen Vektorraum. Dann gilt*

$$\sqrt{G(v_1, \ldots, v_m)} \leq \|v_1\| \cdots \|v_m\|$$

mit Gleichheit genau dann, wenn $v_1, \ldots, v_m$ orthogonal sind.

Beweis. Falls $v_1, \ldots, v_m$ linear abhängig sind, ist die Aussage klar. Also können wir annehmen, dass $v_1, \ldots, v_m$ linear unabhängig sind. Wir folgen genau dem Beweis von Proposition 2.6 in diesem Fall. Also nehmen wir k, so gross wie möglich, so dass $v_1, \ldots, v_k$ orthogonal sind. Ist $k = m$, so haben wir

$$\sqrt{G(v_1, \ldots, v_m)} = \|v_1\| \cdots \|v_m\|.$$

Ist $k < m$, so schreiben wir $v_{k+1} = v'_{k+1} + v''_{k+1}$ mit $v'_{k+1} \in W$ und $v''_{k+1} \in W^\perp$. Nach der Induktionshypothese gilt

$$\sqrt{G(v_1, \ldots, v_m)} \leq \|v_1\| \cdots \|v_k\| \, \|v''_{k+1}\| \, \|v_{k+2}\| \cdots \|v_m\|.$$

Wir haben

$$\|v_{k+1}\|^2 = \|v'_{k+1}\|^2 + \|v''_{k+1}\|^2$$

mit $v'_{k+1} \neq 0$ und $v''_{k+1} \neq 0$ und es folgt,

$$\|v''_{k+1}\| < \|v_{k+1}\|.$$

Die Ungleichung folgt und gilt als echte Ungleichung. □

Definition. Seien X und X' Mengen und d bzw. d' eine Metrik auf X bwz. auf X'. Eine **Isometrie** von X nach X' ist eine bijektive Abbildung $f\colon X \to X'$ mit der Eigenschaft

$$d'(f(x), f(y)) = d(x, y)$$

für alle x, $y \in X$.

So, zum Beispiel, gibt es nach dem Gram-Schmidtschen Orthogonalisierungsverfahren eine lineare Isometrie von einem beliebigen endlichdimensionalen euklidischen Vektorraum V nach $\mathbb{R}^n$ mit der euklidischen Metrik, wobei $n = \dim(V)$.

Im Fall von euklidischen Vektorräumen V und V' ist jede Isometrie $V \to V'$, die $0 \in V$ auf $0 \in V'$ abbildet, tatsächlich linear. Diese Aussage heisst *Satz von Mazur-Ulam.* Es gilt für eine solche Isometrie $f\colon V \to V'$:

$$\|f(v)\|' = \|v\| \qquad \text{und} \qquad \|f(v) - f(w)\|' = \|v - w\|$$

für alle v, $w \in V$. So, zum Beispiel, gilt für eine Isometrie $g\colon V \to V$ mit $g(0) = 0$ und a, $b \in V$:

$$\|g(\frac{a+b}{2}) - g(a)\| = \|\frac{a-b}{2}\|.$$

Es folgt:

$$\|g\big(\frac{a+b}{2}\big) - g(a) + \frac{a-b}{2}\| \leq \|g\big(\frac{a+b}{2}\big) - g(a)\| + \|\frac{a-b}{2}\| \leq \|a-b\|. \tag{2}$$

Erfüllt g zudem die Bedingung

$$g(a) - g(b) = a - b, \tag{3}$$

so behaupten wir, dass g auch

$$g\big(\frac{a+b}{2}\big) = \frac{g(a)+g(b)}{2}$$

erfüllt. Sonst wäre die Norm der Differenz, die wir als den Ausdruck linkerseits in (2) erkennen, positiv. Dann wäre auch $\tilde{g}\colon V \to V$,

$$\tilde{g}(v) := g^{-1}\big(g(a)+g(b)\big) - g^{-1}\big(g(a)+g(b)-g(v)\big),$$

eine Isometrie, die 0 auf 0 abbildet und die Bedingung (3) erfüllt, für die aber

$$\begin{aligned}\|\tilde{g}\big(\frac{a+b}{2}\big) - \tilde{g}(a) + \frac{a-b}{2}\| &= \Big\|\frac{a+b}{2} - g^{-1}\Big(g(a)+g(b)-g\big(\frac{a+b}{2}\big)\Big)\Big\| \\ &= \Big\|g\big(\frac{a+b}{2}\big) - \Big(g(a)+g(b)-g\big(\frac{a+b}{2}\big)\Big)\Big\| \\ &= 2\|g\big(\frac{a+b}{2}\big) - g(a) + \frac{a-b}{2}\|\end{aligned}$$

gilt. In dem wir die Konstruktion von $\tilde{g}$ aus g iterieren, kämen wir zu einem Widerspruch zu (2). Somit ist die Behauptung bewiesen.

Zurück zur Isometrie f, die $0 \in V$ auf $0 \in V'$ abbildet, so definieren wir $\tilde{f}\colon V \to V$ durch

$$\tilde{f}(v) := f^{-1}\big(f(a)+f(b)\big) - f^{-1}\big(f(a)+f(b)-f(v)\big).$$

Dies ist eine Isometrie, die 0 auf 0 abbildet und die Bedingung (3) erfüllt. Deshalb gilt

$$\tilde{f}\big(\frac{a+b}{2}\big) = \frac{\tilde{f}(a)+\tilde{f}(b)}{2}.$$

Wir ziehen mit leichter algebraischer Manipulation die Konsequenz:

$$f\big(\frac{a+b}{2}\big) = \frac{f(a)+f(b)}{2}.$$

Daraus folgt: $f(2a) = 2f(a)$, $f(a+b) = f(a)+f(b)$, $f(ka) = kf(a)$ für alle $k \in \mathbb{Z}$ und weil f stetig ist, $f(\lambda a) = \lambda f(a)$ für alle $\lambda \in \mathbb{R}$. Als Quelle für diesen Beweis dient der Artikel „A proof of the Mazur-Ulam theorem“ von J. Väisälä, *American Mathematical Monthly* Band 110 (2003), S. 633–635.

Für einen euklidischen Vektorraum V haben wir die **orthogonale Gruppe**

$$O(V) := \{f \in GL(V) \mid f \text{ ist eine Isometrie } V \to V\}.$$

Dies ist Untergruppe der Gruppe $GL(V)$ aller invertierbaren linearen Abbildungen $V \to V$. Angenommen, V ist endlichdimensional, $n := \dim(V)$. Es gibt also eine lineare Isometrie nach $\mathbb{R}^n$ mit der euklidischen Metrik. Dadurch wird von einem $f \in O(V)$ ein Element von

$$O(n) := \{A \in GL_n(\mathbb{R}) \mid \mathbf{x} \mapsto A\mathbf{x} \text{ ist eine Isometrie } \mathbb{R}^n \to \mathbb{R}^n\}$$

bestimmt, wobei wir $\mathbb{R}^n$ stets mit der euklidischen Metrik betrachten. Ein $A \in O(n)$ erfüllt die Bedingung $\langle A\mathbf{x}, A\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$ für alle $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$, die nach den Rechenregeln für das Standardskalarprodukt so geschrieben werden kann:

$${}^t\mathbf{x}\,{}^tA\,A\mathbf{y} = {}^t\mathbf{x}\,\mathbf{y}$$

für alle $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$. So folgern wir

$$O(n) = \{A \in GL_n(\mathbb{R}) \mid {}^tA\,A = E_n\}.$$

Als Konsequenz:

Die Determinante einer orthogonalen Matrix ist stets ± 1.

Mit dem Adjektiv *speziell* werden Matrizen der Determinante 1 angedeutet. Es gibt zum Beispiel die **spezielle lineare Gruppe** eines endlichdimensionalen Vektorraums V über einem Körper K bzw. des Vektorraums K^n:

$$SL(V) := \{f \in GL(V) \mid \det(f) = 1\}, \qquad SL_n(K) := \{A \in GL_n(K) \mid \det(A) = 1\}.$$

Wir definieren auch die **spezielle orthogonale Gruppe** eines endlichdimensionalen euklidischen Vektorraums V bzw. des Vektorraums $\mathbb{R}^n$ mit Standardskalarprodukt:

$$SO(V) := \{f \in O(V) \mid \det(f) = 1\}, \qquad SO(n) := \{A \in O(n) \mid \det(A) = 1\}.$$

Proposition 2.8. *Sei V ein endlichdimensionaler euklidischer Vektorraum und $f \in O(V)$. Dann gibt es eine Orthonormalbasis von V, bezüglich derer f durch eine Ma-*

trix der form

$$
\begin{array}{l}
\qquad \overbrace{\hphantom{1 \quad \ddots \quad 1}}^{r} \quad \overbrace{\hphantom{-1 \quad \ddots \quad -1}}^{s} \\
\begin{pmatrix}
1 \\
& \ddots \\
& & 1 \\
& & & -1 \\
& & & & \ddots \\
& & & & & -1 \\
& & & & & & \cos(\theta_1) & -\sin(\theta_1) \\
& & & & & & \sin(\theta_1) & \cos(\theta_1) \\
& & & & & & & & \ddots \\
& & & & & & & & & \cos(\theta_t) & -\sin(\theta_t) \\
& & & & & & & & & \sin(\theta_t) & \cos(\theta_t)
\end{pmatrix}
\end{array}
$$

für $r,\ s,\ t \in \mathbb{N}$ und reelle Zahlen $\theta_1,\ \ldots,\ \theta_t \in (0, \pi)$ dargestellt wird.

Beweis. Wir beweisen die Proposition durch Indukion nach $n := \dim(V)$. Der Fall $n = 0$ ist trivial.

Falls $n > 0$ betrachten wir zwei Fälle. Im ersten Fall gibt es einen Eigenwert $\lambda \in \mathbb{R}$ von f. Sei $v_1 \in V$ ein Eigenvektor, für den wir ohne Beschränkung der Allgemeinheit annehmen können: $\|v_1\| = 1$. Es gilt $1 = \langle f(v), f(v)\rangle = \langle \lambda v, \lambda v\rangle = \lambda^2$, d.h., $\lambda = \pm 1$. Es gilt $f(v_1^\perp) \subset v_1^\perp$, deshalb gehört die Einschranküng von f auf $v_1^\perp$ zur Gruppe $O(v_1^\perp)$. Nach der Induktionshypothese gibt es eine geeignete Basis $v_2, \ldots, v_n$ von $v_1^\perp$ und wir haben nach Korollar 2.5 zusammen mit v_1 eine geeignete Basis von V.

Im zweiten Fall hat f keinen Eigenwert. Das charakteristische Polynom von f hat also irreduzible quadratische Faktoren und nach dem Satz von Cayley-Hamilton gibt es einen von Null verschieden Vektor $v_1 \in V$ und einen irreduziblen Faktor $T^2 + aT + b$ des charakteristischen Polynoms, so dass

$$f(f(v_1)) + af(v_1) + bv_1 = 0.$$

Ohne Beschränkung der Allgemeinheit gilt $\|v_1\| = 1$. Der Untervektorraum $W := \mathrm{span}(v_1, f(v_1))$ hat Dimension 2 und nach dem Gram-Schmidsches Verfahren gibt es ein $v_2 \in V$, so dass v_1, v_2 eine Orthonormalbasis von W ist. Bezüglich v_1, v_2 wird die Einschränkung von f auf W durch eine 2×2-Drehungsmatrix dargestellt und wenn wir falls nötig v_1 und v_2 vertauschen kommen wir zu einem 2×2-Block wie in der Aussage. Wie im ersten Fall geht es weiter mit der Induktionshypothese. □

Für den Fall $n = 3$ gibt es eine weitere Struktur, das **Vektorprodukt** $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$,

$$(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} \times \mathbf{y} := (x_2 y_3 - x_3 y_2, -x_1 y_3 + x_3 y_1, x_1 y_2 - x_2 y_1).$$

Proposition 2.9. *Das Vektorprodukt* $\mathbf{z} := \mathbf{x} \times \mathbf{y}$ *zweier Vektoren* $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^3$ *erfüllt:*
(i) $\mathbf{z} \perp \mathbf{x}$ *und* $\mathbf{z} \perp \mathbf{y}$.
(ii) $\|\mathbf{z}\|^2 = \|\mathbf{x}\|^2\|\mathbf{y}\|^2 - \langle \mathbf{x}, \mathbf{y}\rangle^2$.
(iii) $\mathbf{z} = 0$ *genau dann, wenn* $\mathbf{x}$ *und* $\mathbf{y}$ *linear abhängig sind.*
(iv) *Sind* $\mathbf{x}$ *und* $\mathbf{y}$ *von Null verschieden, so gilt* $\|\mathbf{z}\| = \|\mathbf{x}\|\,\|\mathbf{y}\| \sin\angle(\mathbf{x}, \mathbf{y})$.
(v) (Spatprodukt) *Für* $\mathbf{w} \in \mathbb{R}^3$ *und die Matrix* A *mit Zeilenvektoren* $\mathbf{x}$, $\mathbf{y}$, $\mathbf{w}$ *gilt*

$$\langle \mathbf{z}, \mathbf{w}\rangle = \det(A).$$

(vi) (Antikommutativität) $\mathbf{y} \times \mathbf{x} = -\mathbf{x} \times \mathbf{y}$.
(vi′) (Grassmann-Identität) $\mathbf{w} \times (\mathbf{x} \times \mathbf{y}) = \langle \mathbf{w}, \mathbf{y}\rangle\mathbf{x} - \langle \mathbf{w}, \mathbf{x}\rangle\mathbf{y}$.
(vi″) (Jacobi-Identität) $\mathbf{w} \times (\mathbf{x} \times \mathbf{y}) + \mathbf{x} \times (\mathbf{y} \times \mathbf{w}) + \mathbf{y} \times (\mathbf{w} \times \mathbf{x}) = 0$.

Beweis. Wir verifizieren (i) und (ii) direkt. In (ii) ist der Ausdruck rechts genau dann gleich Null, wenn $\mathbf{x}$ und $\mathbf{y}$ linear abhängig sind. Daraus folgt (iii). In (iv) gilt $\sin\angle(\mathbf{x}, \mathbf{y}) \geq 0$, deshalb genügt es, die Identität $\|\mathbf{z}\|^2 = \|\mathbf{x}\|^2\|\mathbf{y}\|^2 \sin^2\angle(\mathbf{x}, \mathbf{y})$ zu verifizieren. Dies folgt aus (ii), der Identität $\sin^2\angle(\mathbf{x}, \mathbf{y}) + \cos^2\angle(\mathbf{x}, \mathbf{y}) = 1$ und der Definition von $\angle(\mathbf{x}, \mathbf{y})$. Die anderen Identitäten können wir direkt verifizieren. □

2.3 Unitäre Vektorräume

Auch über einem komplexen Vektorraum gibt es eine geeignete Struktur, die zu einer Norm und einer Metrik führt. Die Definition einer Norm ist für einen komplexen Vektorraum V genau dieselbe, wie für einen reellen Vektorraum, nur muss die Formel $\|\lambda v\| = |\lambda|\,\|v\|$ für alle $v \in V$ und $\lambda \in \mathbb{C}$ gelten. Nach wie vor führt eine Norm auf V zu einer Metrik auf V. Wir erinnern uns: fur $\lambda = s + ti$ mit s, $t \in \mathbb{R}$ und $i = \sqrt{-1}$ gilt

$$|\lambda| = \left|\overline{\lambda}\right| = \sqrt{\lambda\,\overline{\lambda}} = \sqrt{s^2 + t^2}.$$

Definition. Eine **Sesquilinearform** auf einem komplexen Vektorraum V ist eine Abbildung $s\colon V \times V \to \mathbb{C}$, die im ersten Faktor linear ist

$$s(v + v', w) = s(v, w) + s(v', w), \qquad s(\lambda v, w) = \lambda s(v, w)$$

für alle v, v', $w \in V$ und $\lambda \in \mathbb{C}$, aber im zweiten Faktor die folgenden modifizierten Bedingungen erfüllen:

$$s(v, w + w') = s(v, w) + s(v, w'), \qquad s(v, \lambda w) = \overline{\lambda} s(v, w)$$

für alle v, w, $w' \in V$ und $\lambda \in \mathbb{C}$. Eine Sesquilinearform heisst **hermitesch** falls

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

für alle v, $w \in V$. Eine hermitesche Sesquilinearform heisst **positiv definit**, falls $s(v, v) > 0$ für alle $0 \neq v \in V$. Eine positiv definite hermitesche Sesquilinearform heisst auch **Skalarprodukt**. Ein **unitärer Vektorraum** ist ein komplexer Vektorraum mit Skalarprodukt.

Für eine hermitesche Sesquilinearform $s\colon V \times V \to \mathbb{C}$ gibt es auch eine Polarisationsformel:

$$s(v,w) = \frac{s(v+w,v+w) - s(v-w,v-w) + is(v+iw,v+iw) - is(v-iw,v-iw)}{4}.$$

Wie im Fall eines euklidischen Vektorraums schreiben wir für einen unitären Vektorraum $\langle \cdot, \cdot \rangle$ für das Skalarprodukt und

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

Proposition 2.10 (Cauchy-Schwarzsche Ungleichung für unitäre Vektorräume)**.** *Ist V ein unitärer Vektorraum, so gilt*

$$|\langle v, w \rangle| \le \|v\| \, \|w\|$$

für alle v, $w \in V$, mit Gleichheit genau dann, wenn v und w linear abhängig sind.

Beweis. Falls v und w linear abhängig sind, gilt $v = \lambda w$ oder $w = \lambda v$ für ein $\lambda \in \mathbb{C}$ und somit $|\langle v, w \rangle| = |\lambda| \, \|w\|^2 = \|v\| \, \|w\|$ bzw. $|\langle v, w \rangle| = |\lambda| \, \|v\|^2 = \|v\| \, \|w\|$.

Angenommen, v und w sind linear unabhängig. Es gilt

$$0 < \langle v, v \rangle \Big\langle w - \frac{\langle w, v \rangle}{\langle v, v \rangle} v, w - \frac{\langle w, v \rangle}{\langle v, v \rangle} v \Big\rangle = \langle v, v \rangle \langle w, w \rangle - |\langle v, w \rangle|^2$$

und es folgt, $|\langle v, w \rangle| < \|v\| \, \|w\|$. □

Korollar 2.11. *Auf einem unitären Vektorraum V ist $\|\cdot\|\colon V \to \mathbb{R}_{\ge 0}$ eine Norm.*

Beweis. Es gilt

$$\|v+w\|^2 = \|v\|^2 + 2\,\mathrm{Re}\langle v, w \rangle + \|w\|^2 \le \|v\|^2 + 2|\langle v, w \rangle| + \|w\|^2 \le (\|v\| + \|w\|)^2$$

für alle v, $w \in V$. □

Definition. Das **Standardskalarprodukt** auf $\mathbb{C}^n$ ist

$$\langle \mathbf{z}, \mathbf{w} \rangle := z_1 \overline{w}_1 + \cdots + z_n \overline{w}_n.$$

Sei $\mathbf{z} = (z_1, \ldots, z_n) \in \mathbb{C}^n$ mit

$$z_1 = x_1 + y_1 i, \qquad \ldots, \qquad z_n = x_n + y_n i$$

und x_1, y_1, $\ldots$, x_n, $y_n \in \mathbb{R}$. Dann gilt für die zum Standardskalarprodukt assoziierte Norm

$$\|\mathbf{z}\| = \sqrt{x_1^2 + y_1^2 + \cdots + x_n^2 + y_n^2}.$$

So bestimmt tatsächlich das Standardskalarprodukt die euklidische Norm und euklidische Metrik auf $\mathbb{C}^n$, wobei $\mathbb{C}^n$ durch $(z_1, \ldots, z_n) \mapsto (x_1, y_1, \ldots, x_n, y_n)$ mit $\mathbb{R}^{2n}$ identifiziert wird.

Es gilt

$$\langle \mathbf{z}, \mathbf{w} \rangle = {}^t\mathbf{z}\,\overline{\mathbf{w}}$$

sowie für $A \in \mathrm{Mat}(n \times n, \mathbb{C})$:

$$\langle A\mathbf{z}, \mathbf{w} \rangle = {}^t\mathbf{z}\,{}^tA\overline{\mathbf{w}} \qquad \text{und} \qquad \langle \mathbf{z}, A\mathbf{w} \rangle = {}^t\mathbf{z}\,\overline{A}\overline{\mathbf{w}}.$$

Die Begriffe orthogonal, orthogonales Komplement und orthonormal sind für einen unitären Vektorraum wie im Fall eines euklidischen Vektorraums definiert. Nach wie vor ist eine aus von Null verschiedenen Vektoren bestehende orthogonale Familie stets linear unabhängig.

Proposition 2.12 (Gram-Schmidtsches Verfahren für unitäre Vektorräume)**.** *Sei V ein endlichdimensionaler unitärer Vektorraum und $v_1, \ldots, v_m$ eine orthonormale Familie von Vektoren aus V. Dann gibt es eine Erweiterung zu einer Orthonormalbasis $v_1, \ldots, v_n$ von V.*

Beweis. Der Beweis von Proposition 2.4 ist auch für unitäre Vektorräume gültig. □

Korollar 2.13. *Sei V ein endlichdimensionaler unitärer Vektorraum und $W \subset V$ ein Untervektorraum. Dann gilt*

$$V = W \oplus W^\perp.$$

Beweis. Der Beweis von Korollar 2.5 ist ebenso für unitäre Vektorräume gültig. □

Wie im Fall eines euklidischen Vektorraums wird die Rolle einer linearen Isometrie eines unitären Vektorraums hervorgehoben. So definieren wir die **unitäre Gruppe** eines unitären Vektorraums V:

$$U(V) := \{f \in GL(V) \,|\, f \text{ ist eine Isometrie } V \to V\}.$$

Angenommen, V ist endlichdimensional, $n := \dim(V)$. Es gibt also eine lineare Isometrie nach $\mathbb{C}^n$ mit der euklidischen Metrik. Dadurch wird von einem $f \in U(V)$ ein Element von

$$U(n) := \{A \in GL_n(\mathbb{C}) \,|\, \mathbf{z} \mapsto A\mathbf{z} \text{ ist eine Isometrie } \mathbb{C}^n \to \mathbb{C}^n\}$$

bestimmt, wobei wir $\mathbb{C}^n$ stets mit der euklidischen Metrik betrachten. Also haben wir

$${}^t\mathbf{z}\,{}^tA\,\overline{A}\overline{\mathbf{w}} = {}^t\mathbf{z}\,\overline{\mathbf{w}}$$

für alle $\mathbf{z}$, $\mathbf{w} \in \mathbb{C}$, und deshalb

$$U(n) = \{A \in GL_n(\mathbb{C}) \,|\, {}^t\overline{A}\, A = E_n.\}$$

Die Determinante einer unitären $n \times n$-Matrix bzw. einer linearen Isometrie eines endlichdimensionalen unitären Vektorraums V ist dehalb stets eine komplexe Zahl vom Absolutbetrag 1. Es gibt auch die **spezielle unitäre Gruppe**

$$SU(V) := \{f \in U(V) \,|\, \det(f) = 1\}, \qquad SU(n) := \{A \in U(n) \,|\, \det(A) = 1\}.$$

Proposition 2.14. *Sei V ein endlichdimensionaler unitärer Vektorraum und $f \in U(V)$. Dann gibt es eine aus Eigenvektoren von f bestehende Orthonormalbasis von V. Die Eigenwerte sind komplexe Zahlen vom Absolutbetrag* 1.

Beweis. Wir beweisen die Proposition durch Induktion nach $n := \dim(V)$. Der Fall $n = 0$ ist trivial. Sei $n > 0$, sei $\lambda \in \mathbb{C}$ ein Eigenwert mit Eigenvektor $v_1 \in V$, $\|v_1\| = 1$. Es gilt $1 = \langle f(v), f(v)\rangle = \langle \lambda v, \lambda v\rangle = \lambda\overline{\lambda} = |\lambda|^2$. Da $f(v_1^\perp) \subset v_1^\perp$ gehört die Einschränkung von f auf $v_1^\perp$ zur Gruppe $U(v_1^\perp)$. Nach der Induktionshypothese gibt es eine Orthonormalbasis $v_2, \ldots, v_n$ von $v_1^\perp$, die die Einschränkung von f auf $v_1^\perp$ diagonalisiert. Nach Korollar 2.13 ist $v_1, v_2, \ldots, v_n$ eine Orthonormalbasis von V. □

2.4 Bilinearformen und Matrizen

Mit der Auswahl einer Basis $v_1, \ldots, v_n$ eines endlichdimensionalen Vektorraums V über einem Körper K wird eine Bilinearform auf V durch eine $n \times n$-Matrix über K dargestellt.

Proposition 2.15. *Sei V ein endlichdimensionaler Vektorraum über einem Körper K und $v_1, \ldots, v_n$ eine Basis von V. Dann wird durch $s \mapsto (s(v_i, v_j))_{1 \le i,j \le n}$ eine bijektive Abbildung*

$$\{\textit{Bilinearformen } s\colon V \times V \to K\} \to \mathrm{Mat}(n \times n, K)$$

definiert. Des Weiteren gilt für die zu s entsprechende Matrix $A \in \mathrm{Mat}(n \times n, K)$:

s ist genau dann …	*wenn …*
symmetrisch	$A = {}^tA$
schiefsymmetrisch	$A = -{}^tA$
alternierend	$A = -{}^tA$ *und die Diagonaleinträge von A sind Null*

Im Fall $K = \mathbb{C}$ wird durch dieselbe Formel eine bijektive Abbildung

$$\{\textit{Sesquilinearformen } s\colon V \times V \to \mathbb{C}\} \to \mathrm{Mat}(n \times n, \mathbb{C})$$

definiert, für die s genau dann hermitesch ist, wenn die zu s entsprechende Matrix $A \in \mathrm{Mat}(n \times n, K)$ die Bedingung $A = {}^t\overline{A}$ erfüllt.

Beweis. Sei $s\colon V\times V \to K$ eine Bilinearform. Weil s bilinear ist, bestimmt die Matrix $(s(v_i,v_j))_{1\le i,j\le n}$ die Bilinearform s vollständig. Deshalb ist die Abbildung injektiv. Für eine beliebige Matrix $A \in \mathrm{Mat}(n\times n, K)$ wird mit

$$(b_1v_1+\dots+b_nv_n, c_1v_1+\dots+c_nv_n) \mapsto {}^t\mathbf{b}\,A\,\mathbf{c}$$

für $\mathbf{b}=(b_1,\dots,b_n)$, $\mathbf{c}=(c_1,\dots,c_n)\in K^n$ eine Bilinearform $s\colon V\times V\to K$ definiert, für die $(s(v_i,v_j))_{1\le i,j\le n} = A$ gilt. Deshalb ist die Abbildung auch surjektiv. Ist s symmetrisch, so gilt $s(v_j,v_i)=s(v_i,v_j)$ für alle i, j, d.h., die entsprechende Matrix ist symmetrisch. Ist A hingegen eine symmetrische Matrix, so gilt

$$s(c_1v_1+\dots+c_nv_n, b_1v_1+\dots+b_nv_n) = {}^t\mathbf{c}\,A\,\mathbf{b} = {}^t\mathbf{b}\,{}^tA\,\mathbf{c} = {}^t\mathbf{b}\,A\,\mathbf{c}$$

und deshalb ist s symmetrisch. Ähnlich argumentieren wir, dass s genau dann schiefsymmetrisch bzw. alternierend ist, wenn die angegebene Bedingung bezüglich der entsprechenden Matrix erfüllt ist.

Im Fall $K=\mathbb{C}$ ist das Argument mit leichter Modifikation gültig. Aus einem $A \in \mathrm{Mat}(n\times n,\mathbb{C})$ wird mit

$$(b_1v_1+\dots+b_nv_n, c_1v_1+\dots+c_nv_n) \mapsto {}^t\mathbf{b}\,A\,\overline{\mathbf{c}}$$

eine Sesquilinearform $s\colon V\times V\to K$ definiert, für die $(s(v_i,v_j))_{1\le i,j\le n}=A$ gilt. Ist s hermitesch, so gilt $s(v_j,v_i)=\overline{s(v_i,v_j)}$ für alle i, j, d.h., die entsprechende Matrix ist hermitesch. Ist A hingegen hermitesch, so gilt

$$s(c_1v_1+\dots+c_nv_n, b_1v_1+\dots+b_nv_n) = {}^t\mathbf{c}\,A\,\overline{\mathbf{b}} = {}^t\overline{\mathbf{b}}\,{}^tA\,\mathbf{c}$$

und weil A hermitesch ist, d.h., ${}^t\overline{A}=A$, gilt ${}^t\overline{\mathbf{b}}\,{}^tA\,\mathbf{c} = {}^t\overline{\mathbf{b}}\,\overline{A}\,\mathbf{c} = \overline{{}^t\mathbf{b}\,A\,\overline{\mathbf{c}}}$. $\square$

Sei V ein endlichdimensionaler Vektorraum über einem Körper K mit Basis v_1, …, v_n und $s\colon V\times V\to K$ eine Bilinearform mit darstellende Matrix A bezüglich der Basis v_1, …, v_n. Ist v'_1, …, v'_n auch eine Basis von V, haben wir auch eine darstellende Matrix A' bezüglich der Basis v'_1, …, v'_n und ein kommutatives Diagramm

$$\begin{tikzcd}
K^n \times K^n \arrow[dd, "(\mathbf{x}\mapsto U\mathbf{x})\times(\mathbf{x}\mapsto U\mathbf{x})"'] \arrow[dr, "(e_i\mapsto v_i)\times(e_i\mapsto v_i)"] \arrow[drr, "(\mathbf{b}{,}\mathbf{c})\mapsto {}^t\mathbf{b}\,A\,\mathbf{c}"] & & \\
& V \times V \arrow[r, "s"] & K \\
K^n \times K^n \arrow[ur, "(e_i\mapsto v'_i)\times(e_i\mapsto v'_i)"'] \arrow[urr, "(\mathbf{b}{,}\mathbf{c})\mapsto {}^t\mathbf{b}\,A'\,\mathbf{c}"'] & &
\end{tikzcd}$$

für ein $U\in GL_n(K)$. Die folgende Transformationsformel folgt:

$$A = {}^tU\,A'\,U.$$

Definition. Zwei quadratische Matrizen A, $A' \in \mathrm{Mat}(n \times n, K)$ heissen **kongruent**, wenn es ein $U \in GL_n(K)$ gibt, so dass $A = {}^tU\, A'\, U$.

Kongruenz ist eine Äquivalenzrelation auf $\mathrm{Mat}(n \times n, K)$. Unter Linksmultiplikation mit ${}^tU$ und Rechtsmultiplikation mit U für ein $U \in GL_n(K)$ bleibt der Rang unverändert, während die Determinante mit $\det(U)^2$ multipliziert wird.

Definition. Der **Rang** einer Bilinearform s auf einem endlichdimensionalen Vektorraum V über einem Körper K ist der Rang einer darstellenden Matrix. Die **Diskriminante** von s ist die Äquivalenzklasse der Determinante einer darstellenden Matrix, unter der Äquivalenzrelation $t \sim u^2 t$ für $t \in K$ und $u \in K^\times$.

Die Diskriminante ist deshalb genau dann Null, wenn der Rang von s kleiner als die Dimension von V ist. Falls $\mathrm{rk}(s) = \dim(V)$ ist die Diskriminante ein Element von

$$K^\times / (K^\times)^2,$$

d.h., ist ein von Null verschiedenes Element von K, definiert bis auf Multiplikation mit einem von Null verschiedenen Quadrat von K.

Im Fall einer Sesquilinearform über einem endlichdimensionalen komplexen Vektorraum sieht das kommutative Diagramm so aus

$$\begin{tikzcd}
\mathbb{C}^n \times \mathbb{C}^n \arrow[dd, "(\mathbf{x}\mapsto U\mathbf{x})\times(\mathbf{x}\mapsto U\mathbf{x})"'] \arrow[dr, "(e_i\mapsto v_i)\times(e_i\mapsto v_i)"] \arrow[drr, "(\mathbf{b},\mathbf{c})\mapsto {}^t\mathbf{b}\,A\,\overline{\mathbf{c}}"] & & \\
& V \times V \arrow[r, "s"] & \mathbb{C} \\
\mathbb{C}^n \times \mathbb{C}^n \arrow[ur, "(e_i\mapsto v_i')\times(e_i\mapsto v_i')"] \arrow[urr, "(\mathbf{b},\mathbf{c})\mapsto {}^t\mathbf{b}\,A'\,\overline{\mathbf{c}}"'] & &
\end{tikzcd}$$

und die Transformationsformel, so:

$$A = {}^tU\, A'\, \overline{U},$$

mit $U \in GL_n(\mathbb{C})$. Insbesondere ist der Rang einer Sesquilinearform wohldefiniert und es gibt eine Diskriminante, eine komplexe Zahl, die bis auf Multiplikation mit einer positiven reellen Zahl definiert ist.

Sei V ein Vektorraum über einem Körper K und $s\colon V \times V \to K$ eine Bilinearform. Die Menge aller Vektoren $v \in V$, mit der Eigenschaft $s(v, w) = 0$ für alle $w \in V$, ist ein Untervektorraum von V. Ebenso ist die Menge aller Vektoren $v \in V$, mit der Eigenschaft $s(u, v) = 0$ für alle $u \in W$. Im Allgemeinen können die beiden Untervektorräume voneinander verschieden sein, aber in wichtigen Fällen sind sie ein und derselbe Untervektorraum.

Proposition 2.16. *Sei V ein Vektorraum über einem Körper K und $s\colon V \times V \to K$ eine Bilinearform oder im Fall $K = \mathbb{C}$ eine Sesquilinearform. In den folgenden Fällen gilt für Vektoren v, $w \in V$, dass $s(v,w) = 0$ genau dann, wenn $s(w,v) = 0$:*

- *s ist eine symmetrische Bilinearform;*
- *s ist eine schiefsymmetrische Bilinearform;*
- *$K = \mathbb{C}$ und s ist eine hermitesche Sesquilinearform.*

Insbesondere gilt in jedem Fall die Gleichheit von Untervektorräumen:

$$\{v \in V \mid s(v,w) = 0 \text{ für alle } w \in V\} = \{v \in V \mid s(u,v) = 0 \text{ für alle } u \in V\}.$$

Beweis. Die Aussage ist klar in jedem Fall, z.B., falls s eine schiefsymmetrische Bilinearform ist: $s(v,w) = 0 \Leftrightarrow -s(v,w) = 0 \Leftrightarrow s(w,v) = 0$. □

In den in Proposition 2.16 geschilderten Fällen schreiben wir $v \perp w$, wenn $s(v,w) = 0$ oder äquivalent, $s(w,v) = 0$. Die Begriffe orthogonales Komplement einer Teilmenge und orthogonale Familie aus §2.2 gelten auch. Der gekennzeichnete Untervektorraum aus Proposition 2.16 heisst **Ausartungsraum** von s. Falls der Ausartungsraum von Null verschieden ist, sagen wir, s ist **ausgeartet**, und sonst **nicht ausgeartet**.

Proposition 2.17. *Sei V ein Vektorraum über einem Körper K und $s\colon V \times V \to K$ eine Bilinearform oder im Fall $K = \mathbb{C}$ eine Sesquilinearform. In jedem der in Proposition 2.16 geschilderten Fälle wird durch*

$$\bar{s}(\bar{v}, \bar{w}) := s(v,w)$$

eine nicht ausgeartete Bilinear- bzw. Sesquilinearform auf V/U definiert, wobei mit U der Ausartungsraum von s bezeichtet wird und der entsprechende Fall (symmetrische Bilinearform usw.) auch für $\bar{s}$ gilt. Zudem gilt falls $\operatorname{char}(K) = 2$: $\bar{s}$ ist genau dann alternierend, wenn s alternierend ist.

Beweis. In jedem Fall verfizieren wir, dass $\bar{s}$ wohldefiniert ist. Sei also $v' = v + u$ mit $u \in U$, dann gilt

$$s(v', w) = s(v,w) + s(u,w) = s(v,w)$$

und ähnlich gilt für $w' = w + u$ in jedem Fall $s(v,w') = s(v,w)$. Es ist leicht zu verifizieren, dass $\bar{s}$ eine symmetrische Bilinearform ist, wenn s eine symmetrische Bilinearform ist, uzw. Gehört ein $\bar{v}$ zum Ausartungsraum von $\bar{s}$, so gilt $s(v,w) = 0$ für alle $w \in V$, d.h., $v \in U$ und deshalb $\bar{v} = 0$. Es gilt $\bar{s}(\bar{v}, \bar{v}) = s(v,v)$, deshalb: $\bar{s}(\bar{v}, \bar{v}) = 0$ für alle $\bar{v} \in V/U$ genau dann, wenn $s(v,v) = 0$ für alle $v \in V$. □

Korollar 2.18. *In der Situation von Proposition 2.17 gilt für ein bielibiges Komplement $W \subset V$ zu U, dass V mit Bilinear- bzw. Sesquilinearform s das orthogonale direkte Summe von U mit trivialer Bilinear- bzw. Sesquilinearform und W mit nicht ausgearteter Einschräunkung von s ist.*

Beweis. Sei $u \in U$. Es gilt $s(u, v) = 0$ für alle $v \in V$, deshalb ist $V = U \oplus W$ ein orthogonales direktes Summe. Die Komposition

$$W \to V \to V/U$$

ist ein Isomorphismus von Vektorräumen und es gilt $\bar{s}(\bar{v}, \bar{w}) = s(v, w)$ für alle v, $w \in W$. Die Einschränkung von s auf W ist also nicht ausgeartet. □

Proposition 2.19. *Sei V ein endlichdimensionaler Vektorraum über einem Körper K und $s\colon V \times V \to K$ eine Bilinearform oder im Fall $K = \mathbb{C}$ eine Sesquilinearform. In jedem der in Proposition 2.16 geschilderten Fälle gilt für einen beliebigen Untervektorraum $W \subset V$:*

- *Ist s nicht ausgeartet, so gilt* $\dim(W) + \dim(W^\perp) = \dim(V)$.
- *Ist zudem die Einschränkung von s auf W nicht ausgeartet, so gilt* $V = W \oplus W^\perp$.

Beweis. Sei $d := \dim(W)$ und $w_1, \dots, w_d$ eine Basis von W. Dann haben wir mit $v \mapsto (s(v, w_1), \dots, s(v, w_d))$ eine lineare Abbildung

$$V \to K^d.$$

Wir behaupten, dies ist eine surjektive lineare Abbildung; weil der Kern $W^\perp$ ist, wird die Aussage über Dimension folgen. Falls dies nicht surjective ist, gibt es ein $0 \neq (a_1, \dots, a_d) \in K^d$, so dass die Komposition

$$V \to K^d \to K$$

mit $(x_1, \dots, x_d) \mapsto a_1x_1 + \dots + a_dx_d$ Null ist. Die Formel für die Komposition ist $v \mapsto s(v, a_1w_1 + \dots + a_dw_d)$ falls s bilinear ist und $v \mapsto s(v, \overline{a}_1w_1 + \dots + \overline{a}_dw_d)$ falls $K = \mathbb{C}$ und s sesquilinear ist. In jedem Fall gibt es dann einen von Null verschiedenen Vektor im Ausartungsraum von s und wir haben einen Widerspruch.

Die Ausartungsraum der Einschränkung von s auf W ist $W \cap W^\perp$. Ist dies Null, so folgt aus der Aussage über Dimension, dass V direkte Summe von W und $W^\perp$ ist. □

In der Situation von Proposition 2.19 können wir eine Basis $v_1, \dots, v_n$ von V auswählen und s durch eine Matrix $A \in \mathrm{Mat}(n \times n, K)$ darstellen. Dann entspricht der Ausartungsraum von s dem Kern von A, also gilt:

$$s \text{ ist ausgeartet} \Leftrightarrow \mathrm{rk}(s) < \dim(V).$$

Ist W der von $v_1, \ldots, v_d$ aufgespannte Untervektorraum für ein $d \leq n$, so gilt ähnlich: die Einschränkung von s auf W ist genau dann ausgeartet, wenn der Rang des linken oberen $d \times d$-Blocks von A kleiner als d ist.

Beispiel. Wir bestimmen eine Normalform für alternierende Billinearformen auf endlichdimensionalen Vektorräumen. Sei K ein Körper, V ein endlichdimensionaler Vektorraum, $n := \dim(V)$ und $s\colon V \times V \to K$ eine alternierende Bilinearform. Dann behaupten wir: $r := \mathrm{rk}(s)$ ist gerade und es gibt eine Basis $v_1, \ldots, v_n$ von V, bezüglich derer s durch die Matrix

$$\begin{pmatrix} 0 & 1 & & & & & \cdots & 0 \\ -1 & 0 & & & & & & \\ & & \ddots & & & & & \vdots \\ & & & 0 & 1 & & & \\ & & & -1 & 0 & & & \\ \vdots & & & & & 0 & & \\ & & & & & & \ddots & \\ 0 & & \cdots & & & & & 0 \end{pmatrix}$$

mit $r/2$ 2×2-Blöcken dargestellt wird. Es genügt nach Korollar 2.18, zu zeigen: ist s eine nicht ausgeartete alternierende Bilinearform auf V, so ist n gerade und s bezüglich einer geeigneten Basis durch eine Blockdiagonalmatrix mit $n/2$ 2×2-Blöcken wie oben dargestellt. Diese Aussage beweisen wir durch Induktion nach n. Die Fälle $n = 0$ (trivial) und $n = 1$ (klar) gelten als Induktionsanfang. Angenommen, $n \geq 2$. Sei $0 \neq v_1 \in V$. Da s nicht ausgeartet ist, gibt es ein $v_2 \in V$ mit $s(v_1, v_2) \neq 0$. In dem wir v_2 durch ein Vielfaches ersetzen können wir annehmen, $s(v_1, v_2) = 1$. Sei $W := \mathrm{span}(v_1, v_2)$, dann haben wir: $\dim(W) = 2$, die Einschränkung von s auf W ist nicht ausgeartet und deshalb (Proposition 2.19) $V = W \oplus W^\perp$. Die Einschränkung von s auf $W^\perp$ ist nicht ausgeartet und nach der Induktionshypothese ist die Dimension $n - 2$ von $W^\perp$ gerade und die Einschränkung von s auf $W^\perp$ bezüglich einer geeigneten Basis durch eine Blockdiagonalmatrix mit $(n-2)/2$ 2×2-Blöcken wie oben dargestellt.

Sei K ein Körper, V ein Vektorraum über K und $s\colon V \times V \to K$ eine symmetrische Bilinearform. Wir definieren die assoziierte **quadratische Form** $q\colon V \to K$ durch

$$q(v) := s(v, v).$$

Im Allgemeinen nennt man eine Abbildung $q\colon V \to K$ quadratische Form, falls gilt:

$$q(\lambda v) = \lambda^2 q(v)\,\forall\,\lambda \in K,\, v \in V \quad \text{und} \quad (v, w) \mapsto q(v + w) - q(v) - q(w) \text{ ist bilinear.}$$

Falls $\mathrm{char}(K) \neq 2$ können wir durch die Polarisationsformel eine symmetrische Bilinearform $s\colon V \times V \to K$ aus einer quadratischen Form gewinnen und wir haben eine bijektive Zuordnung zwischen symmetrischen Bilinearformen und quadratischen Formen auf V.

Falls $K = \mathbb{R}$ gibt es für eine symmetrische Bilinearformen $s\colon V \times V \to K$ den Begriff positiv definite, d.h., $s(v,v) > 0$ für alle $0 \neq v \in V$, oder bezüglich der assoziierten quadratischen Form, $q(v) > 0$ für alle $0 \neq v \in V$. Es gibt auch ähnliche Bedingungen:

negativ definit	$s(v,v) < 0$ für alle $0 \neq v \in V \Leftrightarrow q(v) < 0$ für alle $0 \neq v \in V$,
positiv semidefinit	$s(v,v) \geq 0$ für alle $v \in V \Leftrightarrow q(v) \geq 0$ für alle $v \in V$,
negativ semidefinit	$s(v,v) \leq 0$ für alle $v \in V \Leftrightarrow q(v) \leq 0$ für alle $v \in V$.

Proposition 2.20. *Sei K ein Körper mit $\mathrm{char}(K) \neq 2$, V ein endlichdimensionaler Vektorraum über K und $s\colon V \times V \to K$ eine symmetrische Bilinearform. Dann gibt es eine Basis von V, bezüglich derer s durch eine Diagonalmatrix dargestellt wird.*

Beweis. Es genügt nach Korollar 2.18, die Aussage zu beweisen falls zudem s nicht ausgeartet ist. Also nehmen wir so an und führen den Beweis durch Induktion nach $n := \dim(V)$ aus. Die Aussage für $n = 0$ ist trivial. Ist $n > 0$, so gibt es nach der Polarisationsformel ein $v_1 \in V$ mit $s(v_1, v_1) \neq 0$. Dann gilt nach Proposition 2.19

$$V = \mathrm{span}(v_1) \oplus v_1^{\perp}.$$

Die Einschränkung von s auf $v_1^{\perp}$ ist nicht ausgeartet. Nach der Induktionshypothese gibe es eine Basis $v_2, \ldots, v_n$ von $v_1^{\perp}$, bezüglich derer die Einschränkung von s durch eine Diagonalmatrix dargestellt wird. So kommen wir zur geeigneten Basis $v_1, v_2, \ldots, v_n$ von V. $\square$

Proposition 2.21 (Sylvesterscher Trägheitssatz)**.** *Sei V ein endlichdimensionaler reeller Vektorraum und $s\colon V \times V \to \mathbb{R}$ eine symmetrische Bilinearform. Dann gibt es eindeutige natürliche Zahlen n_+ und n_-, so dass für eine geeignete Basis von V die Bilinearform s durch die Diagonalmatrix*

$$\begin{pmatrix} \overbrace{1 \quad \ddots \quad 1}^{n_+} & & & \\ & \overbrace{-1 \quad \ddots \quad -1}^{n_-} & & \\ & & 0 & \\ & & & \ddots \\ & & & & 0 \end{pmatrix}$$

dargestellt wird. Die Zahlen n_+ und n_- lassen sich charakterisieren als grösste Dimension eines Untervektorraums $U \subset V$, auf dem die Einschränkung von s positiv definit bzw. negativ definit ist.

Beweis. Nach wie vor genügt es, die Aussage zu beweisen falls s nicht ausgeartet ist. Sei $n := \dim(V)$. Die Existenz einer solchen Basis folgt aus Proposition 2.20, denn wir können für eine Basis $v_1, \ldots, v_n \in V$ ein v_i durch λv_i für $\lambda \in \mathbb{R}^\times$ und so den entsprechenden Diagonaleintrag mit λ^2 multiplizieren. So wird jeweils aus einem positiven Diagonaleintrag 1 entstehen, aus einem negativen Diagonaleintrag -1. Es gibt dann offensichtlich Untervektorräume $U_+ \subset V$ und $U_- \subset V$ mit $\dim(U_+) = n_+$ und $\dim(U_-) = n_-$, so dass die Einschränkung von s auf U_+ positiv definit bzw. auf U_- negativ definit ist. Ist $U' \subset V$ ein Untervektorraum, auf dem die Einschränkung von s positiv definit ist, so gilt $U' \cap U_- = 0$. Es folgt, $\dim(U') + \dim(U_-) \le n$ und deshalb $\dim(U') \le n - n_- = n_+$. Ähnlich gilt $\dim(U') \le n_-$ für alle Untervektorräume $U' \subset V$, auf denen die Einschränkung von s negativ definit ist. □

In der Situation von Proposition 2.21 ist $n_0 := \dim(V) - n_+ - n_-$ die Dimension des Ausartungsraums von s. Die symmetrische Bilinearform s auf V ist klassifiziert durch die **Signatur** (n_+, n_-, n_0) im Sinn, dass eine darstellende Matrix von s stets zur Kongruenzklasse der angezeigten Matrix gehört und n_+ und n_- eindeutig sind mit dieser Eigenschaft. Ist $n_0 = 0$ (d.h., s nicht ausgeartet), so wird die Signatur auch als (n_+, n_-) angegeben.

2.5 Selbstadjungierte Endomorphismen

Symmetrische bzw. hermitesche Matrizen lassen sich auf koordinatenfreie Weise betrachten, in dem wir auf einem endlichdimensionalen euklidischen bzw. unitäre Vektorraum V diejenigen Endomorphismen $f\colon V \to V$ herausheben, die die Bedingung $\langle f(v), w\rangle = \langle v, f(w)\rangle$ für alle v, $w \in V$ erfüllen. Im Allgemeinen gibt es zu jedem Endomorphismus $f\colon V \to V$ einen eindeutig bestimmten **adjungierten Endomorphismus** $f^*\colon V \to V$ mit der Eigenschaft

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

für alle v, $w \in V$. Denn wir können durch die Auswahl einer Orthonormalbasis auf die analoge Eigenschaft für Matrizen reduzieren. Für $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ suchen wir $A^* \in \mathrm{Mat}(n \times n, \mathbb{R})$ mit der Eigenschaft $\langle A^*\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, A\mathbf{y}\rangle$ für alle $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$. Tatsächlich ist dies mit $A^* = {}^tA$ gültig:

$$\langle {}^tA\,\mathbf{x}, \mathbf{y}\rangle = {}^t\mathbf{x}\,A\,\mathbf{y} = \langle \mathbf{x}, A\mathbf{y}\rangle.$$

Ähnlich geht es für $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ mit $A^* = {}^t\overline{A}$, denn für alle $\mathbf{z}$, $\mathbf{w} \in \mathbb{C}^n$ gilt

$$\langle {}^t\overline{A}\,\mathbf{z}, \mathbf{w}\rangle = {}^t\mathbf{z}\,\overline{A}\,\overline{\mathbf{w}} = \langle \mathbf{z}, A\mathbf{w}\rangle.$$

Es gilt stets $(f^*)^* = f$.

Ein Endomorphismus $f\colon V \to V$ heisst **selbstadjungiert**, wenn $f^* = f$. Für Matrizen:

$$A \in \mathrm{Mat}(n \times n, \mathbb{R}) \text{ selbstadjungiert} \Leftrightarrow A \text{ symmetrisch},$$
$$A \in \mathrm{Mat}(n \times n, \mathbb{C}) \text{ selbstadjungiert} \Leftrightarrow A \text{ hermitesch}.$$

Proposition 2.22 (Spektralsatz). *Sei V ein endlichdimensionaler euklidischer oder unitärer Vektorraum und $f\colon V \to V$ ein selbstadjungierter Endomorphismus. Dann gibt es eine aus Eigenvektoren von f bestehende Orthonormalbasis von V. Die Eigenwerte sind reelle Zahlen.*

Beweis. Die Proposition beweisen wir durch Induktion nach $n := \dim(V)$. Der Fall $n = 0$ ist trivial. Angenommen, $n > 0$. Wir behaupten, f hat eine reelle Eigenwert. Es genügt, die analoge Behauptung für eine selbstadjungierte Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ bzw. $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ zu beweisen. Wenn wir A als komplexe Matrix betrachten gibt es ohne Zweifel einen Eigenwert $\lambda \in \mathbb{C}$. Sei $v \in \mathbb{C}^n$ ein Eigenvektor. Dann haben wir

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Av, v\rangle = \langle v, Av\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle$$

und deshalb gilt $\lambda \in \mathbb{R}$. Sei $v_1 \in V$ ein Eigenvektor von f mit $\|v_1\| = 1$. Die Einschränkung von f auf $v_1^\perp$ ist ein selbstadjungierter Endomorphismus von $v_1^\perp$. Nach der Induktionshypothese gibt es eine aus Eigenvektoren bestehende Orthonormalbasis von $v_1^\perp$, die zusammen mit v_1 eine Orthonormalbasis von V ist. □

Proposition 2.23 (Hauptachsentransformation). *Sei $s\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ eine symmetrische Bilinearform mit darstellende Matrix A bezüglich der Standardbasis von $\mathbb{R}^n$. Sei $v_1, \ldots, v_n$ eine aus Eigenvektoren von A bestehende Orthonormalbasis von $\mathbb{R}^n$, mit $Av_i = \lambda_i v_i$. Dann wird s bezüglich der Basis $v_1, \ldots, v_n$ durch die Diagonalmatrix*

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

dargestellt.

Beweis. Es gilt $s(v_i, v_j) = {}^t v_i\, A\, v_j = \lambda_j\, {}^t v_i\, v_j = \lambda_j \langle v_i, v_j\rangle = \lambda_j \delta_{ij}$, wobei δ_{ij} das Kronecker-Delta bezeichnet. □

Korollar 2.24. *Notation wie in Proposition 2.23, gilt für die Signatur (n_+, n_-, n_0) von s: die Anzahl i mit $\lambda_i > 0$ ist n_+, die Anzhal mit $\lambda_i < 0$ ist n_- und die Anzahl mit $\lambda_i = 0$ ist n_0.*

Beweis. Wir ersetzen jedes v_i durch ein geeignetes Vielfaches und kommen so zu einer darstellende Matrix wie im Sylvesterschen Trägheitssatz. □

Eine symmetrische Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ heisst positiv definit, wenn die entsprechende Bilinearform $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ positiv definit ist. Ein nützliches Kriterium, zu erkennen ob A positiv definit ist, liefert der folgende Satz.

Proposition 2.25 (Hauptminorenkriterium von Hurwitz-Sylvester)**.** *Eine Matrix $A = (a_{ij})_{1\le i,j\le n} \in \mathrm{Mat}(n \times n, \mathbb{R})$ ist genau dann positiv definit, wenn jede der Untermatrizen $A_k := (a_{ij})_{1\le i,j\le k}$ mit $1 \le k \le n$ positive Determinante hat.*

Die Beträge $\det(A_k)$ heissen *Hauptminoren* von A.

Beweis. Ist A positiv definit, so ist A_k auch positiv definit für alle k, deshalb hat nur positive Eigenwerte und auch positive Determinante.

Wir beweisen die „$\Leftarrow$“-Implikation durch Induktion nach n. Der Fall $n = 1$ ist klar. Sei $n > 1$, sei $V := e_1^\perp$. Dann ist

$$v_2 := e_2 - \frac{a_{12}}{a_{11}} e_1, \qquad \dots, \qquad v_n := e_n - \frac{a_{1n}}{a_{11}} e_1$$

eine Basis von V. Es genügt nach der Induktionshypothese zu zeigen, dass die Hauptminoren der Matrix $({}^t v_i \, A v_j)_{2\le i,j\le n}$ positiv sind. Die Matrix berechnen wir, zeigen wir als rechten unteren $(n-1)\times(n-1)$-Block

$$B := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} - \frac{a_{12}^2}{a_{11}} & \cdots & a_{2n} - \frac{a_{12}a_{1n}}{a_{11}} \\ \vdots & \vdots & & \vdots \\ 0 & a_{2n} - \frac{a_{12}a_{1n}}{a_{11}} & \cdots & a_{nn} - \frac{a_{1n}^2}{a_{11}} \end{pmatrix},$$

und deren Hauptminoren erkennen wir als bis auf Faktor a_{11} gleich den Hauptminoren von B. Aber B entsteht aus A durch eine Folge von elementaren Zeilenumformungen, wobei jeweils ein Vielfaches der ersten Zeile zu einer anderen addiert wird. Also sind die Hauptminoren von B genau die Hauptminoren von A und deshalb positiv. □

Propositionen 2.14 und 2.22 haben ein gemeinsames Ergebnis, Diagonalisierbarkeit durch Orthogonalbasis, aber verschiedene Voraussetzungen. Eine gemeinsame Verallgemeinerung im Fall eines unitären Vektorraums gibt es.

Definition. Sei V ein endlichdimensionaler euklidischer oder unitärer Vektorraum. Ein Endomorphismus $f\colon V \to V$ heisst **normal** falls $f^* \circ f = f \circ f^*$.

Proposition 2.26 (Spektralsatz für normale Endomorphismen)**.** *Sei V ein endlichdimensionaler unitärer Vektorraum. Ein Endomorphismus $f\colon V \to V$ ist genau dann normal, wenn es eine aus Eigenvektoren von f bestehende Orthonormalbasis von V gibt.*

Beweis. Gibt es eine aus Eigenvektoren von f bestehende Orthonormalbasis von V, so wird f und auch f^* bezüglich einer solchen Basis durch eine Diagonalmatrix dargestellt. Es ist ein allgemeiner Fakt, dass zwei Diagonalmatrizen stets kommutieren.

Wir beweisen die „$\Rightarrow$"-Implikation durch Induktion nach n. Der Fall $n = 0$ ist trivial. Angenommen, $n > 0$. Sei $v_1 \in V$ ein Eigenvektor von f mit $\|v_1\| = 1$ und λ der entsprechende Eigenwert. Wir haben

$$\langle \overline{\lambda} v_1, f^*(v_1) - \overline{\lambda} v_1 \rangle = \overline{\lambda} \langle f(v_1), v_1 \rangle - |\lambda|^2 = 0$$

und

$$\|f^*(v_1)\|^2 = \langle v_1, f(f^*(v_1)) \rangle = \langle v_1, f^*(f(v_1)) \rangle = \langle f(v_1), f(v_1) \rangle = |\lambda|^2 = \|\overline{\lambda} v_1\|^2.$$

Deshalb ist v_1 auch Eigenvektor von f^*, mit Eigenvektor $\overline{\lambda}$. Also ist die Einschränkung von f auf $v_1^\perp$ ein normaler Endomorphismus von $v_1^\perp$. Nach der Induktionshypothese gibt es eine aus Eigenvektoren bestehende Orthonormalbasis von $v_1^\perp$. □

3 Multilineare Algebra

Wir führen das Tensorprodukt $V \otimes_K W$ zweier Vektorräume V und W über einem Körper K ein. So werden wir einer Beschreibung begegnen, die anders als bei einer klassischen Definition das Tensorprodukt nur bis auf eindeutigen Isomorphismus charakterisiert.

Wir erinnern uns zuerst an das direkte Summe. Hat V eine Basis $(v_i)_{i\in I}$ und W eine Basis $(w_j)_{j\in J}$, so gibt es für $V \oplus W$ eine Basis, die aus $(v_i, 0)$ für alle $i \in I$ und $(0, w_j)$ für alle $j \in J$ besteht. Indiziert wird diese Basis durch die dijsunkte Vereinigung $I \sqcup J$, die man auf einfache Weise als Vereinigung von $\{(i, 1) \,|\, i \in I\}$ und $\{(j, 2) \,|\, j \in J\}$ beschreibt. In der Praxis hat man aber mehr Flexibilität. Im Fall von endlichdimensionalen Vektorräume V und W mit $\dim(V) = m$ und $\dim(W) = n$ nimmt man oft $I = \{1, \ldots, m\}$ und $J = \{1, \ldots, n\}$. Mit $\dim(V \oplus W) = m + n$ ist $\{1, \ldots, m + n\}$ eine vernünftige Indexmenge einer Basis von $V \oplus W$.

Die disjunkte Vereinigung $I \sqcup J$ lässt sich so charakterisieren: *$I \sqcup J$ ist eine Menge mit Abbildungen $f_1 \colon I \to I \sqcup J$ und $f_2 \colon J \to I \sqcup J$ und der universellen Eigenschaft, dass für eine beliebige Menge T mit Abbildungen $g_1 \colon I \to T$ und $g_2 \colon J \to T$ es eine eindeutige Abbildung $h \colon I \sqcup J \to T$ gibt, so dass $h \circ f_1 = g_1$ und $h \circ f_2 = g_2$.* Zum Beispiel, für $I = \{1, \ldots, m\}$ und $J = \{1, \ldots, n\}$ haben wir mit der Inklusion $f_1 \colon \{1, \ldots, m\} \to \{1, \ldots, m + n\}$ und Addition von m als $f_2 \colon \{1, \ldots, n\} \to \{1, \ldots, m + n\}$, dass $\{1, \ldots, m + n\}$ als disjunkte Vereinigung von I und J gilt. Man kann aber auch die allgemeine Konstruktion benutzen: $I \sqcup J = \{(i, 1) \,|\, 1 \le i \le m\} \cup \{(j, 2) \,|\, 1 \le j \le n\}$ erfüllt ebenso die universelle Eigenschaft. Wenden wir die universelle Eigenschaft für $\{1, \ldots, m + n\}$ auf die Menge $\{(i, 1) \,|\, 1 \le i \le m\} \cup \{(j, 2) \,|\, 1 \le j \le n\}$ (mit

Inkusionen von I und J) an, so kommen wir zur Abbildung

$$k\colon \{1,\dots,m+n\} \to \{(i,1)\,|\,1\le i\le m\} \cup \{(j,2)\,|\,1\le j\le n\},$$

die i auf $(i,1)$ für $1 \le i \le m$ und auf $(i-m,2)$ für $m+1 \le i \le m+n$ abbildet. Wenden wir die universelle Eigenschaft für $\{(i,1)\,|\,1\le i\le m\} \cup \{(j,2)\,|\,1\le j\le n\}$ auf die Menge $\{1,\dots,m+n\}$, so kommen wir zur Abbildung

$$\ell\colon \{(i,1)\,|\,1\le i\le m\} \cup \{(j,2)\,|\,1\le j\le n\} \to \{1,\dots,m+n\},$$

$(i,1) \mapsto i$ und $(j,2) \mapsto m+j$. Obwohl wir direkt verifizieren können, dass diese Abbildungen bijektiv sind, können wir auch mit zwei weiteren Anwendengung der universellen Eigenschaft zum Schluss kommen,

$$\begin{aligned} \ell \circ k &= \mathrm{id}_{\{1,\dots,m+n\}}, \\ k \circ \ell &= \mathrm{id}_{\{(i,1)\,|\,1\le i\le m\} \cup \{(j,2)\,|\,1\le j\le n\}}. \end{aligned}$$

Dieses Muster wird sich im Fall des Tensorprodukts $V \otimes_K W$ wiederholen: eine universelle Eigenschaft bestimmt ein mathematisches Objekt bis auf eindeutigen Isomorphismus. Man hat verschiedene Möglichkeiten, das Objekt zu konstruieren. Entsteht ein Objekt durch eine Konstruktion und ein anderes Objekt durch eine andere Konstruktion, so kommen wir zu eindeutig bestimmten Abbildungen zwischen den Objekten in beide Richtungen, die inverse Isomorphismen sind.

3.1 Das Tensorprodukt

Sei K ein Körper, seien V und W Vektorräume über K. Für einen K-Vektorraum T gilt eine Abbildung

$$f\colon V \times W \to T$$

als bilinear, wenn f einzeln im ersten Faktor und im zweiten Faktor linear ist:

$$\begin{aligned} f(v+v',w) &= f(v,w) + f(v',w), & \qquad f(v,w+w') &= f(v,w) + f(v,w') \\ f(av,w) &= af(v,w), & f(v,aw) &= af(v,w). \end{aligned}$$

für alle v, $v' \in V$, w, $w' \in W$ und $a \in K$.

Definition. Sei K ein Körper, seien V und W Vektorräume über K. Die **universelle Eigenschaft des Tensorprodukts** $V \otimes_K W$ mit bilinearer Abbildung $V \times W \to V \otimes_K W$, $(v,w) \mapsto v \otimes w$ ist die Existenz und Eindeutigkeit, für jeden K-Vektorraum T mit bilinearer Abbildung $f\colon V \times W \to T$, einer linearen Abbildung $\varphi\colon V \otimes_K W \to T$, so dass

$$\varphi(v \otimes w) = f(v,w)$$

für alle $v \in V$ und $w \in W$.

Mit der universellen Eigenschaft ist klar: falls $V \otimes_K W$ existiert ist $V \otimes_K W$ bestimmt bis auf eindeutigen Isomorphismus. Wir werden mit Proposition 3.1, unten, die Existenz von $V \otimes_K W$ belegen. Die Kombination von universeller Eigenschaft und Existenzbeweis macht eine Definition bis auf eindeutigen Isomorphismus aus.

Beispiel. Seien $m, n \in \mathbb{N}$, sei $V := K^m$ mit Standardbasis $e_1, \ldots, e_m$ und $W := K^n$ mit Standardbasis $e'_1, \ldots, e'_n$. Wir zeigen, $V \otimes_K W \cong K^{mn}$. Mit Standardbasis $\tilde{e}_1, \ldots, \tilde{e}_{mn}$ von K^{mn} behaupten wir: es gibt eine eindeutige bilineare Abbildung

$$K^m \times K^n \to K^{mn},$$

unter welcher (e_i, e'_j) auf

$$e_i \otimes e'_j := \tilde{e}_{(i-1)n+j}$$

für alle i und j abgebildet wird und mit dieser bilinearen Abbildung erfüllt K^{mn} die universelle Eigenschaft des Tensorprodukts. Der erste Teil der Behauptung, Existenz und Eindeutigkeit einer bilinearen Abbildung mit $(e_i, e'_j) \mapsto \tilde{e}_{(i-1)n+j}$ für alle i und j, folgt aus der Eigenschaft, bilinear zu sein. Denn für $\mathbf{x} = (x_1, \ldots, x_m) \in K^m$ und $\mathbf{y} = (y_1, \ldots, y_n) \in K^n$ muss $(\mathbf{x}, \mathbf{y})$ auf

$$\sum_{i=1}^{m} \sum_{j=1}^{n} x_i y_j \, \tilde{e}_{(i-1)n+j} \tag{1}$$

abgebildet werden. So haben wir eine eindeutige bilineare Abbildung. Nun zur Überprüfung der universellen Eigenschaft. Sei T ein K-Vektorraum und $f\colon K^m \times K^n \to T$ eine bilineare Abbildung. Weil $(\tilde{e}_{(i-1)n+j})_{1 \le i \le m,\, 1 \le j \le n}$ eine Basis von K^{mn} ist, gibt es eine eindeutige lineare Abbildungen $\varphi\colon K^{mn} \to T$ mit $\varphi(\tilde{e}_{(i-1)n+j}) = f(e_i, e'_j)$ für alle i und j. Wir müssen verifizieren, dass $\varphi(\mathbf{x} \otimes \mathbf{y}) = f(\mathbf{x}, \mathbf{y})$ gilt für alle $\mathbf{x} \in K^m$ und $\mathbf{y} \in K^n$. Dabei ist $\mathbf{x} \otimes \mathbf{y}$ der Vektor (1). Also haben wir

$$\varphi(\mathbf{x} \otimes \mathbf{y}) = \sum_{i=1}^{m} \sum_{j=1}^{n} x_i y_j \varphi(\tilde{e}_{(i-1)n+j}) = \sum_{i=1}^{m} \sum_{j=1}^{n} x_i y_j f(e_i, e'_j) = f(\mathbf{x}, \mathbf{y}),$$

weil φ linear und f bilinear ist.

Proposition 3.1. *Sei K ein Körper, seien V und W Vektorräume über K, sei $(v_i)_{i \in I}$ eine Basis von V und $(w_j)_{j \in J}$ eine Basis von W. Dann gibt es eine eindeutige bilinear Abbildung*

$$V \times W \to \bigoplus_{(i,j) \in I \times J} K,$$

die (v_i, w_j) auf $e_{(i,j)}$ für alle $i \in I$ und $j \in J$ abbildet, wobei $(e_{(i,j)})_{(i,j) \in I \times J}$ die Standardbasis der direkten Summe von K über $(i,j) \in I \times J$ bezeichnet, und mit dieser bilinearen Abbildung wird die universelle Eigenschaft des Tensorprodukts erfüllt.

Beweis. Sei $v \in V$ und $w \in W$. Also können wir auf eindeutige Weise $v = \sum_{i \in I} a_i v_i$ schreiben mit $a_i \in K$ für alle $i \in I$ und $a_i \neq 0$ nur für endlich viele $i \in I$ und auf eindeutige Weise $w = \sum_{j \in J} b_j w_j$ schreiben mit $b_j \in K$ für alle $j \in J$ und $b_j \neq 0$ nur für endlich viele $j \in J$. Dann gilt für eine bilineare Abbildung $V \times W \to \bigoplus_{(i,j) \in I \times J} K$, die (v_i, w_j) auf $e_{(i,j)}$ für alle $i \in I$ und $j \in J$ abbildet,

$$(v, w) \mapsto \sum_{i \in I} \sum_{j \in J} a_i b_j \, e_{(i,j)}. \tag{2}$$

Tatsächlich wird mit (2) eine bilinear Abbildung definiert. Also haben wir die Existenz und Eindeutigkeit der bilinearen Abbildung $V \times W \to \bigoplus_{(i,j) \in I \times J} K$.

Sei T ein K-Vektorraum und $f\colon V \times W \to T$ eine bilineare Abbildung. Die Standardbasis von $\bigoplus_{(i,j) \in I \times J} K$ ist $(e_{(i,j)})_{(i,j) \in I \times J}$, deshalb gibt es eine eindeutige lineare Abbildung

$$\varphi\colon \bigoplus_{(i,j) \in I \times J} K \to T$$

mit $\varphi(e_{(i,j)}) = f(v_i, w_j)$ für alle $i \in I$ und $j \in J$. Oben haben wir

$$v \otimes w = \sum_{i \in I} \sum_{j \in J} a_i b_j \, e_{(i,j)}$$

berechnet. Weil φ bilinear und f linear ist, gilt

$$\varphi(v \otimes w) = \sum_{i \in I} \sum_{j \in J} a_i b_j \varphi(e_{(i,j)}) = \sum_{i \in I} \sum_{j \in J} a_i b_j f(v_i, w_j) = f(v, w).$$

Somit ist die universelle Eigenschaft bewiesen. □

Es ist gewöhnungsbedürftig, mit einem Vektorraum zu arbeiten, die nur bis auf eindeutigen Isomorphismus definiert ist. Egal welchen Vektorraum wir als $V \otimes_K W$ nehmen, gibt es die einheitliche Notation $v \otimes w$ für das Element, auf das (v, w) unter der bilinearen Abbildung $V \times W \to V \otimes_K W$ abgebildet wird. Aus der Bilinearität folgt die *Rechenregeln für Tensoren*, wobei mit Tensor ein Element von $V \otimes_K W$ bezeichnet wird:

$$\begin{gathered} v \otimes w + v' \otimes w = (v + v') \otimes w, \\ v \otimes w + v \otimes w' = v \otimes (w + w'), \\ av \otimes w = a(v \otimes w) = v \otimes aw \end{gathered}$$

für v, $v \in V$, w, $w' \in W$ und $a \in K$. Mit den Rechenregeln und der Information aus Proposition 3.1, dass eine Basis von V und eine Basis von W stets eine Basis von $V \otimes_K W$ bestimmen, kommt man gut durch.

Beispiel. Sind $f\colon V \to V'$ und $g\colon W \to W'$ lineare Abbildungen von Vektorräumen über einem Körper K, so gibt es eine lineare Abbildung

$$f \otimes g\colon V \otimes_K W \to V' \otimes_K W',$$

definiert durch

$$(f \otimes g)(v \otimes w) := f(v) \otimes g(w),$$

und so haben wir eine bilineare Abbildung

$$\mathrm{Hom}_K(V, V') \times \mathrm{Hom}_K(W, W') \to \mathrm{Hom}_K(V \otimes_K W, V' \otimes_K W').$$

(Die Komposition $V \times W \to V' \times W' \to V' \otimes_K W'$, $(v, w) \mapsto f(v) \otimes g(w)$ ist bilinear und wir wenden die universelle Eigenschaft für $V \otimes_K W$ an.)

Beispiel. Seien V und W Vektorräume über einem Körper K. Dann haben wir durch $v \otimes w \mapsto w \otimes v$ einen Isomorphismus

$$V \otimes_K W \cong W \otimes_K V.$$

(Wir haben aus der universellen Eigenschaft Abbildungen in beide Richtungen und die Kompositionen sind identische Abbildungen.)

Beispiel. Seien V und W Vektorräume über einem Körper K. Ist $(v_i)_{i\in I}$ eine Basis von V, so haben wir durch $v_i \otimes w \mapsto (0, \dots, 0, w, 0, \dots, 0)$ mit w in der i-ten Stelle einen Isomorphismus

$$V \otimes_K W \cong \bigoplus_{i\in I} W.$$

Ist $(w_j)_{j\in J}$ eine Basis von W, so haben wir durch $v \otimes w_j \mapsto (0, \dots, 0, v, 0, \dots, 0)$ mit v in der j-ten Stelle einen Isomorphismus

$$V \otimes_K W \cong \bigoplus_{j\in J} V.$$

(In beiden Fällen haben wir aus der universellen Eigenschaft eine Vorwärts-Abbildung und weil wir $\mathrm{Hom}_K(\bigoplus_{i\in I} W, T)$ mit $\prod_{i\in I} \mathrm{Hom}_K(W, T)$ identifizieren können für beliebigen K-Vektorraum T und analog für $\bigoplus_{j\in J} V$, haben wir jeweils eine Rückwärts-Abbildung.)

Ist L/K ein Körpererweiterung, so hat L die Struktur von K-Vektorraum. Es gibt also für einen beliebigen K-Vektorraum V das Tensorprodukt $L \otimes_K V$.

Proposition 3.2. *Sei K ein Körper, L/K ein Körpererweiterung und V ein K-Vektorraum. Dann hat $L \otimes_K V$ mit $b \cdot (a \otimes v) := ab \otimes v$ für alle a, $b \in L$ und $v \in V$ die Struktur von L-Vektorraum. Ist W ein zweiter K-Vektorraum und $f\colon V \to W$ eine lineare Abbildung, so ist die Abbildung*

$$\mathrm{id}_L \otimes f$$

eine L-lineare Abbildung $L \otimes_K V \to L \otimes_K W$.

Beweis. Sei $b \in L$. Weil $L \times W \to L \otimes_K W$, $(a, v) \mapsto ab \otimes v$, bilinear ist, haben wir daraus eine K-lineare Abbildung $L \otimes_K V \to L \otimes_K V$. So kommen wir zur Abbildung

$$L \times (L \otimes_K V) \to L \otimes_K V,$$

die die Struktur von L-Vektorraum auf $L \otimes_K V$ definieren soll. Die Axiome, eine L-Vektorraum zu sein, sind klar, nur braucht es für eines der Distributivgesetze eine detaillierte Erläuterung:

$$(b+b')\cdot(a\otimes v) = a(b+b')\otimes v = (ab+ab')\otimes v = ab\otimes v + ab'\otimes v = b\cdot(a\otimes v) + b'\cdot(a\otimes v).$$

Mit der weiteren Notation wird $a \otimes v$ auf $a \otimes f(v)$ und $ab \otimes v$ auf $ab \otimes f(v)$ unter $\mathrm{id}_L \otimes f$ abgebildet. Also ist $\mathrm{id}_L \otimes f$ eine lineare Abbildung von Vektorräumen über L. $\square$

3.2 Der Dualraum

Der Dualraum $V^* := \mathrm{Hom}_K(V, K)$ eines Vektoraums V über einem Körper K liefert eine Interpretation der transponierten Matrix. Wir erinnern uns, im Fall $V = K^n$ mit Standardbasis $e_1, \dots, e_n$ hat V^* auch eine Standardbasis $e_1^*, \dots, e_n^*$, mit $e_j^*(e_i) = \delta_{ij}$ (Kronecker-Delta). Dann gibt es zu einer linearen Abbildung $f\colon K^n \to K^m$ eine duale Abbildung $f^*\colon (K^m)^* \to (K^n)^*$ und bezüglich der Standardbasen sind die entsprechenden Matrizen zueinander transponiert.

Der Dualraum des Dualraums V^{**} heisst Bidualraum. Mit

$$v \mapsto ((f\colon V \to K) \mapsto f(v))$$

haben wir eine Abbildung

$$V \to V^{**}, \tag{1}$$

die offensichtlich linear ist.

Proposition 3.3. *Sei V ein Vektorraum über einem Körper K. Die in* (1) *definierte Abbildung $V \to V^{**}$ ist injektiv und im folgenden Sinn natürlich: ist W ein zweiter K-Vektorraum und $h\colon V \to W$ eine lineare Abbildung, so haben wir ein kommutatives Diagramm*

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

wo die Abbildung (1) *und die analoge Abbildung für W als vertikale Pfeile erscheinen.*

Beweis. Ist $v \in V$ von Null verschieden, so gibt es eine Basis von V, in der v enthalten ist. Deshalb gibt es ein $f \in V^*$ mit $f(v) \neq 0$ und die Abbildung (1) ist injektiv. Mit der weiteren Notation haben wir $h(v)$, das auf

$$((g\colon W \to K) \mapsto g(h(v))) \in W^{**}$$

abgebildet wird. Wir haben

$$h^{**}((f\colon V \to K) \mapsto f(v)) = ((f\colon V \to K) \mapsto f(v)) \circ h^*,$$

under dem ein $g\colon W \to K$ auf $g(h(v))$ abgebildet wird. Das Diagramm ist deshalb kommutativ. □

Weil eine injektive lineare Abbildung von endlichdimensionalen Vektorräumen derselben Dimension stets bijektiv ist, gilt im Fall $\dim(V) < \infty$, dass (1) für endlichdimensionale Vektorräume ein Isomorphismus ist.

Wir können Dualräume mit dem Tensorprodukt kombinieren. Mit dem Tensorprodukt von Abbildungen (Beispiel aus §3.1) haben wir, für Vektorräume V und W über einem Körper K:

$$\mathrm{Hom}_K(V, K) \times \mathrm{Hom}_K(W, K) \to \mathrm{Hom}_K(V \otimes_K W, K \otimes_K K).$$

Dies ist eine bilineare Abbildung, und mit $a \otimes b \mapsto ab$ haben wir $K \otimes_K K \cong K$, deshalb gibt es eine induzierte lineare Abbildung

$$V^* \otimes_K W^* \to (V \otimes_K W)^*. \tag{2}$$

Komposition von linearen Abbildung ist eine Abbildung

$$\mathrm{Hom}_K(V, K) \times \mathrm{Hom}_K(K, W) \to \mathrm{Hom}_K(V, W),$$

die ebenfalls bilinear ist, also haben wir

$$V^* \otimes_K W \to \mathrm{Hom}_K(V, W). \tag{3}$$

Proposition 3.4. *Ist mindestens einer der Vektorräume V und W endlichdimensional, so sind die linearen Abbildungen* (2) *und* (3) *Isomorphismen.*

Beweis. Angenommen, V ist endlichdimensional mit Basis $v_1, \dots, v_n$ und duale Basis $v_1^*, \dots, v_n^*$ von V^*. Mit denen identifizeren wir die linke Seite von (2) mit $(W^*)^n$ und die rechte Seite mit $(W^n)^*$. Dann ergibt (2) den Isomorphismus $(W^*)^n \to (W^n)^*$,

$$(g_1, \dots, g_n) \mapsto \big((w_1, \dots, w_n) \mapsto g_1(w_1) + \dots + g_n(w_n)\big).$$

Wir haben auch Isomorphismen $V^* \otimes_K W \cong W^n$ und $\mathrm{Hom}_K(V, W) \cong W^n$, die mit der Abbildung (3) kompatibel sind.

Ist W endlichdimensional mit Basis $w_1, \dots, w_m$, so gehen wir analog mit (2) vor und so haben wir Isomorphismen $V^* \otimes_K W \cong (V^*)^n$ und $\mathrm{Hom}_K(V, W) \cong (V^*)^n$, die mit (3) kompatibel sind. □

3.3 Symmetrische und äussere Potenz

Es gibt auch multilineare Abbildungen

$$V_1 \times \cdots \times V_n \to T$$

für Vektorräume $V_1, \ldots, V_n$ und T über einem Körper K. Zum Beispiel für $n = 3$ haben wir trilineare Abbildungen $V_1 \times V_2 \times V_3 \to T$. Mit einer analogen universellen Eigenschaft kommen wir zu einem dreifachen Tensorprodukt $V_1 \otimes_K V_2 \otimes_K V_3$, das allerdings isomorph zu $(V_1 \otimes_K V_2) \otimes_K V_3$ und zu $V_1 \otimes_K (V_2 \otimes_K V_3)$ ist. Dabei ist der Isomorphismus

$$(V_1 \otimes_K V_2) \otimes_K V_3 \cong V_1 \otimes_K (V_2 \otimes_K V_3)$$

nennenswert, aber sobald bekannt haben wir auch mit n Faktoren nichts anders als iteriertes Tensorprodukt.

Wie im Fall der Determinante ist es interessant, den Blick auf den Fall $V_1 = \cdots = V_n$ zu fokussieren. Also betrachten wir multilineare Abbildung

$$\overbrace{V \times \cdots \times V}^{n} \to T$$

für Vektorräume V und T und ein $n \in \mathbb{N}_{>0}$. Es gibt dabei das n-fache Tensorprodukt

$$\bigotimes^n V := \overbrace{V \otimes_K \cdots \otimes_K V}^{n}.$$

Aus der Diskussion bezüglich der Determinante können wir uns erinnern: Eine multilineare Abbildung $f\colon V^n \to T$

heisst ...	falls stets gilt
symmetrisch	$f(\ldots, v, \ldots, v', \ldots) = f(\ldots, v', \ldots, v, \ldots)$
schiefsymmetrisch	$f(\ldots, v, \ldots, v', \ldots) = -f(\ldots, v', \ldots, v, \ldots)$
alternierend	$f(\ldots, v, \ldots, v, \ldots) = 0$

Falls $\operatorname{char}(K) \neq 2$ ist schiefsymmetrisch äquivalent zu alternierend. Falls $\operatorname{char}(K) = 2$ ist schiefsymmetrisch äquivalent zu symmetrisch. Deshalb wird der Schwerpunkt auf symmetrische und alternierende multilineare Abbildungen gesetzt.

Definition. Sei K ein Körper, V ein Vektorraum über K und $n \in \mathbb{N}_{>0}$. Die **universelle Eigenschaft der symmetrischen Potenz** $\operatorname{Sym}^n V$ mit symmetrischer multilinearer Abbildung $V^n \to \operatorname{Sym}^n V$, $(v_1, \ldots, v_n) \mapsto v_1 \cdots v_n$ ist die Existenz und Eindeutigkeit, für jeden K-Vektorraum T mit symmetrischer multilinearer Abbildung $f\colon V^n \to T$, einer linearen Abbildung $\varphi\colon \operatorname{Sym}^n V \to T$, so dass

$$\varphi(v_1 \cdots v_n) = f(v_1, \ldots, v_n)$$

für alle $v_1, \ldots, v_n \in V$. Die **universelle Eigenschaft der äusseren Potenz $\bigwedge^n V$** mit alternierender multilinearer Abbildung $V^n \to \bigwedge^n V$, $(v_1, \ldots, v_n) \mapsto v_1 \wedge \cdots \wedge v_n$ ist die Existenz und Eindeutigkeit, für jeden K-Vektorraum T mit alternierender multilinearer Abbildung $f\colon V^n \to T$, einer linearen Abbildung $\varphi\colon \bigwedge^n V \to T$, so dass

$$\varphi(v_1 \wedge \cdots \wedge v_n) = f(v_1, \ldots, v_n)$$

für alle $v_1, \ldots, v_n \in V$.

Beispiel. Sei $m \in \mathbb{N}$, $n \in \mathbb{N}_{>0}$ und $V := K^m$ mit Standardbasis $e_1, \ldots, e_m$. Wir zeigen,

$$\operatorname{Sym}^n V \cong \bigoplus_{1 \le i_1 \le \cdots \le i_n \le m} K \qquad \text{und} \qquad \bigwedge^n V \cong \bigoplus_{1 \le i_1 < \cdots < i_n \le m} K.$$

Wir behaupten, es gibt eine eindeutige symmetrische multilineare Abbildung

$$\overbrace{K^m \times \cdots \times K^m}^{n} \to \bigoplus_{1 \le i_1 \le \cdots \le i_n \le m} K,$$

under welcher $(e_{i_1}, \ldots, e_{i_n})$ auf

$$e_{i_1} \cdots e_{i_n} := e_{(i_{\sigma(1)}, \ldots, i_{\sigma(n)})}$$

abgebildet wird, wobei $\sigma \in S_n$ ausgewählt wird, so dass

$$i_{\sigma(1)} \le \cdots \le i_{\sigma(n)},$$

und mit dieser symmetrischen multilinearen Abbildung erfüllt $\bigoplus_{1 \le i_1 \le \cdots \le i_n \le m} K$ die universelle Eigenschaft der symmetrischen Potenz. Ähnlich geht der Fall von $\bigwedge^n V$, mit

$$e_{i_1} \wedge \cdots \wedge e_{i_n} := \begin{cases} \operatorname{sgn}(\sigma) e_{(i_{\sigma(1)}, \ldots, i_{\sigma(n)})} & \text{falls } i_a \ne i_b \text{ für alle } a \ne b, \\ 0 & \text{sonst.} \end{cases}$$

Die Verifikation ist in beiden Fällen ähnlich, wie im Beispiel von $K^m \otimes_K K^n$.

Sei K ein Körper, V ein Vektorraum über K und $n \in \mathbb{N}_{>0}$. Wir definieren die folgenden Untervektorräume von $\bigotimes^n V$:

$$\begin{aligned} S^n(V) &:= \operatorname{span}(\{v_1 \otimes \cdots \otimes v_i \otimes \cdots \otimes v_j \otimes \cdots \otimes v_n - v_1 \otimes \cdots \otimes v_j \otimes \cdots \otimes v_i \otimes \cdots \otimes v_n \mid \\ &\qquad\qquad v_1, \ldots, v_n \in V,\ 1 \le i < j \le n\}), \\ A^n(V) &:= \operatorname{span}(\{v_1 \otimes \cdots \otimes v_i \otimes \cdots \otimes v_j \otimes \cdots \otimes v_n \mid v_1, \ldots, v_n \in V, \\ &\qquad\qquad 1 \le i < j \le n \text{ mit } v_i = v_j\}). \end{aligned}$$

Proposition 3.5. *Sei K ein Körper, V ein Vektorraum über K und $n \in \mathbb{N}_{>0}$. Dann bekommen wir aus der multilinearen Abbildung $V^n \to \bigotimes^n V$, $(v_1, \dots, v_n) \mapsto v_1 \otimes \cdots \otimes v_n$ durch Komposition mit der kanonischen linearen Abbildung eine symmetrische multilineare Abbildung*

$$V^n \to \bigotimes\nolimits^n V/S^n(V),$$

mit der $\bigotimes^n V/S^n(V)$ die universelle Eigenschaft der symmetrischen Potenz erfüllt. Ebebfalls bekommen wir aus der multilinearen Abbildung $V^n \to \bigotimes^n V$, $(v_1, \dots, v_n) \mapsto v_1 \otimes \cdots \otimes v_n$ durch Komposition mit der kanonischen linearen Abbildung eine alternierende multilineare Abbildung

$$V^n \to \bigotimes\nolimits^n V/A^n(V),$$

mit der $\bigotimes^n V/A^n(V)$ die universelle Eigenschaft der äusseren Potenz erfüllt.

Beweis. Die angegebenen multilinearen Abbildungen sind symmetrisch bzw. alternierend. Ist T ein K-Vektorraum mit symmetrischer multilinearer Abbildung $f\colon V^n \to T$, so haben wir aus der universellen Eigenschaft des Tensorprodukts eine eindeutige lineare Abbildung $\psi\colon \bigotimes^n V \to T$, so dass $\psi(v_1 \otimes \cdots \otimes v_n) = f(v_1, \dots, v_n)$ für alle $v_1, \dots, v_n \in V$. Weil f symmetrisch ist, gilt

$$\psi(v_1 \otimes \cdots \otimes v_i \otimes \cdots \otimes v_j \otimes \cdots \otimes v_n) = \psi(v_1 \otimes \cdots \otimes v_j \otimes \cdots \otimes v_i \otimes \cdots \otimes v_n)$$

für alle $v_1, \dots, v_n$ und $1 \le i < j \le n$. Deshalb ist die Einschränkung von ψ auf $S^n(V)$ Null. Es gibt also eine lineare Abbildung $\varphi\colon \bigotimes^n V/S^n(V) \to T$, deren Komposition $\bigotimes^n V \to \bigotimes^n V/S^n(V) \to T$ mit der kanonischen linearen Abbildung gleich ψ ist, und φ ist eindeutig bestimmt. Im Fall eines K-Vektorraums mit alternierender multilinearer Abbildung haben wir auf ählinche Weise, dass es eine eindeutige kompatible lineare Abbildung $\varphi\colon \bigotimes^n V/A^n(V) \to T$ gibt. □

Wie im Fall des Tensorprodukts gibt es zu einer linearen Abbildung $f\colon V \to V'$ lineare Abbildungen

$$\mathrm{Sym}^n f\colon \mathrm{Sym}^n V \to \mathrm{Sym}^n V' \qquad \text{und} \qquad \bigwedge\nolimits^n f\colon \bigwedge\nolimits^n V \to \bigwedge\nolimits^n V'.$$

Die folgenden Resultate legen Verbindungen zwischen der symmetrischen oder äusseren Potenz und der Determinante oder dem Dualraum vor.

Proposition 3.6. *Sei K ein Körper, V ein endlichdimensionaler K-Vektorraum, $n := \dim V$ und $f\colon V \to V$ ein Endomorphismus. Angenommen, $n > 0$. Dann ist $\bigwedge^n f\colon \bigwedge^n V \to \bigwedge^n V$ gegeben durch Skalarmultiplikation mit* $\det(f)$.

Beweis. Es genügt, den Fall $V = K^n$ zu behandeln. Wir haben durch $e_1 \wedge \cdots \wedge e_n \mapsto 1$ einen Isomorphismus $\bigwedge^n V \cong K$, unter dem $\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_n$ auf die Determinante der Matrix mit Spaltenvektoren $\mathbf{x}_1, \dots, \mathbf{x}_n$ für $\mathbf{x}_1, \dots, \mathbf{x}_n \in V$ abgebildet wird. Ist A die entsprechende Matrix zum Endomorphismus f, so wird $\bigwedge^n f$ durch Multiplikation mit $\det(A)$ gegeben. □

Proposition 3.7. *Sei K ein Körper und V ein endlichdimensionaler K-Vektorraum. Dann wird für $n \in \mathbb{N}_{>0}$ ein Isomorphismus $\bigwedge^n(V^*) \to (\bigwedge^n V)^*$ von K-Vektorräumen durch*

$$f_1 \wedge \cdots \wedge f_n \mapsto \big(v_1 \wedge \cdots \wedge v_n \mapsto \det((f_i(v_j))_{1\le i,j\le n})\big)$$

definiert.

Beweis. Es genügt, den Fall $V = K^m$ zu behandeln. Dann berechnen wir mit expliziten Basen (vgl. Beispiel, oben). □

Proposition 3.8. *Sei K ein Körper, V ein endlichdimensionaler K-Vektorraum und $n \in \mathbb{N}_{>0}$ mit $\mathrm{char}(K) \notin \{\textit{Primzahlen } p \,|\, p \le n\}$. Dann wird ein Isomorphismus $\mathrm{Sym}^n(V^*) \to (\mathrm{Sym}^n V)^*$ von K-Vektorräumen durch*

$$f_1 \cdots f_n \mapsto \Big(v_1 \cdots v_n \mapsto \sum_{\sigma \in S_n} \prod_{i=1}^{n} f_i(v_{\sigma(i)})\Big).$$

definiert.

Beweis. Es genügt, den Fall $V = K^m$ zu behandeln. Wir berechnen mit expliziten Basen (vgl. Beispiel, oben) und zeigen, dass die Komposition

$$(\mathrm{Sym}^n V)^* \to (\bigotimes^n V)^* \cong \bigotimes^n(V^*) \to \mathrm{Sym}^n(V^*) \to (\mathrm{Sym}^n V)^*$$

durch Skalarmultiplikation mit $n!$ gegeben ist, wobei wir als letzte Abbildung diejenige aus der Proposition einsetzen und die anderen aus Propositionen 3.4 und 3.5 erlangen. □