



seit 1558

Friedrich-Schiller-Universität Jena  
Mathematisches Institut

Lineare Algebra und Analytische Geometrie 2  
Sommersemester 2007

David J. Green  
8. Februar 2008

# Inhaltsverzeichnis

10	Das Minimalpolynom einer Matrix	2
11	Der Satz von Cayley-Hamilton	9
12	Algebraische Strukturen: Quotienten	15
13	Algebraische und Geometrische Vielfachheit	22
14	Nilpotente Endomorphismen	28
15	Die Jordansche Normalform	33
16	Der Dualraum	37
17	Bilinearformen	43
18	Affine Unterräume 2	60
19	Affine Quadriken	65
20	Projektive Räume: Eine kurze Einführung	71
21	Zwei Anwendungen	75
21.1	Die Dreifärbungszahl in der Knotentheorie . . . . .	75
21.2	Lineare Codes . . . . .	76

# Einleitung

Bei unserer ersten Begegnung mit Eigenwerten & Co sind einige Fragen offen geblieben bzw. sie wurden gar nicht erst gestellt.

- Die Matrix  $A \in M_4(\mathbb{R})$  habe das charakteristische Polynom  $X^4 - X^2 = X^2(X - 1)(X + 1)$ . Wir wissen, dass  $0, 1, -1$  Eigenwerte sind, und dass es keine weiteren Eigenwerte gibt. Wir werden sehen, dass die Eigenräume zu den Eigenwerten  $1, -1$  eindimensional sind, und dass der Eigenraum zum Eigenwert  $0$  höchstens zweidimensional ist. Stichwort: *Algebraische und geometrische Vielfachheit*.
- Vorher werden wir sehen, dass diese Matrix die Gleichung  $A^4 = A^2$  erfüllen muss. Dies ist der *Satz von Cayley–Hamilton*.
- Später werden wir sehen, dass es im wesentlichen nur zwei verschiedene Möglichkeiten für  $A$  gibt, und zwar

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Diese beiden *Jordan-Normalformen* lassen sich durch ihre *Minimalpolynome* unterscheiden, denn die erste der beiden Matrizen erfüllt  $A^3 = A$ , die zweite nicht.

## 10 Das Minimalpolynom einer Matrix

**Lemma 10.1** Sei  $k$  ein Körper und  $A \in M_n(k)$  eine quadratische Matrix. Dann gibt es ein  $m \geq 1$  und ein Polynom  $f = a_m X^m + a_{m-1} X^{m-1} + \dots + a_1 X + a_0$  vom Grad  $m$  (d.h.  $a_m \neq 0$ ) mit Koeffizienten aus  $k$  derart, dass  $f(A) = 0$  gilt, d.h. es ist  $a_m A^m + a_{m-1} A^{m-1} + \dots + a_1 A + a_0 E_n = 0$  in  $M_n(k)$ .

*Beweis.*  $M_n(k)$  ist ein  $k$ -Vektorraum, und zwar von Dimension  $n^2$ : eine Basis stellen die Matrizen  $E(r, s)$  für  $r, s \in \{1, \dots, n\}$  dar, wobei  $E(r, s)_{ij} = \delta_{ir} \delta_{js}$ , d.h.  $E(r, s)$  hat eine 1 an der  $(r, s)$ -Stelle, und alle anderen Einträge sind Null. Diese Matrizen sind linear unabhängig (Vergleich der  $(r, s)$ -Einträge), und wegen  $A = \sum_{r,s=1}^n A_{rs} E(r, s)$  spannen Sie  $M_n(k)$  auch auf.

Somit sind die  $n^2 + 1$  Matrizen  $E_n, A, A^2, A^3, \dots, A^{n^2}$  linear abhängig über  $k$ , d.h. es gibt Skalare  $a_0, a_1, \dots, a_{n^2}$ , die nicht alle Null sind und  $a_{n^2} A^{n^2} + \dots + a_2 A^2 + a_1 A + a_0 E_n = 0$  erfüllen. Da  $E_n \neq 0$  muss dazu  $a_i \neq 0$  sein für mindestens ein  $i \geq 1$ . Also  $f = a_{n^2} X^{n^2} + \dots + a_2 X^2 + a_1 X + a_0$  ist ein solches Polynom mit  $1 \leq m \leq n^2$ . ■

*Beispiel* Sei  $A \in M_3(\mathbb{R})$  die Matrix

$$A = \begin{pmatrix} 0 & 1 & 3 \\ -1 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Nach dem Lemma muss es ein Polynom vom Grad höchstens 9 geben, das in  $A$  verschwindet. Dies ist tatsächlich so, denn

$$A^2 = \begin{pmatrix} -1 & 0 & 5 \\ 0 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \quad A^3 = \begin{pmatrix} 0 & -1 & 2 \\ 1 & 0 & -3 \\ 0 & 0 & 1 \end{pmatrix} \quad A^4 = E_3.$$

Sofort erkennt man, dass  $f(A) = 0$  für  $f = X^4 - 1$ . Schaut man etwas genauer hin, so sieht man, dass  $A^3 + A = A^2 + E_3$ , weshalb auch  $g(A) = 0$  auch für  $g = X^3 - X^2 + X - 1$  gilt.

Es stellt sich heraus, dass  $g = p_A$ , das charakteristische Polynom von  $A$ . Nach dem Satz von Cayley–Hamilton gilt  $p_A(A) = 0$  immer.

Später werden wir den Satz von Cayley–Hamilton beweisen. Vorher werden wir aber zeigen, dass es zu jede Matrix  $A \in M_n(k)$  ein besonderes Polynom  $m_A$  mit  $m_A(A) = 0$  gibt: das Minimalpolynom, das den kleinstmöglichen Grad hat. Vorher bauen wir aber unsere Kenntnisse über Polynome ein bisschen aus.

## Polynome

*Bezeichnung* Sei  $R$  ein kommutativer Ring; Beispiele sind  $R = \mathbb{Z}$  oder  $R$  ein Körper. Der Ring aller Polynome  $f(X) = a_n X^n + a_{n-1} X^{n-1} + \dots + a_2 X^2 + a_1 X + a_0$  mit Koeffizienten aus  $k$  bezeichnet man mit  $R[X]$ .

**Lemma 10.2** *Sei  $R$  ein Körper oder  $R = \mathbb{Z}$ . Allgemeiner sei  $R$  irgendein kommutativer Ring derart, dass aus  $ab = 0$  in  $R$  folgt, dass mindestens eins aus  $a = 0, b = 0$  gelten muss.*

*Für Polynome  $f, g \in R[X] \setminus \{0\}$  gilt dann  $fg \neq 0$  und*

$$\text{grad}(fg) = \text{grad}(f) + \text{grad}(g).$$

*Beweis.* Sei  $f = \sum_{r=0}^m a_r X^r$  und  $g = \sum_{s=0}^n b_s X^s$  mit  $a_m, b_n \neq 0$ . Es ist also  $\text{grad}(f) = m$  und  $\text{grad}(g) = n$ . Ferner ist

$$fg = \left( \sum_{r=0}^m a_r X^r \right) \cdot \left( \sum_{s=0}^n b_s X^s \right) = \sum_{r=0}^m \sum_{s=0}^n a_r b_s X^{r+s}.$$

Die höchste Potenz von  $X$ , die vorkommt, ist also  $X^{m+n}$ , und diese Potenz kommt nur im Fall  $r = m, s = n$  vor. Also

$$fg = a_m b_n X^{m+n} + \text{Terme von kleinerem Grad}.$$

Nach den Voraussetzungen folgt  $a_m b_n \neq 0$  aus  $a_m, b_n \neq 0$ . Also  $\text{grad}(fg) = m+n$ . ■

Polynome über einen Körper haben einen größten gemeinsamen Teiler, der sich recht ähnlich zum ggT von ganzen Zahlen verhält.

**Lemma 10.3 (Divisionsalgorithmus)** *Sei  $k$  ein Körper und seien  $f, g \in k[X]$  Polynome mit  $g \neq 0$ . Dann gibt es Polynome  $q, r \in k[X]$  derart, dass  $f = qg + r$  gilt und außerdem entweder  $r = 0$  oder  $\text{grad}(r) < \text{grad}(g)$ .*

*Beweis.* Um Fallunterscheidungen zu vermeiden, ist es hilfreich, für diesen Beweis den Grad des Nullpolynoms als  $-\infty$  zu setzen;  $-1$  geht aber genau so gut.

Induktion über  $\text{grad}(f)$ . Ist ( $f = 0$  oder)  $\text{grad}(f) < \text{grad}(g)$ , so sind wir fertig mit  $q = 0$  und  $r = f$ . Nun sei  $\text{grad}(f) \geq \text{grad}(g)$ , also  $f = a_n X^n + a_{n-1} X^{n-1} + \dots + a_0$  und  $g = b_m X^m + b_{m-1} X^{m-1} + \dots + b_0$  mit  $a_i, b_j \in k$  und außerdem  $a_n, b_m \neq 0$  sowie  $n \geq m$ . Setzen wir  $f_1 := f - \frac{a_n}{b_m} X^{n-m} g$ . Dann

$$f_1 = \left( a_{n-1} - \frac{a_n b_{m-1}}{b_m} \right) X^{n-1} + \left( a_{n-2} - \frac{a_n b_{m-2}}{b_m} \right) X^{n-2} + \dots + \left( a_{n-m} - \frac{a_n b_0}{b_m} \right) X^{n-m} \\ + a_{n-m-1} X^{n-m-1} + \dots + a_0,$$

denn die Koeffizienten von  $X^n$  heben einander nach Konstruktion auf. Es ist somit  $\text{grad}(f_1) < \text{grad}(f)$ , weshalb es nach der Induktionsannahme Polynome  $q_1, r \in k[X]$  gibt mit  $f_1 = q_1g + r$  und  $\text{grad}(r) < \text{grad}(g)$ . Mit  $q := q_1 + \frac{a_n}{b_m}X^{n-m}$  ist dann  $f = qg + r$ , und ( $r = 0$  oder)  $\text{grad}(r) < \text{grad}(g)$ . ■

*Beispiel* In  $\mathbb{R}[X]$  sei  $f = X^4 - 8X^2 + 16$  sowie  $g = X^3 + 2X^2 + X + 2$ . Es ist  $\text{grad}(f) \geq \text{grad}(g)$ , also setzen wir  $f_1 = f - Xg = -2X^3 - 9X^2 - 2X + 16$ . Es ist auch  $\text{grad}(f_1) \geq \text{grad}(g)$ , weshalb wir  $f_2 = f_1 + 2g = -5X^2 + 20$  setzen. Da  $\text{grad}(f_2) < \text{grad}(g)$  ist, folgern wir: es ist  $f_1 = -2g + r$  und  $f = qg + r$  für  $q = X - 2$  und  $r = -5X^2 + 20$ .

## Ideale in Polynomringen

*Bezeichnung* Ist  $h$  ein Teiler von  $f$ , so schreibt man  $h \mid f$ .

**Lemma 10.4** Sei  $k$  ein Körper und  $I \subseteq k[X]$  eine Teilmenge mit den folgenden Eigenschaften:

- (I1)  $f + g \in I$  für alle  $f, g \in I$ .
- (I2)  $fg \in I$  für alle  $g \in I$  und für alle  $f \in k[X]$ .
- (I3) Es gibt mindestens ein  $0 \neq g \in I$ .

Dann gibt es genau ein normiertes Polynom  $g_0 \in I$  mit der folgenden Eigenschaft:  $I = \{fg_0 \mid f \in k[X]\}$ , d.h.  $I = \{f \in k[X] \mid f \text{ ist durch } g_0 \text{ teilbar}\}$ .

*Beispiele* Eine solche Teilmenge nennt man ein *Ideal* in  $k[X]$ , wobei  $\{0\}$  eigentlich auch als ein Ideal gilt. Das Polynom  $g_0$  nennt man den normierten *Erzeuger* des Ideals  $I$ . Wir werden insbesondere mit den folgenden beiden Ideale  $\neq 0$  arbeiten:

- Für eine Matrix  $A \in M_n(k)$  sei  $I$  die Menge  $I = \{f \in k[X] \mid f(A) = 0\}$ .  
Zu (I1): ist  $f(A) = g(A) = 0$ , so ist auch  $(f + g)(A) = f(A) + g(A) = 0$ .  
Zu (I2): sind  $f, g$  Polynome mit  $g(A) = 0$ , so ist  $(fg)(A) = f(A)g(A) = 0$ .  
Bedingung (I3) ist die Aussage von Lemma 10.1.

In diesem Fall werden wir  $g_0$  das *Minimalpolynom*  $m_A$  von  $A$  nennen.

- Seien  $f, g \in k[X]$  zwei Polynome, nicht beide = 0. Sei  $I = \{af + bg \mid a, b \in k[X]\}$ .  
Zu (I1): ist  $p = af + bg$  und  $q = cf + dg$ , so ist  $p + q = (a + c)f + (b + d)g$ .  
Zu (I2): sind  $p, q$  Polynome mit  $q = af + bg$ , so ist  $pq = a'f + b'g$  für  $a' = pa$ ,  $b' = pb$ .  
Zu (I3): sowohl  $f = 1 \cdot f + 0 \cdot g$  als auch  $g = 0 \cdot f + 1 \cdot g$  liegen in  $I$ ; aber  $f, g$  sind nicht beide = 0.

Hier wird es sich herausstellen, dass  $g_0$  der größte gemeinsame Teiler von  $f$  und  $g$  ist.

*Beweis.* Sei  $n = \min\{\text{grad}(f) \mid 0 \neq f \in I\}$ . Wegen (I3) existiert dieses Minimum. Sei  $g \in I$  ein Polynom vom Grad  $n$ , und sei  $g_0 = \frac{1}{\lambda}g$  für  $\lambda =$  führender Koeffizient von  $g$ . Dann ist  $g_0$  normiert vom Grad  $n$ , und es ist  $g_0 \in I$  wegen (I2).

Wäre  $g'_0 \neq g_0$  ein weiteres normiertes Polynom vom Grad  $n$  in  $I$ , so wäre  $0 \neq g'_0 - g_0$  ein Element aus  $I$  (wegen (I1) und (I2)) vom Grad  $< n$ , ein Widerspruch. Somit gibt es genau ein solches  $g_0$ .

Wir müssen noch zeigen: es ist  $I = \{fg_0 \mid f \in k[X]\}$ . Die Inklusion  $\supseteq$  folgt direkt aus (I2). Ist wiederum  $h \in I$ , so gibt es nach Lemma 10.3 Polynome  $q, r \in k[X]$  mit  $h = qg_0 + r$  und  $r = 0$  oder  $\text{grad}(r) < \text{grad}(g_0)$ . Es ist  $r = h - qg_0$ , weshalb  $r \in I$  wegen (I1) und (I2). Aufgrund der Minimalität von  $\text{grad}(g_0)$  muss also  $r = 0$  sein, und  $h = qg_0$ . ■

## ggT und euklidischer Algorithmus

**Lemma 10.5** *Sei  $k$  ein Körper, und seien  $f, g \in k[X]$  zwei Polynome, die nicht beide  $= 0$  sind. Dann gibt es genau ein Polynom  $h \in k[X]$  mit den folgenden Eigenschaften:*

- $h$  ist normiert;
- $h$  teilt sowohl  $f$  als auch  $g$ .
- Ist  $h'$  ein weiterer gemeinsamer Teiler von  $f$  und  $g$ , so teilt  $h'$  auch  $h$ .

Dieses  $h$  nennt man den größten gemeinsamen Teiler  $\text{ggT}(f, g)$  von  $f$  und  $g$ .

*Zusatz:* Es gibt Polynome  $a, b \in k[X]$  mit  $h = af + bg$ .

*Bemerkung* Es ist Geschmackssache, ob man verlangt, dass der ggT normiert sei. Verlangt man dies nicht, so ist der ggT natürlich nicht mehr eindeutig definiert.

*Beweis.* Sei  $I$  die Menge  $I = \{h \in k[X] \mid \exists a, b \in k[X] \text{ mit } h = af + bg\}$ . In den Beispielen zu Lemma 10.4 wurde gezeigt, dass  $I$  ein Ideal  $\neq 0$  ist. Nach diesem Lemma gibt es also genau ein normiertes Polynom  $h \in I$  mit  $I = \{p \in k[X] \mid h \text{ teilt } p\}$ . Wegen  $h \in I$  erfüllt  $h$  den Zusatz. Wegen  $f, g \in I$  ist  $h$  ein gemeinsamer Teiler von  $f$  und  $g$ . Wegen  $h = af + bg$  ist jeder gemeinsamer Teiler von  $f, g$  gleichzeitig ein Teiler von  $h$ . Somit ist  $h$  ein ggT; wir müssen nur noch die Eindeutigkeit zeigen.

Ist  $h_1$  ein weiterer größter gemeinsamer Teiler, so ist  $h_1$  durch  $h$  teilbar und auch umgekehrt. Somit haben beide den gleichen Grad, d.h. der Quotient ist ein Skalar. Da beide normiert sind, muss dieses Skalar 1 betragen, also  $h_1 = h$ . ■

Sei  $k$  ein Körper und  $f, g \in k[X]$  zwei Polynome, beide  $\neq 0$ . Der folgende Algorithmus konstruiert den größten gemeinsamen Teiler  $\text{ggT}(f, g)$ .

### Algorithmus 10.6 (Der euklidische Algorithmus)

Input: Zwei Polynome  $f, g \in k[X]$ , beide  $\neq 0$ .

Output: Polynome  $h, a, b$  mit  $h = \text{ggT}(f, g)$  und  $h = af + bg$ .

- Ggf.  $f, g$  vertauschen, um  $\text{grad}(g) \leq \text{grad}(f)$  sicherzustellen.
- $q, r \in k[X]$  berechnen mit  $f = qg + r$ , und  $r = 0$  oder  $\text{grad}(r) < \text{grad}(g)$ .
- Ist  $r = 0$ , dann fertig mit  $h = \frac{1}{\lambda}g$ ,  $a = 0$ ,  $b = \frac{1}{\lambda}$  ( $\lambda$  der führende Koeffizient von  $g$ ).
- Ist  $r \neq 0$ , dann Algorithmus auf  $g, r$  anwenden: erhalte  $h', a', b'$  mit  $h' = a'g + b'r$  und  $h' = \text{ggT}(g, r)$ . Also fertig mit  $h = h'$ ,  $a = b'$ ,  $b = a' - b'q$ .

**Hilfssatz 1** *Nach endlich vielen Schritten hört dieser Algorithmus auf. Am Ende haben tatsächlich  $h, a, b$  die behaupteten Eigenschaften.*

*Beweis.* Induktion über  $\text{grad}(f) + \text{grad}(g)$ . Den Induktionsanfang stellt den Fall  $r = 0$  dar. Ist  $r \neq 0$ , dann ist  $\text{grad}(r) < \text{grad}(g) \leq \text{grad}(f)$  und deshalb  $\text{grad}(g) + \text{grad}(r) < \text{grad}(f) + \text{grad}(g)$ . Nach Induktionsannahme also haben  $h', a', b'$  die erwünschten Eigenschaften für  $g, r$ : man beachte, dass  $g, r$  nie vertauscht werden im ersten Schritt. Es ist aber

$$h' = a'g + b'r = a'g + b'(f - qg) = b'f + (a' - b'q)g,$$

und aus  $h' \mid g$  und  $h' \mid r$  folgt  $h' \mid f$ , denn  $f = qg + r$ . Aus dem Beweis von Lemma 10.5 folgt jetzt  $h = \text{ggT}(g, f)$ , denn  $h$  ist normiert, ein gemeinsamer Teiler und liegt in der Menge  $I$ . ■

*Beispiel* Für  $k = \mathbb{R}$  führen wir den Algorithmus für  $f = X^3 + 2X^2 + X + 2$  und  $g = X^4 - 8X^2 + 16$  durch. Da  $\text{grad}(f) < \text{grad}(g)$  vertauschen wir  $f, g$  und setzen  $f_1 = X^4 - 8X^2 + 16$ ,  $f_2 = X^3 + 2X^2 + X + 2$ . Wie oben berechnet ist  $f_1 = q_1f_2 + f_3$  für  $q_1 = X - 2$ ,  $f_3 = -5X^2 + 20$ . Jetzt berechnen wir: es ist  $f_2 = q_2f_3 + f_4$  für  $q_2 = -\frac{1}{5}X - \frac{2}{5}$ ,  $f_4 = 5X + 10$ . Dann sehen wir: es ist  $f_3 = q_3f_4 + 0$  für  $q_3 = -X + 2$ . Wir haben den Fall  $r = 0$  erreicht. Also  $h = X + 2 = 0 \cdot f_3 + \frac{1}{5}f_4$  ist ein gemeinsamer Teiler von  $f_3, f_4$  und deshalb von  $f, g$ . Wegen  $f_4 = f_2 - q_2f_3$  folgt  $h = \frac{1}{5}f_2 + \frac{1}{25}(X + 2)f_3$ . Wegen  $f_3 = f_1 - q_1f_2$  folgt  $h = \frac{1}{25}(X + 2)f_1 - \frac{1}{25}(X^2 - 9)f_2$ . Wegen  $f_1 = g$ ,  $f_2 = f$  folgt: es ist  $h = X + 2 = \frac{1}{25}((9 - X^2)f + (X + 2)g)$ , und  $h$  teilt sowohl  $f$  als auch  $g$ . Es ist also

$$\text{ggT}(X^3 + 2X^2 + X + 2, X^4 - 8X^2 + 16) = X + 2.$$

Es lohnt sich, an dieser Stelle den Ausdruck für  $h$  zur Kontrolle auszurechnen. Bei meinem ersten Versuch musste ich feststellen, dass ich mich verrechnet hatte.



## Das Minimalpolynom

*Definition* Sei  $k$  ein Körper und  $A \in M_n(k)$  eine quadratische Matrix. Nach Lemma 10.4 und dem ersten Beispiel dazu gibt es genau ein normiertes Polynom, das man das *Minimalpolynom*  $m_A(X)$  von  $A$  nennt, derart, dass  $m_A(A) = 0$  und  $\{f \in k[X] \mid f(A) = 0\} = \{gm_A \mid g \in k[X]\}$  erfüllt. Nach dem Beweis des Lemmas ist  $m_A(X)$  zugleich das einzige normierte Polynom kleinsten möglichen Grades, das in  $X = A$  verschwindet.

*Beispiel* Im Beispiel am Anfang des Kapitels hat

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

das Minimalpolynom  $m_A(X) = X^3 - X^2 + X - 1 = p_A(X)$ ; ein Vergleich der Einträge an der  $(2, 1)$ -Stelle und danach an der  $(2, 3)$ -Stelle zeigt, dass  $E_3, A, A^2$  linear unabhängig sind.

Die Matrix  $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$  dagegen hat Minimalpolynom  $m_A(X) = (X - 2)(X - 1) = X^2 - 3X + 2$ : dies erkennt man auch daran, dass  $Ae_1 = 2e_1$ ,  $Ae_2 = e_2$  und  $Ae_3 = e_3$ , also  $(A - 2E_3)(A - E_3)e_i = 0$  für alle  $i$ . Hier beträgt das charakteristische Polynom  $p_A(X) = (X - 2)(X - 1)^2$ , es muss ja Grad 3 haben. In diesem Fall ist  $m_A$  zwar ein Teiler von  $p_A$ , aber die beiden Polynome sind nicht gleich.

Eine Formulierung des Satzes von Cayley–Hamilton lautet: das Minimalpolynom ist immer ein Teiler des charakteristischen Polynoms.

## Das Minimalpolynom eines Endomorphismus

**Lemma 10.7** Sei  $F$  ein Endomorphismus des endlich dimensionalen  $k$ -Vektorraums  $V$ . Sei  $B$  eine Basis von  $V$ , und sei  $A$  die quadratische Matrix  $A = {}_B M_B(F)$ . Für jedes Polynom  $f \in k[X]$  gilt dann

$$f(F) = 0 \iff f(A) = 0.$$

Somit hat auch der Endomorphismus  $F$  ein Minimalpolynom  $m_F(X) \in k[X]$ , und für jede Matrix  $A = {}_B M_B(F)$  von  $F$  gilt  $m_A(X) = m_F(X)$ .

*Bemerkung* Ist  $f(X) = a_m X^m + \dots + a_1 X + a_0$ , so ist  $f(F) = a_m F^m + \dots + a_1 F + a_0 \text{Id}$ , denn  $a_0 = a_0 X^0$  und  $F^0 = \text{Id}$ . Aus dem gleichen Grund ist  $f(A) = a_m A^m + \dots + a_1 A + a_0 E_n$ , denn  $A^0 = E_n$ .

*Beweis.* Seien  $b_1, \dots, b_n$  die Vektoren der Basis  $B$ . Sei  $v \in V$  beliebig. Dann gibt es Skalare  $\lambda_1, \dots, \lambda_n \in k$  mit  $v = \sum_{i=1}^n \lambda_i b_i$ . Die Matrix  ${}_B M_B(F)$  ist so

definiert, dass  $F(v) = \sum_{i=1}^n \mu_i b_i$  gilt für

$$\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = A \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Also  $F^r(v) = \sum_{i=1}^n \nu_i b_i$  für

$$\begin{pmatrix} \nu_1 \\ \vdots \\ \nu_n \end{pmatrix} = A^r \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix},$$

und so  $f(F)(v) = \sum_{i=1}^n \sigma_i b_i$  für

$$\begin{pmatrix} \sigma_1 \\ \vdots \\ \sigma_n \end{pmatrix} = f(A) \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Diese Gleichung gilt auch dann, wenn man  $(\lambda_1, \dots, \lambda_n) \in k^n$  beliebig wählt und dann  $v := \sum_{i=1}^n \lambda_i b_i$  setzt. Also wie behauptet ist  $f(F) = 0$  genau dann, wenn  $f(A) = 0$  ist. Somit erfüllt die Menge  $I_F := \{f \in k[X] \mid f(F) = 0\}$  die Voraussetzungen von Lemma 10.4, außerdem ist  $I_F = I_A$ , wobei  $I_A$  das bereits untersuchte Ideal  $I_A = \{f \in k[X] \mid f(A) = 0\}$  ist. Also:  $F$  hat ein Minimalpolynom, genau wie  $A$ ; und es ist  $m_F = m_A$ . ■

## 11 Der Satz von Cayley-Hamilton

In diesem Kapitel wollen wir den folgenden Satz beweisen:

**Satz von Cayley–Hamilton** Sei  $k$  ein Körper und  $A \in M_n(k)$  eine quadratische Matrix. Dann gilt  $p_A(A) = 0$ , d.h. das charakteristische Polynom  $p_A(X)$  verschwindet in  $X = A$ .

In der Sprache von Endomorphismen heißt das: für jeden Endomorphismus  $F$  eines endlich-dimensionalen  $k$ -Vektorraums  $V$  gilt  $p_F(F) = 0$ .

Wir werden *drei* Beweise sehen. Der erste ist überschaubar und nachvollziehbar, funktioniert aber nur für  $k = \mathbb{C}$  und  $k = \mathbb{R}$ . Für den zweiten und den dritten müssen wir zuerst unsere Kenntnisse über Determinanten ausbauen. Haben wir dies einmal getan, so ist der zweite sehr kurz: aber es benutzt ein Trick, man ist zwar von der Richtigkeit des Beweises überzeugt, aber man versteht trotzdem nicht, *warum* das Ergebnis wahr ist. Der dritte Beweis ist etwas länger und zwingt uns dazu, den Begriff „Modul“ erstmals kennenzulernen.

Die Aussagen für Matrizen und für Endomorphismen sind bekanntlich äquivalent (vgl. Lemma 7.5 und Lemma 10.7).

### Der erste Beweis

Dieser Beweis benutzt die Tatsache, dass jede Matrix mit Einträgen aus  $\mathbb{C}$  mindestens einen Eigenwert hat, um den Satz per Induktion über  $n$  nachzuweisen. Das folgende Lemma wird benutzt für den Induktionsschritt.

**Lemma 11.1** Seien  $r, s \geq 1$ . Wir beschäftigen uns mit Blockmatrizen der Gestalt  $A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix} \in M_{r+s}(k)$ , wobei  $B \in M_r(k)$ ,  $D \in M_s(k)$  und  $C \in M(r \times s, k)$ .

a) Die charakteristischen Polynome der Matrizen  $A, B, D$  erfüllen die Gleichung

$$p_A(X) = p_B(X)p_D(X).$$

b) Sind  $A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$  und  $A' = \begin{pmatrix} B' & C' \\ 0 & D' \end{pmatrix}$  zwei Matrizen dieser Blockgestalt, so ist

$$AA' = \begin{pmatrix} BB' & BC' + CD' \\ 0 & DD' \end{pmatrix}.$$

c) Für jedes Polynom  $f \in k[X]$  gibt es eine Matrix  $K \in M(r \times s, k)$  derart, dass die folgende Gleichung gilt:

$$f(A) = \begin{pmatrix} f(B) & K \\ 0 & f(D) \end{pmatrix}.$$

*Beweis.* a) Sei  $n = r + s$ . Es ist

$$XE_n - A = \begin{pmatrix} XE_r - B & -C \\ 0 & XE_s - D \end{pmatrix},$$

weshalb gilt nach der Determinantenregel für Blockmatrizen

$$p_A(X) = \det(XE_n - A) = \det(XE_r - B) \cdot \det(XE_s - D) = p_B(X)p_D(X).$$

b) Dies lässt sich nachrechnen.

c) Gilt für  $f = X^m$  für alle  $m \geq 0$ , wegen (b) per Induktion über  $m$ . Deshalb gilt die Aussage für alle Polynome. ■

1. *Beweis des Satzes von Cayley–Hamilton.* In diesem Beweis setzen wir voraus, es ist  $k = \mathbb{C}$  oder  $k = \mathbb{R}$ . Jede Matrix mit Einträgen aus  $\mathbb{R}$  ist gleichzeitig eine Matrix mit Einträgen aus  $\mathbb{C}$ , und das charakteristische Polynom ändert sich nicht, ich eine Matrix  $A \in M_n(\mathbb{R})$  stattdessen als eine Matrix  $A \in M_n(\mathbb{C})$  betrachte. Es reicht also, den Fall  $k = \mathbb{C}$  zu behandeln.

Sei also  $F$  ein Endomorphismus eines endlich dimensionalen  $\mathbb{C}$ -Vektorraums  $V$ . Nach dem Fundamentalsatz der Algebra hat das charakteristische Polynom  $p_F(X)$  mindestens eine Nullstelle  $\lambda \in \mathbb{C}$ , d.h.  $\lambda$  ist ein Eigenwert von  $F$ . Sei  $B: b_1, \dots, b_n$  eine Basis von  $V$  derart, dass  $b_1$  ein Eigenvektor zum Eigenwert  $\lambda$  ist. Sei  $A$  die Matrix  $A = {}_B M_B(F)$ . Dann ist  $A$  vom Blockgestalt  $A = \begin{pmatrix} \lambda & C \\ 0 & D \end{pmatrix}$  mit  $D \in M_{n-1}(\mathbb{C})$  und  $C \in M(1 \times n-1, \mathbb{C})$ . Nach Lemma 11.1 (a) ist  $p_A(X) = (X - \lambda)p_D(X)$ , und deshalb  $p_A(A) = (A - \lambda E_n)p_D(A)$ . Nun,  $A - \lambda E_n = \begin{pmatrix} 0 & C \\ 0 & D' \end{pmatrix}$  für  $D' = D - \lambda E_{n-1}$ ; und nach Lemma 11.1 (c) gibt es ein  $K \in M(1 \times n-1, \mathbb{C})$  mit  $p_D(A) = \begin{pmatrix} \mu & K \\ 0 & 0 \end{pmatrix}$  für  $\mu = p_D(\lambda)$ , denn nach Induktionsannahme ist  $p_D(D) = 0$ . Also

$$p_A(A) = \begin{pmatrix} 0 & C \\ 0 & D' \end{pmatrix} \begin{pmatrix} \mu & K \\ 0 & 0 \end{pmatrix} = 0, \quad \text{nach Lemma 11.1 (b).}$$

Der Induktionsanfang ist der Fall  $n = 1$ . Dieser Fall ist klar. ■

*Bemerkung* Damit man diese Beweismethode für den allgemeinen Fall beweisen können, muss man folgendes zeigen:

Zu jedem Körper  $k$  und zu jedem normierten Polynom  $f \in k[X]$  gibt es einen größeren Körper  $K \supseteq k$  und ein  $\lambda \in K$  derart, dass  $f(\lambda) = 0$  gilt.

Diese Aussage stimmt zwar, ihr Beweis kommt aber erst in der Hauptstudiums-Vorlesung Algebra 1.

## Die Determinante für Matrizen über Ringe

In diesem Abschnitt sei  $R$  ein kommutativer Ring, d.h.  $rs = sr$  für alle  $r, s \in R$ . Sei  $A \in M_n(R)$  eine quadratische Matrix mit Einträgen aus  $R$ .

In Kapitel 6 konstruierten wir die Determinante einer Matrix mit Einträgen aus einem Körper. Der größte Teil der Konstruktion gilt auch für Einträge aus einem kommutativen Ring.

*Definition* Sei  $R$  ein kommutativer Ring. Eine Abbildung  $\Delta: M_n(R) \rightarrow k$  heißt eine *Determinantenfunktion*, wenn folgende drei Bedingungen erfüllt sind:

(D1)  $\Delta(A)$  ist  $n$ -fach multilinear in den Zeilen, d.h. für alle  $1 \leq i \leq n$  ist

$$\Delta(\dots, \underbrace{rv + sw}_{\text{ite Zeile}}, \dots) = r\Delta(\dots, \underbrace{v}_{\text{ite Zeile}}, \dots) + s\Delta(\dots, \underbrace{w}_{\text{ite Zeile}}, \dots)$$

für alle  $\lambda, \mu \in R$  und für alle Zeilenvektoren  $v, w \in R^n$ .

(D2)  $\Delta(A)$  ist alternierend: sind zwei Zeilen gleich, so ist  $\Delta(A) = 0$ .

(D3) Normierung:  $\Delta(E_n) = 1$ .

**Lemma 11.2** *Sei  $R$  ein kommutativer Ring. Es gibt genau eine Determinantenfunktion auf  $M_n(R)$ , nämlich*

$$\det(A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) A_{1\sigma(1)} A_{2\sigma(2)} \cdots A_{n\sigma(n)}$$

*Diese Funktion erfüllt die Determinanten-Eigenschaften (D1)–(D7) sowie (D9)–(D11). Auch Laplace-Entwicklung (s. Satz 6.5) gilt für diese Determinantenfunktion. Als Ersatz für (D8) erhält man*

(D8')  *$A$  ist genau dann invertierbar in  $M_n(R)$ , wenn es ein  $r \in R$  gibt mit  $r \det(A) = 1$ .*

*Beweis.* Bitte nehmen Sie Kapitel 6 von der LAAG1 zur Hand. Genau wie dort (Lemma 6.1) weist man als erstes die Eigenschaften (D4) und (D5) nach, dann überlegt man, dass die angegebene Funktion  $\det$  die einzige mögliche Determinantenfunktion ist. Mittels der Eigenschaften des Vorzeichens einer Permutation zeigt man dann, dass  $\det$  tatsächlich eine Determinantenfunktion ist. Als nächstes weist man die Eigenschaften (D6), (D7), (D10) und (D11) wie in Lemma 6.4 nach, danach die Laplace-Entwicklung (Satz 6.5).

Im Beweis von Lemma 6.4 benutzt man Gaußsche Elimination für Eigenschaft (D8). Dies benutzt wiederum die Tatsache, dass man in einem Körper jeden Bruch  $\frac{a}{b}$  für  $b \neq 0$  bilden kann. In einem kommutativen Ring ist das nicht so. Der dortige Beweis der Produktregel (D9) benutzt wiederum (D8), wir benötigen also einen neuen Beweis.

(D9): Die Produktregel (D9) besagt: es ist  $\det(AB) = \det(A) \det(B)$ . Sei  $S$  der Ring aller Polynome mit Koeffizienten aus  $\mathbb{Z}$  in  $2n^2$  Unbestimmten  $X_{ij}, Y_{ij}$  für  $i, j \in \{1, \dots, n\}$ . Somit ist z.B.  $3X_{13}Y_{21}^2Y_{22} - 2X_{12}^5Y_{11} + 4X_{21} - 5Y_{31}$  ein Element aus  $S$  für  $n = 3$ . Sei  $X \in M_n(S)$  bzw.  $Y \in M_n(S)$  die Matrix, dessen  $(i, j)$ -Eintrag der Unbestimmte  $X_{ij}$  bzw.  $Y_{ij}$  ist. Es reicht zu zeigen, dass  $\det(XY) = \det(X) \det(Y)$  ist, denn dann können wir Werte aus  $R$  für  $X_{ij}, Y_{ij}$  einsetzen und die Gleichung gilt weiterhin. Um  $\det(XY) = \det(X) \det(Y)$  nachzuweisen, beobachtet man, dass alle Brüche  $\frac{f}{g}$  mit  $f, g \in S$  und  $g \neq 0$  einen Körper  $k$  bilden (vgl. rationale Funktionen aus der Analysis), und dass dieser Körper  $S$  enthält. Dann sind wir aber fertig, denn die Produktregel gilt schon für Matrizen mit Einträgen aus einem Körper.

(D8'): Ist  $A$  invertierbar, so ist  $r \det(A) = 1$  für  $r = \det(A^{-1})$ , aufgrund der Produktregel. Gibt es ein  $r \in R$  mit  $r \det(A) = 1$ , so ist  $BA = AB = E_n$  für  $B = r \operatorname{Adj}(A)$ , nach der Produktregel und Lemma 11.3 unten. Erst im nächsten Abschnitt wird die Adjunkte  $\operatorname{Adj}(A)$  eingeführt. ■

## Die Adjunkte

*Definition* Sei  $R$  ein kommutativer Ring und  $A \in M_n(R)$  eine quadratische Matrix. Für  $i, j \in \{1, \dots, n\}$  sei  $A(i, j) \in M_{n-1}(R)$  wie in Satz 6.5 die Matrix, die entsteht, wenn man aus  $A$  die  $i$ te Zeile und die  $j$ te Spalte entfernt. Die *Adjunkte*  $\operatorname{Adj}(A) \in M_n(R)$  ist per Definition die durch

$$(\operatorname{Adj} A)_{ij} := (-1)^{i+j} \det A(j, i)$$

gegebene Matrix. Ein anderer Name ist der *komplementäre* Matrix zu  $A$ .

**Lemma 11.3** *Sei  $R$  ein kommutativer Ring, und sei  $A \in M_n(R)$ . Dann*

$$A \cdot \operatorname{Adj}(A) = \operatorname{Adj}(A) \cdot A = \det(A) E_n.$$

*Beweis.* Es ist

$$(A \cdot \operatorname{Adj}(A))_{ik} = \sum_{j=1}^n A_{ij} \operatorname{Adj}(A)_{jk} = \sum_{j=1}^n (-1)^{j+k} A_{ij} \det A(k, j).$$

Für  $k = i$  ist dies  $\det(A)$  nach Satz 6.5 (a). Nun sei  $k \neq i$ . Wir müssen zeigen, dass 0 als Ergebnis rauskommt. Sei  $B \in M_n(R)$  die Matrix, die aus  $A$  entsteht, wenn man die  $k$ te Zeile durch eine Kopie der  $i$ ten Zeile ersetzt: wegen (D2) ist  $\det(B) = 0$ . Also (Laplace-Entwicklung nach der  $k$ ten Zeile)

$$0 = \sum_{j=1}^n (-1)^{k+j} B_{kj} \det B(k, j) = (-1)^{i+k} \sum_{j=1}^n (-1)^{i+j} A_{ij} \det A(k, j),$$

denn nach Konstruktion ist  $B_{kj} = A_{ij}$  und  $B(k, j) = A(k, j)$ .

Wir haben also gezeigt, dass  $A \cdot \operatorname{Adj}(A) = \det(A) E_n$ . Der zweite Teil wird ähnlich gezeigt, diesmal mit Laplace-Entwicklung nach Spalten. ■

*Beispiel* Für  $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 1 & 2 \\ 2 & -3 & -4 \end{pmatrix} \in M_3(\mathbb{Z})$  ist  $\text{Adj}A = \begin{pmatrix} 2 & -1 & 1 \\ 20 & -10 & 10 \\ -14 & 7 & -7 \end{pmatrix}$ .

Es ist  $\det(A) = 0$  und  $A \cdot \text{Adj}(A) = \text{Adj}(A) \cdot A = 0$ .

*Beispiel* Für  $A = \begin{pmatrix} 3 & X-2 & X^2+1 \\ 0 & X & 1 \\ 0 & 0 & X-2 \end{pmatrix} \in M_3(\mathbb{R}[X])$  ist

$$\text{Adj}(A) = \begin{pmatrix} X(X-2) & -(X-2)^2 & -X^3-2 \\ 0 & 3(X-2) & -3 \\ 0 & 0 & 3X \end{pmatrix}.$$

Es ist  $\det(A) = 3X(X-2)$  und  $A \cdot \text{Adj}(A) = \text{Adj}(A) \cdot A = 3X(X-2)E_3$ .

*Beispiel* Als Element von  $M_2(\mathbb{Z})$  ist  $A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$  nicht invertierbar, denn  $\det(A) = 2$  und  $\frac{1}{2} \notin \mathbb{Z}$ . Zwar gibt es eine Inverse-Matrix, diese liegt aber nicht in  $M_2(\mathbb{Z})$ .

Dagegen ist  $B = \begin{pmatrix} 3 & 4 \\ 4 & 5 \end{pmatrix}$  in  $M_2(\mathbb{Z})$  invertierbar, denn  $\det(B) = -1$  und für  $r = -1$  ist  $r \det(B) = 1$ . Es ist  $B^{-1} = -\text{Adj}(B) = \begin{pmatrix} -5 & 4 \\ 4 & -3 \end{pmatrix}$ .

*Bemerkung* Für große Matrizen dauert die Berechnung der Adjunkte zu lang; man sollte die Inverse-Matrix eher mittels Gaußschen Elimination ermitteln. Das Nutzen der Adjunkte ist eher für theoretische Fragen.

2. *Beweis des Satzes von Cayley-Hamilton.* Sei  $A \in M_n(k)$  eine Matrix. Zu jedem  $m \geq 0$  gibt es eine Matrix  $B \in M_n(k[X])$  mit  $A^m - X^m E_n = B \cdot (A - X E_n)$ , und zwar  $B = \sum_{i=0}^{m-1} A^i X^{m-1-i}$ . Nun sei  $f \in k[X]$  ein Polynom,  $f = \sum_{r=0}^m c_r X^r$ . Dann

$$f(A) - f(X E_n) = \sum_{r=0}^m c_r (A^r - X^r E_n),$$

weshalb es eine Matrix  $C \in M_n(k[X])$  gibt mit  $f(A) - f(X E_n) = C \cdot (A - X E_n)$ , d.h.

$$f(A) = f(X) E_n + C \cdot (A - X E_n).$$

Dies wenden wir im Fall  $f = p_A$  an. Es ist

$$p_A(X) E_n = \det(X E_n - A) E_n = \text{Adj}(X E_n - A) \cdot (X E_n - A).$$

Mit  $D = C - \text{Adj}(X E_n - A) \in M_n(k[X])$  ist also

$$p_A(A) = D \cdot (A - X E_n).$$

Diese ist eine Gleichung in  $M_n(k[X])$ , auf der linken Seite kommt allerdings  $X$  gar nicht vor. Es folgt hieraus, dass  $D = 0$  und deshalb  $p_A(A) = 0$ , wie erwünscht. Ist  $D \neq 0$ , so gibt es ein  $m \geq 0$  und Matrizen  $D_0, D_1, \dots, D_m \in M_n(k)$  mit  $D = \sum_{i=0}^m D_i X^i$  und  $D_m \neq 0$ : dann ist  $X^m$  die höchste Potenz von  $X$ , die in  $D$  vorkommt. Die Beweismethode von Lemma 10.2 zeigt hier, dass

$$D \cdot (A - XE_n) = -D_m X^{m+1} + \text{Terme vom Grad } \leq m.$$

Dann wäre aber  $D \cdot (X - XE_n)$  kein Element von  $M_n(k)$ , ein Widerspruch. Also  $D = 0$  und  $p_A(A) = 0$ . ■

### Die 3. Beweismethode und der Modulbegriff

#### Satz von Cayley–Hamilton für Matrizen über kommutative Ringe

Sei  $R$  ein kommutativer Ring und  $A \in M_n(R)$  eine quadratische Matrix. Dann gilt  $p_A(A) = 0$ , d.h. das charakteristische Polynom  $p_A(X)$  verschwindet in  $X = A$ .

3. *Beweis des Satzes von Cayley–Hamilton.* Für ein Polynom  $f \in R[X]$  und ein (Spalten-)vektor  $v \in R^n$  werden wir  $f * v$  für das Element  $f(A) \cdot v \in R^n$  schreiben. Somit haben wir eine Abbildung  $R[X] \times M \rightarrow M$ ,  $(f, v) \mapsto f * v$ , wobei ich  $M := R^n$  setze. Ist nun  $B$  eine Matrix aus  $M_n(R[X])$  und  $\underline{v} = (v_1, \dots, v_n) \in M^n$ , so können wir durch Matrixmultiplikation das Produkt  $B \cdot \underline{v} \in M^n$  bilden: es ist  $B \cdot \underline{v} = \underline{w}$  für  $\underline{w} = (w_1, \dots, w_n)$  mit  $w_i = \sum_{j=1}^n B_{ij} * v_j$ . Beachten Sie, dass  $B_{ij} \in R[X]$  liegt und  $w_i, v_j \in M = R^n$ .

Betrachten wir jetzt den Fall  $B = XE_n - A^T$  und  $\underline{v} = (e_1, \dots, e_n)$ , wobei  $e_i \in R^n$  durch  $e_i = (0, \dots, 1, \dots, 0)$  mit der 1 an der  $i$ ten Stelle gegeben ist. Es ist  $B_{ij} = X\delta_{ij} - A_{ji}$ , also  $B \cdot \underline{v} = \underline{w}$  für

$$w_i = \sum_{j=1}^n B_{ij} * e_j = X * e_i - \sum_{j=1}^n A_{ji} e_j = A \cdot e_i - \sum_{j=1}^n A_{ji} e_j = 0.$$

Das heißt, es ist  $B \cdot \underline{v} = 0$ . Jetzt multiplizieren wir diese Gleichung von Links mit  $\text{Adj}(B)$ : es ist  $\text{Adj}(B) \cdot B = \det(B)E_n$ , weshalb  $\det(B) * e_i = 0$  für jedes  $1 \leq i \leq n$ . Wegen  $\det(B) = p_{A^T}(X)$  ist  $\det(B) * e_i = p_{A^T}(A) \cdot e_i$ , also  $p_{A^T}(A) \cdot e_i = 0$  für alle  $i$ . Hieraus folgt, dass  $p_{A^T}(A) = 0$ . Es ist aber  $p_A = p_{A^T}$ , denn  $p_A = \det(XE_n - A) = \det(XE_n - A)^T = \det(XE_n - A^T)$ . ■

*Bemerkung* Moduln sind – mit einigen Abstrichen – das für Ringe, was Vektorräume für Körper sind. Im obigen Beweis kommen zwei Moduln vor: zuerst machten wir  $M = R^n$  durch  $*$  zu einem  $R[X]$ -Modul, dann machten wir  $M^n$  durch Matrixmultiplikation zu einem  $M_n(R[X])$ -Modul.



## 12 Algebraische Strukturen: Quotienten

Der Restklassenring  $\mathbb{Z}/6$  sowie die Körper  $\mathbb{F}_2$  und  $\mathbb{F}_3$  sind eigentlich als Quotientenringen zu konstruieren. Es gibt auch Quotientenvektorräume: ist etwa  $F$  ein Endomorphismus von  $V$  und  $U \subseteq V$  ein Unterraum, der  $F(U) \subseteq U$  erfüllt –  $U$  könnte z.B. ein Eigenraum von  $F$  sein –, so induziert  $F$  einen Endomorphismus des Quotientenraums  $V/U$ . Hat die Matrix von  $F$  die Blockgestalt  $\begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$  bezüglich einer Basis, die mit einer Basis von  $U$  anfängt, so ist  $B$  die Matrix von  $F$  eingeschränkt auf  $U$ , und  $D$  ist die Matrix des induzierten Endomorphismus von  $V/U$ . Dagegen gibt es im allgemeinen Fall kein Komplement  $W$  von  $U$  derart, dass  $F(W) \subseteq W$  und  $D$  die Matrix von  $F$  auf  $W$  ist.

Um diesen beiden Beispielen möglichst einheitlich behandeln zu können, fangen wir mit Quotienten von abelschen Gruppen an. Der Fall von *beliebigen* Gruppen ist etwas komplizierter, denn dort kann man nicht für jede Untergruppe  $H \leq G$  eine Quotientengruppe  $G/H$  bilden.

Zur Erinnerung: eine additive Gruppe ist eine abelsche Gruppe, deren Gruppenoperation mit  $+$  und deren neutrales Element mit  $0$  bezeichnet wird.

*Definition* Sei  $G$  eine Gruppe mit Gruppenoperation  $*$ . Eine Teilmenge  $H \subseteq G$  heißt genau dann eine Untergruppe von  $G$ , wenn auch  $H$  bezüglich  $*$  eine Gruppe ist. Dies bedeutet: für alle  $h_1, h_2 \in H$  ist auch  $h_1 * h_2 \in H$ ; das neutrale Element  $e$  liegt in  $H$ ; und für jedes  $h \in H$  liegt auch das Inverse  $h'$  in  $H$ . Bezeichnung:  $H \leq G$ .

**Lemma 12.1** *Sei  $A$  eine additive Gruppe und  $H \leq A$  eine Untergruppe. Sei  $\sim$  die Relation auf  $A$  gegeben durch: es ist  $a \sim b$  genau dann, wenn  $b - a$  in  $H$  liegt. Dann gelten:*

a)  $\sim$  ist eine Äquivalenzrelation auf  $A$ .

*Die Äquivalenzklasse von  $a \in A$  bezeichnet man mit  $a + H$ . Man nennt sie die Nebenklasse von  $a$  bezüglich  $H$ , oder die Restklasse von  $a$  modulo  $H$ .*

b) Ist  $a_1 \sim a_2$  und  $b_1 \sim b_2$ , so ist auch  $a_1 + b_1 \sim a_2 + b_2$ .

c) Die Menge  $A/\sim$  der Äquivalenzklassen bildet eine additive Gruppe bezüglich der Operation  $(a + H) + (b + H) := (a + b) + H$ . Diese Gruppe nennt man die Quotientengruppe  $A/H$ .

d) Die Abbildung  $p: A \rightarrow A/H$ ,  $a \mapsto a + H$  ist ein surjektiver Gruppenhomomorphismus. Manchmal bezeichnet man diese Abbildung als die kanonische Projektion.

*Beweis.* a) Reflexiv:  $a \sim a$ , denn  $a - a = 0 \in H$ . Symmetrisch: Ist  $a \sim b$ , dann  $b - a \in H$ , also  $a - b = -(b - a) \in H$ , also  $b \sim a$ . Transitiv: Ist  $a \sim b$  und  $b \sim c$ , dann  $b - a, c - b \in H$ , also  $c - a = (c - b) + (b - a) \in H$ , also  $a \sim c$ .

b) Es ist  $(a_2 + b_2) - (a_1 + b_1) = (a_2 - a_1) + (b_2 - b_1) \in H$ .

c) Wegen b) ist die Operation  $(a + H) + (b + H) := (a + b) + H$  repräsentantenunabhängig. Assoziativ:

$$((a + H) + (b + H)) + (c + H) = ((a + b) + H) + (c + H) = ((a + b) + c) + H.$$

Analog ist  $(a + H) + ((b + H) + (c + H)) = (a + (b + c)) + H$ . Aber  $a + (b + c) = (a + b) + c$ , da  $A$  assoziativ ist. Das neutrale Element ist  $0 + H$ , und  $-(a + H) = (-a) + H$ . ■

**Korollar 12.2** a) Sei  $R$  ein kommutativer Ring und  $I \subseteq R$  ein Ideal in  $R$ , d.h.  $I$  ist eine Teilmenge von  $R$  mit den folgenden drei Eigenschaften:

(I1) Aus  $a, b \in I$  folgt  $a + b \in I$ ;

(I2) Aus  $r \in R, a \in I$  folgt  $ra \in I$ ;

(I3')  $0 \in I$ .

Dann ist  $I \leq R$ , und die Quotientengruppe ist bezüglich der Multiplikation  $(r + I)(s + I) := rs + I$  selbst ein kommutativer Ring, der Quotientenring  $R/I$ .

Nun sei  $U$  ein Untervektorraum des  $k$ -Vektorraums  $V$ .

b) Es ist  $U \leq V$ , und die Quotientengruppe ist bezüglich der Skalarmultiplikation  $\lambda(v + U) := (\lambda v) + U$  selbst ein  $k$ -Vektorraum, der Quotientenvektorraum  $V/U$ . In diesem Fall ist die kanonische Projektion  $V \rightarrow V/U, v \mapsto v + U$  eine surjektive lineare Abbildung.

c) Ist  $V$  endlich dimensional, so ist  $\dim V/U = \dim V - \dim U$ . Außerdem gilt: ist  $v_1, \dots, v_n$  eine Basis von  $V$  derart, dass die ersten  $r$  Elemente eine Basis von  $U$  bilden, so ist  $v_{r+1} + U, v_{r+2} + U, \dots, v_n + U$  eine Basis von  $V/U$ .

d) Sei  $F$  ein Endomorphismus des  $k$ -Vektorraums  $V$ . Sei  $U \subseteq V$  ein Unterraum, der bezüglich  $F$  invariant ist, d.h. es ist  $F(U) \subseteq U$ . Dann: durch  $\bar{F}(v + U) = F(v) + U$  induziert  $F$  einen Endomorphismus  $\bar{F}$  des Quotientenvektorraums.

*Beweis.* a) Für  $a, b \in I$  sind  $a + b$  und  $-a = (-1)a$  auch Elemente von  $I$ , also ist  $I \leq R$ . Ist  $r \sim r'$  und  $s \sim s'$ , so gibt es  $a, b \in I$  mit  $r' = r + a, s' = s + b$ . Also  $r's' - rs = rb + sa + ab$ . Wegen (I2) liegen  $rb, sa, ab$  in  $I$ . Wegen (I1) ist also  $r's' \sim rs$ . Die Multiplikation ist also wohldefiniert. Dass die Axiome für einen kommutativen Ring erfüllt sind, folgt jetzt aus den gleichen Axiomen für  $R$ : wie es für Assoziativität im Beweis des Lemmas der Fall war.

- b) Vektorräume sind insbesondere additive Gruppen, also  $U \leq V$ . Ist  $v' = v + u$  für  $u \in U$ , dann  $\lambda v' - \lambda v = \lambda u \in U$ . Somit ist die Skalarmultiplikation wohldefiniert.

Die kanonische Projektion ist bekanntlich surjektiv. Linear:  $\lambda v + \mu w \mapsto (\lambda v + \mu w) + U = \lambda(v + U) + \mu(w + U)$ .

- c) Dimension von  $V/U$ : die Projektion hat Bild  $V/U$  und Kern  $U$ . Man wendet also die Dimensionsformel an.

Basis: Aus Dimensionsgründen reicht es, zu zeigen, dass die angebliche Basis ein Erzeugendensystem von  $V/U$  ist. Sei  $v + U$  ein beliebiges Element von  $V/U$ . Da  $v_1, \dots, v_n$  eine Basis von  $V$  ist, gibt es  $\lambda_1, \dots, \lambda_n \in k$  mit  $v = \sum_{i=1}^n \lambda_i v_i$ . Also  $v + U = \sum_{i=1}^n \lambda_i (v_i + U)$ . Nun,  $v_i + U = 0$  für  $i \leq r$ , denn  $v_i \in U$  für solches  $i$ . Also  $v + U = \sum_{i=r+1}^n \lambda_i (v_i + U)$ .

- d)  $\bar{F}$  ist repräsentantenunabhängig: ist  $v + U = v' + U$ , so gibt es ein  $u \in U$  mit  $v' = v + u$ . Also  $\bar{F}(v' + U) = F(v') + U = F(v) + F(u) + U = F(v) + U = \bar{F}(v + U)$ , denn  $F(u) \in U$ . Da  $F$  linear ist, ist auch  $\bar{F}$ . ■

*Beispiel: Restklassen modulo  $n$*  Sei  $n \geq 2$  eine ganze Zahl. Mit  $n\mathbb{Z}$  bezeichnet man die Menge

$$n\mathbb{Z} = \{nm \mid m \in \mathbb{Z}\} = \{r \in \mathbb{Z} \mid n \text{ teilt } r\}.$$

Dieses  $n\mathbb{Z}$  ist ein Ideal in  $\mathbb{Z}$ . Die Nebenklassen  $a + n\mathbb{Z}$  heißen Restklassen modulo  $n$ , und  $\mathbb{Z}/n\mathbb{Z}$  heißt der Restklassenring modulo  $n$ . Weiter Bezeichnungen für diesen Ring sind  $\mathbb{Z}/n$  und  $\mathbb{Z}_n$ ; allerdings bezeichnete  $\mathbb{Z}_n$  manchmal etwas anderes.

Der Ring  $\mathbb{Z}/n\mathbb{Z}$  hat  $n$  Elemente. Sie sind  $0 + n\mathbb{Z}, 1 + n\mathbb{Z}, 2 + n\mathbb{Z}, \dots, (n-1) + n\mathbb{Z}$ .

In  $\mathbb{Z}/6\mathbb{Z}$  gilt  $(2 + \mathbb{Z})(3 + \mathbb{Z}) = 6 + \mathbb{Z} = 0 + \mathbb{Z}$ . Es kann also sein, dass  $ab = 0$  gilt, obwohl  $a \neq 0$  und  $b \neq 0$  gelten.

*Der Körper  $\mathbb{F}_p$*  Sei  $p$  eine Primzahl. Dann ist der Restklassenring  $\mathbb{Z}/p\mathbb{Z}$  ein Körper, den man mit  $\mathbb{F}_p$  bezeichnet: der Körper mit  $p$  Elementen.

$\mathbb{F}_p$  ist ein Körper, denn es ist ein kommutativer Ring;  $1 \neq 0$ ; und ist  $a + p\mathbb{Z} \neq 0$ , so gibt es ein  $b$  mit  $(a + p\mathbb{Z})(b + p\mathbb{Z}) = 1 + p\mathbb{Z}$ . Dies ist der Fall, denn die Abbildung  $b + p\mathbb{Z} \mapsto ab + p\mathbb{Z}$  von der  $p$ -elementigen Menge  $\mathbb{Z}/p\mathbb{Z}$  nach sich selbst ist injektiv und deshalb surjektiv; die Abbildung ist injektiv, denn aus  $ab + p\mathbb{Z} = ac + p\mathbb{Z}$  folgt, dass  $a(c - b)$  durch  $p$  teilbar ist. Da  $p$  eine Primzahl ist und  $a$  nicht teilt, muss also  $c - b$  durch  $p$  teilbar sein, d.h.  $b + p\mathbb{Z} = c + p\mathbb{Z}$ .

Die Körper  $\mathbb{F}_2$  und  $\mathbb{F}_3$  haben wir bereits gesehen. Der Körper  $\mathbb{F}_4$  aus Übungsblatt 3 gehört nicht zu dieser Familie von Beispielen, denn 4 ist keine Primzahl.

*Beispiel: Invariante Unterräume* Sei  $F$  ein Endomorphismus des  $k$ -Vektorraums  $V$ , und sei  $U \subseteq V$  ein invarianter Unterraum. Nach dem Korollar induziert

$F$  einen Endomorphismus  $\bar{F}$  des Quotientenraums  $V/U$  durch  $\bar{F}(v+U) = F(v) + U$ .

Jeder Unterraum von  $V$  hat mindestens ein Komplement  $W$ , d.h.  $W$  ist selbst ein Unterraum von  $V$ , und es ist  $V = U \oplus W$ . Für jedes solche Komplement ist die lineare Abbildung  $W \rightarrow V/U$ ,  $w \mapsto w + U$  ein Isomorphismus. Ist  $W$  selbst invariant, so ist  $\bar{F}$  nichts anderes als die Einschränkung von  $F$  auf  $W$ .

In den meisten Fällen gibt es aber kein invariantes Komplement  $W$ . Ist zum Beispiel  $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  die Abbildung  $F(x, y) = (x + y, y)$ , so ist  $U = \{(x, 0) \mid x \in \mathbb{R}\}$  ein invarianter Unterraum. Jedes Komplement  $W$  von  $U$  in  $\mathbb{R}^2$  ist eindimensional mit Basis  $w = (x, 1)$  für ein  $x \in \mathbb{R}$ ; aber  $F(w) = w + (1, 0) \notin \text{Spann}(w)$ .

In diesem Beispiel ist  $\mathbb{R}^2/U$  eindimensional mit Basis  $(0, 1) + U$ . Es ist

$$\bar{F}: (0, 1) + U \mapsto (1, 1) + U = (0, 1) + U.$$

Somit ist  $\bar{F}$  die Identitätsabbildung auf  $\mathbb{R}^2/U$ .

*Invariante Unterräume und Matrizen* Sei  $F$  ein Endomorphismus eines endlich dimensionalen  $k$ -Vektorraums  $V$ , und sei  $U \subseteq V$  ein invarianter Unterraum. Wir wählen eine Basis  $t_1, \dots, t_r$  für  $U$  und setzen dies zu einer Basis  $T: t_1, \dots, t_n$  von  $V$  fort. Die Matrix von  $F$  hat dann die Blockgestalt

$${}_T M_T(F) = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix} \quad (*)$$

mit  $B \in M_r(k)$  die Matrix von  $F|_U$  und  $D \in M_{n-r}(k)$ . Es ist dann  $D$  die Matrix des Endomorphismus  $\bar{F}$  von  $V/U$  bezüglich der Basis  $t_{r+1} + U, \dots, t_n + U$  von  $V/U$ .

Berechnen wir jetzt das charakteristische Polynom von beiden Seiten der Gleichung (\*). Wir erhalten  $p_F(X) = p_B(X)p_D(X)$ . Da  $B$  bzw.  $D$  die Matrix von  $F|_U$  bzw.  $\bar{F}$  ist, haben wir das folgende Lemma bewiesen:

**Lemma 12.3** *Sei  $F$  ein Endomorphismus des endlich dimensionalen  $k$ -Vektorraums  $V$ , sei  $U \subseteq V$  ein invarianter Unterraum, und sei  $\bar{F}$  der induzierte Endomorphismus des Quotientenraums  $V/U$ . Dann erfüllen die charakteristischen Polynome der drei Endomorphismen  $F$ ,  $F|_U$  und  $\bar{F}$  die Gleichung  $p_F(X) = p_{F|_U}(X)p_{\bar{F}}(X)$ . ■*

## Äquivalente und ähnliche Matrizen

Die Matrix einer linearen Abbildung hängt bekanntlich von der gewählten Basis ab. Zwei Matrizen der gleichen linearen Abbildung dürfen sich nicht zu stark voneinander unterscheiden.

*Definition* Sei  $k$  ein Körper.

- a) Zwei Matrizen  $A, B \in M(m \times n, k)$  heißen *äquivalent*, wenn es invertierbare Matrizen  $P \in M_m(k)$  und  $Q \in M_n(k)$  gibt derart, dass  $B = PAQ$  gilt.
- b) Zwei Matrizen  $A, B \in M_n(k)$  heißen *ähnlich*, wenn es eine invertierbare Matrix  $T \in M_n(k)$  gibt derart, dass  $B = TAT^{-1}$  gilt.

**Lemma 12.4** a) „äquivalent“ und „ähnlich“ sind Äquivalenzrelationen.

- b) Sei  $f: V \rightarrow W$  eine lineare Abbildung. Sind  $B, B'$  bzw.  $C, C'$  zwei Basen von  $V$  bzw.  $W$ , so sind die Matrizen  ${}_B M_C(f)$  und  ${}_{B'} M_{C'}(f)$  äquivalent. Umgekehrt gilt: ist die Matrix  $A$  zu  ${}_B M_C(f)$  äquivalent, so gibt es Basen  $B'$  für  $V$  und  $C'$  für  $W$  derart, dass  $A = {}_{B'} M_{C'}(f)$  ist.
- c) Sei  $f$  ein Endomorphismus des  $k$ -Vektorraums  $V$ . Sind  $B, B'$  zwei Basen von  $V$ , so sind die Matrizen  ${}_B M_B(f)$  und  ${}_{B'} M_{B'}(f)$  ähnlich. Umgekehrt gilt: ist die Matrix  $A$  zu  ${}_B M_B(f)$  ähnlich, so gibt es eine Basis  $B'$  für  $V$  derart, dass  $A = {}_{B'} M_{B'}(f)$  ist.

*Beweis.* a) Äquivalenz: Reflexiv:  $P = E_m, Q = E_n$ . Symmetrisch:  $B = P^{-1}AQ^{-1}$ . Transitiv: ist  $C = RBS$ , dann  $C = (RP)B(QS)$ . Ähnlichkeit: Reflexiv:  $T = E_n$ . Symmetrisch:  $B = T^{-1}AT$ . Transitiv: ist  $C = SBS^{-1}$ , dann  $C = (ST)B(ST)^{-1}$ .

- b) Es ist  ${}_{B'} M_{C'}(f) = P {}_B M_C(f) Q$  für  $P = {}_C M_{C'}(\text{Id}), Q = {}_{B'} M_B(\text{Id})$ . Diese Matrizen sind invertierbar, z.B.  $P^{-1} = {}_C M_C(\text{Id})$ . Nun sei  $A = P {}_B M_C(f) Q$ . Es ist  $Q = {}_{B'} M_B(\text{Id})$  für die Matrix  $B': v'_1, \dots, v'_n$  von  $V$  gegeben durch  $v'_i = \sum_{j=1}^n Q_{ji} v_j$ , wobei  $B$  die Basis  $v_1, \dots, v_n$  ist. Da  $B$  eine Basis ist, und die Spalten der invertierbaren Matrix  $Q$  linear unabhängig sind, ist  $B'$  linear unabhängig und deshalb eine Basis. Analog gibt es eine Basis  $C'$  von  $W$  mit  $P^{-1} = {}_C M_C(\text{Id})$ . Also  $A = {}_C M_{C'}(\text{Id}) {}_B M_C(f) {}_{B'} M_B(\text{Id}) = {}_{B'} M_{C'}(f)$ .
- c) Analog. ■

## Triangulierbarkeit

*Definition* Ein Endomorphismus  $F$  des endlich dimensionalen  $k$ -Vektorraums  $V$  heißt *triangulierbar*, wenn es eine Basis  $B$  von  $V$  gibt derart, dass die Matrix  ${}_B M_B(F)$  obere Dreiecksgestalt hat.

Entsprechend heißt eine Matrix  $A \in M_n(k)$  triangulierbar, wenn sie zu einer Matrix in oberer Dreiecksgestalt ähnlich ist. Nach Lemma 12.4 ist dies genau dann der Fall, wenn der Endomorphismus  $L_A$  des  $k^n$  triangulierbar ist.

*Beispiele* Diagonalmatrizen haben obere Dreiecksgestalt, somit ist jede diagonalisierbare Matrix triangulierbar. Etwa  $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{R})$ : es gibt zwei Eigenwerte, nämlich  $1, -1$ , also ist die Matrix diagonalisierbar und triangulierbar.

Die Matrix  $A = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \in M_2(\mathbb{R})$  dagegen ist nicht diagonalisierbar, denn  $p_A(X) = X^2$ , d.h. 0 ist der einzige Eigenwert, aber der Eigenraum ist nur eindimensional. Nun sei  $B: b_1, b_2$  die Basis von  $\mathbb{R}^2$  gegeben durch  $b_1 = (1, 1)$ ,  $b_2 = (1, 0)$ . Dann  $A \cdot b_1 = 0$  und  $A \cdot b_2 = b_1$ . Bezüglich der Basis  $B$  hat der Endomorphismus  $L_A$  deshalb die Matrix  ${}_B M_B(L_A) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ , d.h.  $A$  ist zu dieser Matrix ähnlich und deshalb triangulierbar.

Dagegen ist  $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in M_2(\mathbb{R})$  nicht einmal triangulierbar: denn jede Matrix in oberer Dreiecksgestalt hat in  $e_1$  einen Eigenvektor, somit muss jede triangulierbare Matrix mindestens einen Eigenvektor haben, aber  $A$  hat keine Eigenwerte, denn  $p_A(X) = X^2 + 1$  hat keine Nullstellen in  $\mathbb{R}$ .

**Satz 12.5** *Ein Endomorphismus  $F$  des  $n$ -dimensionalen  $k$ -Vektorraums  $V$  bzw. eine quadratische Matrix  $A \in M_n(k)$  ist genau dann triangulierbar, wenn das charakteristische Polynom in  $k[X]$  als ein Produkt von linearen Faktoren zerfällt:*

$$p_F(X) = \prod_{i=1}^n (X - \lambda_i), \quad \text{mit } \lambda_1, \dots, \lambda_n \in k.$$

Für  $k = \mathbb{C}$  ist dies immer der Fall.

*Beweis.* Fall  $k = \mathbb{C}$ : nach dem Fundamentalsatz der Algebra hat jedes nicht-konstante Polynom  $f \in \mathbb{C}[X]$  mindestens eine Nullstelle in  $\mathbb{C}$ . Per Induktion über  $\text{grad}(f)$  folgt es, dass jedes nichtkonstante Polynom in  $\mathbb{C}[X]$  sich als ein Produkt von linearen Faktoren schreiben lässt.

Hat die Matrix  $A \in M_n(k)$  obere Dreiecksgestalt, so ist  $p_A(X) = \prod_{i=1}^n (X - A_{ii})$  ein Produkt von linearen Faktoren. Ähnliche Matrizen haben das gleiche charakteristische Polynom. Somit ist für jede triangulierbare Matrix das charakteristische Polynom ein Produkt von linearen Faktoren.

Umgekehrt sei  $F$  ein Endomorphismus des  $n$ -dimensionalen Vektorraums  $V$ , dessen charakteristische Polynom als  $p_F(X) = \prod_{i=1}^n (X - \lambda_i)$  zerfällt. Sei  $U \subseteq V$  der Eigenraum  $U = E_{\lambda_1}(A)$ . Dann  $r := \dim(U) \geq 1$ . Dieser Unterraum ist invariant, d.h.  $F(U) \subseteq U$ , denn  $F(u) = \lambda_1 u$  für jedes  $u \in U$ . Nach Lemma 12.3 oben ist  $p_{\bar{F}}(X)$  ein Teiler von  $p_F(X)$  und deshalb selbst ein Produkt von linearen Faktoren. Nach Induktion über  $n = \dim V$  ist deshalb  $\bar{F}$  triangulierbar. Sei  $T: \bar{v}_{r+1}, \dots, \bar{v}_n$  eine Basis von  $V/U$ , die den Endomorphismus  $\bar{F}$  trianguliert, d.h. mit  ${}_T M_T(\bar{F})$  in oberer Dreiecksgestalt. Wählen wir dann  $v_{r+1}, \dots, v_n \in V$  derart, dass  $\bar{v}_i = v_i + U$  für  $i \geq r+1$ . Sei  $S: v_1, \dots, v_r$  eine Basis von  $U$ . Sei  $R$  die Basis  $v_1, \dots, v_n$  von  $V$ : um zu sehen, dass  $R$  eine Basis ist, beachten Sie, dass  $R$  linear unabhängig ist: ist  $\sum_{i=1}^n \lambda_i v_i = 0$ , dann  $\sum_{i=1}^n \lambda_i (v_i + U) = 0$ , also  $\sum_{i=r+1}^n \lambda_i \bar{v}_i = 0$ , also  $\lambda_i = 0$  für alle  $i \geq r+1$ , weshalb  $\sum_{i=1}^r \lambda_i v_i = 0$  und  $\lambda_i = 0$  für alle  $i$ .

Es ist dann  ${}_R M_R(F) = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$  mit  $B = {}_S M_S(F|_U)$  eine Diagonalmatrix, und  $D = {}_T M_T(\bar{F})$  in oberer Dreiecksgestalt. Somit ist auch  ${}_R M_R(F)$  in oberer Dreiecksgestalt. ■

**Korollar 12.6** Sei  $F$  ein Endomorphismus des endlich dimensionalen  $k$ -Vektorraums  $V$ . Die folgenden drei Aussagen sind äquivalent:

- a)  $F$  ist triangulierbar.
- b) Das charakteristische Polynom  $p_F(X)$  zerfällt als ein Produkt linearer Faktoren.
- c) Das Minimalpolynom  $m_F(X)$  zerfällt als ein Produkt linearer Faktoren.

*Zusatz:* Jede Nullstelle des charakteristischen Polynoms ist auch eine Nullstelle des Minimalpolynoms.

*Beweis.* Zusatz: Ist  $p_F(\lambda) = 0$ , so gibt es einen Eigenvektor  $v \neq 0$  mit  $F(v) = \lambda v$ . Dann  $0 = m_F(F)(v) = m_F(\lambda) \cdot v$ , weshalb  $m_F(\lambda) = 0$  sein muss, denn  $v \neq 0$ .

Nach dem Satz sind a) und b) äquivalent. Aus b) folgt c), denn  $m_F(X)$  teilt  $p_F(X)$  nach dem Satz von Cayley–Hamilton. Wir setzen jetzt c) voraus und werden per Induktion über  $\dim V$  zeigen, dass a) folgt. Wegen c) hat  $m_F(X)$  mindestens eine Nullstelle  $\lambda \in k$ . Also  $p_F(\lambda) = 0$ , denn  $m_F$  teilt  $p_F$ , also ist  $\lambda$  ein Eigenwert von  $F$ . Wie im Beweis des Satzes setzen wir  $U = E_\lambda(F)$ , ein invarianter Unterraum. Der eingeschränkte Endomorphismus  $F|_U$  ist diagonalisierbar. Für den induzierten Endomorphismus  $\bar{F}$  des Quotientenraums  $V/U$  gilt  $m_F(\bar{F}) = 0$ , denn für alle  $v + U \in V/U$  ist  $m_F(\bar{F})(v + U) = m_F(F)(v) + U = 0 + U$ . Somit ist das Minimalpolynom von  $\bar{F}$  ein Teiler von  $m_F$  und deshalb ein Produkt von linearen Faktoren. Nach Induktionsannahme ist  $\bar{F}$  triangulierbar. Wie im Beweis des Satzes folgt jetzt, dass auch  $F$  triangulierbar ist. ■

## 13 Algebraische und Geometrische Vielfachheit

### Die Vielfachheit einer Nullstelle

Sei  $k$  ein Körper und  $f \in k[X]$  ein Polynom mit  $f(a) = 0$  für ein  $a \in k$ . Was bedeutet es, wenn man behauptet, dass  $a$  eine 3-fache Nullstelle von  $f$  ist?

In der Analysis heißt das: es ist  $f(a) = f'(a) = f''(a) = 0$ . Meint man, dass die Vielfachheit der Nullstelle genau 3 beträgt und nicht noch höher ist, so fügt man  $f'''(a) \neq 0$  hinzu. Dies setzt voraus, dass  $k = \mathbb{R}$  oder  $\mathbb{C}$  ist.

In der Algebra heißt es, dass  $f(X)$  durch  $(X - a)^3$  teilbar ist. Meint man, dass die Vielfachheit genau 3 beträgt, verlangt man, dass  $f(X)$  genau dreimal durch  $(X - a)$  teilbar ist, d.h. es ist  $f(X) = (X - a)^3 g(X)$  für ein Polynom  $g$  mit  $g(a) \neq 0$ , was wiederum äquivalent dazu ist, dass  $g(X)$  nicht durch  $X - a$  teilbar ist.

Diese algebraische Charakterisierung der Vielfachheit hat den Vorteil, dass sie für jeden Körper funktioniert. Anders als im analytischen Fall ist es nicht sofort klar, dass die algebraische Charakterisierung eindeutig ist.

**Lemma 13.1** *Sei  $k$  ein Körper,  $a \in k$  ein Skalar, und  $0 \neq f \in k[X]$  ein Polynom.*

- a) *Es ist  $f(a) = 0$  genau dann, wenn  $f(X)$  durch  $X - a$  teilbar ist.*
- b) *Ist  $f(a) = 0$ , so gibt es eine ganze Zahl  $d \geq 1$  und ein Polynom  $g \in k[X]$  derart, dass  $f(X) = (X - a)^d g(X)$  gilt und  $g(a) \neq 0$  ist. Sowohl  $d$  als auch  $g$  sind eindeutig durch  $f, a$  definiert. Man nennt  $d$  die Vielfachheit der Nullstelle  $a$  von  $f$ .*

*Beweis.* a) Diese Aussage wurde bereits in der LAAG1 als Lemma 7.3 bewiesen, per Induktion über  $\text{grad}(f)$ . Alternativ kann man sie bequem aus den Divisionsalgorithmus (Lemma 10.3) hergeleitet werden: denn dies besagt, dass es ein Polynom  $q \in k[X]$  und ein Skalar  $r \in k$  gibt derart, dass  $f(X) = (X - a)q(X) + r$  ist. Setzt man  $X = a$  ein, so ist  $r = f(a)$ . Also gilt: ist  $f(a) = 0$ , so folgt  $(X - a) \mid f(X)$ .

- b) Existenz: Wegen a) existieren Faktorisierungen  $f(X) = (X - a)^d g(X)$  mit  $d \geq 1$ , andererseits ist  $d \leq \text{grad}(f)$  für jede solche Faktorisierung. Wählen wir eine solche Faktorisierung mit dem größtmöglichen Wert von  $d$ . Dann  $g(a) \neq 0$ , sonst würden wir aus a) folgern, dass  $g(X)$  durch  $X - a$  teilbar sein müsste, weshalb  $f(X)$  durch  $(X - a)^{d+1}$  teilbar wäre, in Widerspruch zur Maximalität von  $d$ .

Eindeutigkeit: Angenommen es ist  $f(X) = (X - a)^d g(X) = (X - a)^e h(X)$  mit  $g(a) \neq 0 \neq h(a)$ . Ohne Einschränkung ist  $d \leq e$ . Dann  $0 = f(X) - f(X) = (X - a)^d (g(X) - (X - a)^{e-d} h(X))$ . Aus Lemma 10.2 ist bekannt:



ist das Produkt von zwei Polynome gleich Null, so muss eins der beiden Polynome Null sein. Es ist also  $0 = g(X) - (X - a)^{e-d}h(X)$ . Ist  $e = d$ , so heißt das  $g(X) - h(X) = 0$ , und wir sind fertig. Ist  $e \neq d$ , so setzen wir  $X = a$  ein und erhalten  $0 = g(a) - 0 \neq 0$ , ein Widerspruch. ■

## Algebraische und geometrische Vielfachheit

*Definition* Sei  $V$  ein  $n$ -dimensionaler  $k$ -Vektorraum. Sei  $F$  ein Endomorphismus von  $V$ . Bekanntlich ist ein Skalar  $\lambda \in k$  genau dann ein Eigenwert von  $F$ , wenn  $\lambda$  eine Nullstelle des charakteristischen Polynoms  $p_F(X)$  ist.

- a) Die Dimension des Eigenraums  $E_\lambda(F)$  nennt man die *geometrische Vielfachheit* des Eigenwerts  $\lambda$ .
- b) Die Vielfachheit von  $\lambda$  als eine Nullstelle des charakteristischen Polynoms  $p_F(X)$  nennt man die *algebraische* Vielfachheit des Eigenwerts  $\lambda$ .

**Satz 13.2** *Sei  $F$  ein Endomorphismus eines  $n$ -dimensionalen  $k$ -Vektorraums  $V$ , und sei  $\lambda$  ein Eigenwert von  $k$ . Dann gilt*

$$1 \leq \text{Geometrische Vielfachheit}(\lambda) \leq \text{Algebraische Vielfachheit}(\lambda) \leq n.$$

*Beweis.* Ist  $\lambda$  ein Eigenwert, so gibt es mindestens einen Eigenvektor mit Eigenwert  $\lambda$ , d.h. der Eigenraum enthält mindestens einen Vektor  $\neq 0$  und hat also Dimension  $> 0$ . Das charakteristische Polynom hat Grad  $n$ . Ist also  $p_F(X)$  durch  $(X - \lambda)^d$  teilbar, so ist  $d \leq n$ .

Sei  $t_1, \dots, t_r$  eine Basis des Eigenraums  $E_\lambda(F)$ . Dieses System setzen wir zu einer Basis  $T: t_1, \dots, t_n$  von  $V$  fort. Sei  $A$  die Matrix  $A = {}_T M_T(F)$ , dann ist  $p_F(X) = p_A(X)$ . Nun,  $A$  hat die Blockmatrix-Gestalt

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

mit  $B = \lambda E_r$  und  $D \in M_{n-r}(k)$ . Also gilt  $p_A(X) = p_B(X)p_D(X)$ , wegen Lemma 11.1 a). Somit ist  $p_A(X)$  durch  $p_B(X) = (X - \lambda)^r$  teilbar, weshalb die Vielfachheit der Nullstelle  $\lambda$  mindestens  $r$  beträgt. Aber  $r$  ist die geometrische Vielfachheit. ■

## Das Fitting-Lemma

Auch die algebraische Vielfachheit ist die Dimension eines Unterraums, der durch den Eigenwert  $\lambda$  definiert wird. Um dies zu erkennen, benötigen wir ein Lemma.

**Lemma 13.3 (Fitting-Lemma)** *Sei  $F$  ein Endomorphismus des endlich dimensionalen  $k$ -Vektorraums  $V$ . Dann:*

a) Es ist

$$\{0\} \subseteq \text{Kern}(F) \subseteq \text{Kern}(F^2) \subseteq \dots \subseteq \text{Kern}(F^r) \subseteq \text{Kern}(F^{r+1}) \subseteq \dots \subseteq V$$

und

$$V \supseteq \text{Bild}(F) \supseteq \text{Bild}(F^2) \supseteq \dots \supseteq \text{Bild}(F^r) \supseteq \text{Bild}(F^{r+1}) \supseteq \dots \supseteq \{0\}.$$

Beide Türme bestehen aus invarianten Unterräumen und sind nach endlicher Zeit konstant, d.h. es gibt ein  $m \geq 1$  mit  $\text{Kern}(F^{m+r}) = \text{Kern}(F^m)$  und  $\text{Bild}(F^{m+r}) = \text{Bild}(F^m)$  für alle  $r \geq 0$ .

b) Für jedes solche  $m$  gilt  $V = \text{Kern}(F^m) \oplus \text{Bild}(F^m)$ .

*Beweis.* Für alle  $v \in V$  ist  $F^r(F(v)) = F^{r+1}(v) = F(F^r(v))$ . Hieraus folgt einerseits, dass  $\text{Kern}(F^{r+1}) \supseteq \text{Kern}(F^r)$  ist, und  $\text{Kern}(F^r)$  invariant ist; und andererseits, dass  $\text{Bild}(F^{r+1}) \subseteq \text{Bild}(F^r)$  ist, und dass  $\text{Bild}(F^r)$  invariant ist. Da  $\dim \text{Kern}(F^r)$  nur endlich oft wachsen und  $\dim \text{Bild}(F^r)$  nur endlich oft fallen darf, sind beide Türme nach endlicher Zeit konstant.

Es gibt also ein  $m \geq 1$  mit  $\text{Kern}(F^{m+r}) = \text{Kern}(F^m)$  und  $\text{Bild}(F^{m+r}) = \text{Bild}(F^m)$  für alle  $r \geq 0$ . Insbesondere ist dann  $\text{Kern}(F^{2m}) = \text{Kern}(F^m)$ . Ist  $v \in \text{Kern}(F^m) \cap \text{Bild}(F^m)$ , so gibt es ein  $w \in V$  mit  $v = F^m(w)$ . Aus  $v \in \text{Kern}(F^m)$  folgt  $F^m(v) = 0$ , also  $F^{2m}(w) = 0$ . Wegen  $\text{Kern}(F^{2m}) = \text{Kern}(F^m)$  folgt  $F^m(w) = 0$ , d.h.  $v = 0$ . Also  $\text{Kern}(F^m) \cap \text{Bild}(F^m) = \{0\}$ . Nach der Dimensionsformel für  $F^m$  ist also  $V = \text{Kern}(F^m) \oplus \text{Bild}(F^m)$ . ■

*Beispiel* Sei  $A \in M_4(\mathbb{R})$  die Matrix

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

Es ist

$$A^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 2 & 2 & -1 & -1 \\ 2 & 2 & -1 & -1 \\ 1 & 1 & -1 & 0 \end{pmatrix} \quad A^3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & -1 \\ 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad A^4 = A^3.$$

Also  $\dim \text{Kern}(A^r) = 1, 2, 3$  für  $r = 1, 2, 3$ , und  $\text{Kern}(A^r) = \text{Kern}(A^3)$  für alle  $r \geq 3$ . Außerdem ist  $\dim \text{Bild}(A^r) = 3, 2, 1$  für  $r = 1, 2, 3$ , und  $\text{Bild}(A^r) = \text{Bild}(A^3)$  für alle  $r \geq 3$ . Also  $\mathbb{R}^4 = \text{Kern}(A^3) \oplus \text{Bild}(A^3)$ . Der Vektor  $(0, 1, 1, 0)$  ist eine Basis für  $\text{Bild}(A^3)$ , und  $(1, -1, 0, 0), (0, 0, 1, 0), (1, 0, 0, 1)$  ist eine Basis für  $\text{Kern}(A^3)$ .

## Haupträume

Wir deuten die algebraische Vielfachheit als die Dimension eines bestimmten invarianten Unterraums.

*Definition* Sei  $V$  ein endlich dimensionaler  $k$ -Vektorraum. Sei  $F$  ein Endomorphismus von  $V$ . Sei  $\lambda$  ein Eigenwert von  $F$ . Sei

$$U = \{v \in V \mid \text{Es gibt ein } n \geq 1 \text{ mit } (F - \lambda \text{Id})^n(v) = 0\}.$$

Nach dem Fitting-Lemma 13.3 ist  $U = \text{Kern}((F - \lambda \text{Id})^m)$  für alle  $m$  groß genug, und deshalb ein invarianter Unterraum von  $V$ .

Diesen invarianten Unterraum  $U$  von  $V$  nennt man den *Hauptraum* von  $F$  zum Eigenwert  $\lambda$ . Anstelle von Hauptraum wird  $U$  manchmal der *verallgemeinerte Eigenraum* genannt.

*Beispiel* Wie oben sei  $A \in M_4(\mathbb{R})$  die Matrix

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

Dann gilt  $A \cdot v = 0$  für  $v = (0, 1, 1, 1)$ , weshalb 0 ein Eigenwert von  $A$  ist. Wir sahen oben, dass  $\text{Kern}(A^m) = \text{Kern}(A^3)$  für alle  $m \geq 3$ . Also der Hauptraum ist  $\text{Kern}(A^3)$ , mit Basis  $(1, -1, 0, 0)$ ,  $(0, 0, 1, 0)$ ,  $(1, 0, 0, 1)$ .

**Lemma 13.4** *Die Dimension des Hauptraums zum Eigenwert  $\lambda$  stimmt mit der algebraischen Vielfachheit von  $\lambda$  überein.*

*Beweis.* Für  $m$  groß genug ist der Hauptraum  $U$  der invariante Unterraum  $U = \text{Kern}((F - \lambda \text{Id})^m) = \text{Kern}((F - \lambda \text{Id})^{m+1})$ . Nach Lemma 12.3 ist  $p_F(X) = p_{F|_U}(X)p_{\bar{F}}(X)$ , wobei  $\bar{F}$  der induzierte Endomorphismus des Quotientenraums  $V/U$  ist.

Das Polynom  $(X - \lambda)^m$  verschwindet auf  $U = \text{Kern}((F - \lambda \text{Id})^m)$  wenn man  $X = F|_U$  einsetzt. Somit ist  $F|_U$  triangulierbar nach Korollar 12.6, denn das Minimalpolynom teilt  $(X - \lambda)^m$ . Folglich ist auch das charakteristische Polynom  $p_{F|_U}(X)$  ein Produkt von linearen Faktoren  $X - \lambda_i$ . Jedes  $\lambda_i$  ist ein Eigenwert und somit eine Nullstelle des Minimalpolynoms: ist  $v \in U$  ein Eigenvektor, so ist  $m_{F|_U}(\lambda_i)v = m_{F|_U}(F)(v) = 0$ . Also  $p_{F|_U}(X) = (X - \lambda)^r$  für  $r = \dim U$ .

Somit ist die Dimension  $r$  des Hauptraums  $\leq$  die algebraische Vielfachheit. Ist die algebraische Vielfachheit größer, so ist  $p_{\bar{F}}(X)$  durch  $X - \lambda$  teilbar, also gibt es  $v \in V$  derart, dass  $v + U$  ein Eigenvektor von  $\bar{F}$  mit Eigenwert  $\lambda$  ist. Also  $(F - \lambda \text{Id})(v) \in U$ , weshalb  $(F - \lambda \text{Id})^{m+1}(v) = 0$ , weshalb  $v \in U$ . Dies kann nicht sein, denn  $v + U$  ist ein Eigenvektor und deshalb  $\neq 0$ . ■

## Die kanonische Zerlegung eines Endomorphismus

**Lemma 13.5** Sei  $F$  ein Endomorphismus des endlich dimensionalen Vektorraums  $V$ .

- a) Jede Summe von Haupträumen von  $F$  ist direkt.
- b)  $F$  ist genau dann triangulierbar, wenn  $V$  die direkte Summe aller Haupträume von  $F$  ist.

*Beweis.* a) Per Induktion über  $n = \dim(V)$ . Sei  $\lambda$  ein Eigenwert von  $F$ . Dann  $p_F(X) = (X - \lambda)^r g(X)$  für ein normiertes Polynom  $g(X)$  mit  $g(\lambda) \neq 0$ . Nach Lemma 13.4 ist  $r = \dim(U)$ , wobei  $U$  der Hauptraum zum Eigenwert  $\lambda$  ist.

Nach dem Fitting-Lemma 13.3 ist  $V = U \oplus W$ , wobei  $W$  der invariante Unterraum  $W = \text{Bild}((F - \lambda \text{Id})^m)$  ist für  $m$  groß genug. Wählt man Basen für  $U$  und für  $W$ , so erhält man eine Basis von  $V$ . Bezüglich dieser Basis hat die Matrix von  $F$  die Blockgestalt  $\begin{pmatrix} B & 0 \\ 0 & D \end{pmatrix}$ , wobei  $B$  bzw.  $D$  die Matrix von  $F|_U$  bzw.  $F|_W$  ist. Also

$$p_F(X) = p_{F|_U}(X)p_{F|_W}(X).$$

Ist  $\mu \neq \lambda$  ein weiterer Eigenwert von  $F$ , so sind  $p_{F|_W}(X)$  und  $p_F(X)$  genau so oft wie einander durch  $X - \mu$  teilbar, denn  $X - \mu$  und  $X - \lambda$  sind teilerfremd. Somit haben die Haupträume von  $F$  und von  $F|_W$  zum Eigenwert  $\mu$  die gleiche Dimension und sind somit gleich. Nach Induktionsannahme ist also die Summe der weiteren Haupträume direkt, und ein Unterraum von  $W$ . Also bleibt die Summe direkt, wenn man  $U$  dazu nimmt.

- b) Zerfällt das charakteristische Polynom als ein Produkt von linearen Faktoren, so stimmt die Summe der algebraischen Vielfachheiten der verschiedenen Eigenwerten mit  $\dim V$  überein, und  $V$  ist deshalb die direkte Summe aller Haupträume.

Umgekehrt ist das charakteristische Polynom durch  $(X - \lambda)^r$  teilbar für jeden Eigenwert  $\lambda$ , wobei  $r$  die algebraische Vielfachheit ist, also  $r =$  Dimension des Hauptraums. Ist  $V$  die direkte Summe des Hauptraums, so ist das Produkt aller  $(X - \lambda)^r$  ein normiertes Polynom, das das charakteristische Polynom teilt und den gleichen Grad hat, d.h. dieses Produkt ist das charakteristische Polynom. ■

*Bemerkung* Meistens hat ein Vektorraum mehrere Basen, und ein Unterraum hat mehrere Komplemente. Dagegen gibt es für festes  $F$  nur eine Zerlegung von  $V$  als direkte Summe von Haupträumen. Diese Zerlegung heißt die *kanonische Zerlegung*.

*Beispiel* Wie oben sei  $A \in M_4(\mathbb{R})$  die Matrix

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

Die Matrix ist triangulierbar, denn es ist  $A^4 = A^3$ , weshalb das Minimalpolynom ein Teiler von  $X^3(X - 1)$  ist und somit ein Produkt von linearen Faktoren ist.

Die Eigenwerte sind 0, 1. Sei  $H_0$  bzw.  $H_1$  der Hauptraum zum Eigenwert 0 bzw. 1. Wir sahen oben, dass  $H_0$  dreidimensional ist mit Basis  $(1, -1, 0, 0)$ ,  $(0, 0, 1, 0)$ ,  $(1, 0, 0, 1)$ . Es ist  $\dim H_0 + \dim H_1 = 4$ , also  $\dim H_1 = 1$ . Der Vektor  $(0, 1, 1, 0)$  liegt im Eigenraum  $E_1(A)$ , der ein  $H_1$  liegt, und bildet deshalb eine Basis von  $H_1$ .

Die kanonische Zerlegung von  $\mathbb{R}^4$  bezüglich der Matrix  $A$  ist also  $\mathbb{R}^4 = H_0 \oplus H_1$  mit  $H_0 = \text{Spann}((1, -1, 0, 0), (0, 0, 1, 0), (1, 0, 0, 1))$  und  $H_1 = \text{Spann}((0, 1, 1, 0))$ . Bei den Basen von  $H_0, H_1$  haben wir einiges an Wahlfreiheit; die Unterräume  $H_0, H_1$  dagegen sind eindeutig festgelegt.

Bezüglich dieser Basis  $(1, -1, 0, 0), (0, 0, 1, 0), (1, 0, 0, 1), (0, 1, 1, 0)$  von  $\mathbb{R}^4$  operiert die Matrix  $A$  als

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

## 14 Nilpotente Endomorphismen

Sei  $F$  ein Endomorphismus von  $V$ . Ist  $F$  triangulierbar, so kann man eine Basis  $B$  von  $V$  derart wählen, dass die Matrix  $A = {}_B M_B(F)$  in oberer Dreiecksgestalt ist. Im nächsten Kapitel werden wir sehen, dass man diese Basis so wählen kann, dass nur die Diagonaleinträge  $A_{ii}$  und die Einträge  $A_{i,i+1}$  direkt oberhalb der Diagonale nicht Null sind. Die Zutaten hierfür sind die kanonische Zerlegung in Haupträumen sowie eine Analyse von *nilpotenten* Endomorphismen.

*Definition* Ein Endomorphismus  $F$  eines  $n$ -dimensionalen  $k$ -Vektorraums  $V$  heißt *nilpotent*, wenn es ein  $r \geq 1$  gibt derart, dass  $F^r = 0$  ist.

**Entwurf eines Lemmas** Sei  $F$  ein nilpotenter Endomorphismus von  $V$ . Dann gibt es eine Basis  $B: b_1, \dots, b_n$  derart, dass die Matrix  $A = {}_B M_B(F)$  die folgenden Bedingungen erfüllt:

- Jeder Eintrag der Art  $A_{i,i+1}$  ist entweder 0 oder 1.
- Alle weitere Einträge sind Null.

*Überlegung* Gibt es eine solche Basis, so ist  $F(b_1) = 0$ , und  $F(b_{i+1}) = A_{i,i+1}b_i$  für alle  $1 \leq i \leq n-1$ . Hier ist eine solche Matrix:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Es ist  $F(b_6) = 0$ ,  $F(b_5) = b_4$ ,  $F(b_4) = 0$ ,  $F(b_3) = b_2$ ,  $F(b_2) = b_1$  und  $F(b_1) = 0$ . Seien  $U_1, U_2, U_3$  die folgenden Unterräume des  $k^6$ :

$$U_1 = \text{Spann}(b_6) \quad U_2 = \text{Spann}(b_5, b_4) \quad U_3 = \text{Spann}(b_3, b_2, b_1).$$

Dann  $k^6 = U_1 \oplus U_2 \oplus U_3$ , und es ist  $F(U_i) \subseteq U_i$  für  $i = 1, 2, 3$ . Ferner ist  $U_3$  der kleinste invariante Unterraum, der  $b_3$  enthält: ist  $F(U) \subseteq U$  und  $b_3 \in U$ , dann liegt auch  $b_2 = F(b_3)$  und deshalb auch  $b_1 = F(b_2)$  in  $U$ , weshalb  $U_3 = \text{Spann}(b_3, b_2, b_1) \subseteq U$ . Analog ist  $U_2$  bzw.  $U_1$  der kleinste invariante Unterraum, der  $b_5$  bzw.  $b_6$  enthält. Man nennt die invarianten Unterräume  $U_1, U_2$  und  $U_3$  deshalb *zyklisch*.

*Definition* Sei  $F: V \rightarrow V$  ein Endomorphismus und  $U \subseteq V$  ein invarianter Unterraum. Man nennt  $U$  ein *zyklischer* invarianter Unterraum, falls es ein  $u \in U$  gibt derart, dass folgendes gilt:

$$U = \text{Spann}(u, F(u), F^2(u), \dots) = \text{Spann}\{F^r(u) \mid r \geq 0\}.$$

Ein solches  $u$  nennt man ein *Erzeuger* des zyklischen Unterraums  $U$ . Diese Tatsache werden wir manchmal in der Bezeichnung  $U = Z_F(u)$  festhalten.

*Beispiel* Für  $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  bilden  $e_1$  und  $e_2 = A \cdot e_1$  eine Basis des  $\mathbb{R}^2$ . Somit ist  $\mathbb{R}^2$  zyklisch als  $A$ -invarianter Raum, und  $e_1$  ist ein Erzeuger dieses zyklischen Raums. Dagegen ist  $e_1 + e_2$  kein Erzeuger, denn  $A \cdot (e_1 + e_2) = e_1 + e_2$ .

**Lemma 14.1** Sei  $F: V \rightarrow V$  ein nilpotenter Endomorphismus, und sei  $0 \neq v \in V$  ein Vektor. Es ist  $F^m = 0$  für alle  $m$  groß genug, also gibt es ein  $r \geq 1$  derart, dass  $F^r(v) = 0$  aber  $F^{r-1}(v) \neq 0$ .

Dann sind die  $r$  Vektoren  $v, F(v), F^2(v), \dots, F^{r-1}(v)$  linear unabhängig, und sie bilden sogar eine Basis des zyklischen Unterraums  $Z_F(v)$ . Insbesondere gilt dann

$$\dim Z_F(v) = \text{das kleinste } r \geq 1 \text{ mit } F^r(v) = 0.$$

*Beweis.* Es ist  $Z_F(v) = \text{Spann}(v, F(v), \dots, F^{r-1}(v))$ , denn es ist  $Z_F(v) = \text{Spann}(v, F(v), \dots, F^{r-1}(v), F^r(v), \dots)$  per Definition, und  $F^s(v) = 0$  für alle  $s \geq r$ .

Lineare Unabhängigkeit: ist  $\sum_{i=0}^{r-1} \lambda_i F^i(v) = 0$  für Skalare  $\lambda_0, \dots, \lambda_{r-1} \in k$ , so ist  $\lambda_i = 0$  für alle  $i$ : denn sonst sei  $s$  die kleinste Zahl mit  $\lambda_s \neq 0$ , dann ist  $\lambda_i = 0$  für alle  $i < s$ , und  $\sum_{i=s}^{r-1} \lambda_i F^i(v) = 0$ . Man wendet jetzt  $F^{r-1-s}$  auf beiden Seiten dieser Gleichung an. Ist  $i > s$ , so ist  $F^{r-1-s} F^i(v) = F^{r+(i-s-1)}(v) = 0$ . Also  $\lambda_s F^{r-1}(v) = 0$  und deshalb  $\lambda_s = 0$  wegen  $F^{r-1}(v) \neq 0$ . Dies ist ein Widerspruch, denn  $\lambda_s \neq 0$ . ■

**Lemma 14.2** Sei  $F$  ein nilpotenter Endomorphismus eines endlich dimensionalen Vektorraums  $V$ . Dann lässt sich  $V$  als eine direkte Summe von zyklischen Unterräumen zerlegen:

$$V = Z_F(v_1) \oplus Z_F(v_2) \oplus \dots \oplus Z_F(v_t).$$

*Zusatz:* Sei  $m \geq 1$  die kleinste Zahl mit  $F^m = 0$ , dann ist  $\text{Kern}(F^{m-1}) \subsetneq V$ . Seien  $u_1, \dots, u_s \in V$  beliebige Vektoren derart, dass die Restklassen

$$u_1 + \text{Kern}(F^{m-1}), \dots, u_s + \text{Kern}(F^{m-1})$$

eine Basis für den Quotientenraum  $V/\text{Kern}(F^{m-1})$  bilden. Dann kann man die obige Zerlegung von  $V$  so wählen, dass  $v_i = u_i$  gilt für  $1 \leq i \leq s$ .

*Beweis.* Wir beweisen den Zusatz (einschl. Hauptaussage) per Induktion über  $n = \dim(V)$ . Den Induktionsanfang stellt der Fall  $F = 0$  dar, der insbesondere im Fall  $n = 1$  gegeben ist. Ist  $F = 0$ , so ist  $Z_F(u) = \text{Spann}(u)$  für jedes  $u \in V$ , und für jede Basis  $v_1, \dots, v_n$  von  $V$  ist  $V = Z_F(v_1) \oplus \dots \oplus Z_F(v_n)$ .

Induktionsschritt: Seien  $u_1, \dots, u_s$  wie im Zusatz, also liegt keine nichttriviale Linearkombination der  $u_i$  in  $\text{Kern}(F^{m-1})$ . Wir zeigen zunächst, dass die Summe

$$W := Z_F(u_1) + Z_F(u_2) + \dots + Z_F(u_s)$$

direkt ist. Das heißt, wir zeigen dass die Vektoren  $F^j(u_i)$  sind linear unabhängig für  $1 \leq i \leq s$  und  $0 \leq j \leq m-1$ , denn nach Wahl der  $u_i$  ist immer erst  $F^m(u_i) = 0$ . Seien also  $\lambda_{ij} \in k$  Skalare für  $1 \leq i \leq s$  und  $0 \leq j \leq m-1$  derart, dass  $\sum_{i=1}^s \sum_{j=0}^{m-1} \lambda_{ij} F^j(u_i) = 0$ . Sei  $j_0$  der kleinste Wert von  $j$  derart, dass  $\lambda_{ij_0} \neq 0$  ist für mindestens ein  $i$ . Also  $\sum_{i=1}^s \sum_{j=j_0}^{m-1} \lambda_{ij} F^j(u_i)$ . Wendet man  $F^{m-1-j_0}$  auf beiden Seiten an, so erhalten wir  $\sum_{i=1}^s \lambda_{ij_0} F^{m-1}(u_i) = 0$  und deshalb  $F^{m-1}(\sum_{i=1}^s \lambda_{ij_0} u_i) = 0$ . Nach der Annahme im Zusatz ist also  $\lambda_{ij_0} = 0$  für alle  $i$ , ein Widerspruch zur Wahl von  $j_0$ .

Nun betrachten wir den invarianten Unterraum  $U = \text{Kern}(F^{m-1})$  des  $V$ , und den eingeschränkten Endomorphismus  $G = F|_U$  des  $U$  mit  $G^{m-1} = 0$ . Die Vektoren  $F(u_1), \dots, F(u_s)$  sind Elemente von  $U$ . Sie sind auch linear unabhängig modulo  $\text{Kern}(G^{m-2})$ : denn wäre  $\sum_{i=1}^s \mu_i F(u_i) = w$  für ein  $w \in V$  mit  $F^{m-2}(w) = 0$ , dann wenden wir  $F^{m-2}$  an und erhalten  $\sum_{i=1}^s \mu_i F^{m-1}(u_i) = 0$ , woraus folgt wie oben, dass  $\mu_i = 0$  für alle  $i$ .

Nach Induktionsannahme gilt der Zusatz einschl. Hauptaussage für  $U, G$ . Es gibt also ein  $t \geq s$  und Vektoren  $v_{s+1}, \dots, v_t \in U$  derart, dass

$$U = \bigoplus_{i=1}^s Z_G(F(u_i)) \oplus \bigoplus_{i=s+1}^t Z_G(v_i)$$

gilt, das heißt

$$U = \bigoplus_{i=1}^s Z_F(F(u_i)) \oplus \bigoplus_{i=s+1}^t Z_F(v_i). \quad (*)$$

Es ist  $Z_F(u_i) = \text{Spann}(u_i) \oplus Z_F(F(u_i))$ , und deshalb

$$\text{Spann}(u_1, \dots, u_s) + U = \sum_{i=1}^s Z_F(u_i) \oplus \sum_{i=s+1}^t Z_F(v_i). \quad (**)$$

Da  $u_1 + U, \dots, u_s + U$  eine Basis für  $V/U$  ist, ist  $\text{Spann}(u_1, \dots, u_s) + U = V$  und  $\dim U = \dim V - s$ . Da die Summe in (\*) direkt ist, ist  $\dim U$  die Summe der Dimensionen der Summanden auf der rechten Seite von (\*). Die Summe der Dimensionen der Summanden in (\*\*) ist um  $s$  größer, beträgt also  $s + \dim U = \dim V$ , d.h. die Dimension der linken Seite. Also ist auch die Summe in (\*\*) direkt. ■



Beispiel Sei  $A \in M_{10}(\mathbb{R})$  die Matrix

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

Es ist dann

$$A^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 & -1 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad A^3 = 0.$$

Wahrscheinlich wird die gesuchte Zerlegung einige dreidimensionale Summanden  $Z_F(u)$  enthalten. Die Vektoren  $e_1 + e_3, e_3 + e_4$  sind linear unabhängig, und für diesen beiden Werten von  $u$  ist  $A^2(u) \neq 0$ , also  $\dim Z_F(u) = 3$ . Aber die Summe  $Z_F(e_1 + e_3) + Z_F(e_3 + e_4)$  ist nicht direkt, denn beide Summanden enthalten  $A^2(e_1 + e_3) = e_3 + e_4 + e_9 = A^2(e_3 + e_4)$ . Wir brauchen also eine Strategie, um Erzeuger zu wählen. Der Zusatz zu Lemma 14.2 und seine Beweismethode liefert eine.

Mit  $m = 3$  also ist  $A^m = 0$  und  $A^{m-1} \neq 0$ . Der Kern von  $A^2$  hat Basis  $e_1, e_2, e_4, e_7, e_5 + e_3, e_6 - e_3, e_9 + e_3, e_{10} + e_8$ . Somit bilden  $e_3, e_8$  eine Basis eines Komplements, d.h.  $e_3 + \text{Kern}(A^2), e_8 + \text{Kern}(A^2)$  ist eine Basis von  $\mathbb{R}^{10} / \text{Kern}(A^2)$ . Somit werden  $Z_A(e_3), Z_A(e_8)$  zwei drei-dimensionale Summanden von  $\mathbb{R}^{10}$  sein. Sei  $b_3 = e_3$  und  $b_6 = e_8$ . Sei  $b_2 = A \cdot b_3, b_5 = A \cdot b_6, b_1 = A \cdot b_2$  und  $b_4 = A \cdot b_5$ . Also  $b_2 = e_3 + e_{10}, b_1 = e_3 + e_4 + e_9, b_5 = e_4, b_4 = e_1 - e_2 + e_7$ .

Eine Basis für  $\text{Kern}(A)$  ist  $e_2 + e_1, e_5 + e_3, e_6 - e_3 + e_1, e_7 + 2e_1, e_9 + e_4 + e_3$ . Ein Komplement von  $\text{Kern}(A)$  in  $\text{Kern}(A^2)$  hat Basis  $e_1, e_4, e_{10} + e_8$ . Zum Glück enthält diese Basis bereits  $b_2 = A \cdot e_3$  und  $b_5 = A \cdot e_8$ . Übrig bleibt  $e_1$ . Dies wird also einen zweidimensionalen Summanden  $Z_A(e_1)$  beitragen. Wir setzen  $b_8 = e_1$  und  $b_7 = A \cdot b_8 = -e_1 - e_5 - e_6$ .

Zum Schluss suchen wir eindimensionale Summanden. Hierfür müssen wir  $b_1, b_4, b_7$  d.h.  $e_9 + e_4 + e_3, e_7 - e_2 + e_1, -e_6 - e_5 - e_1$  zu einer Basis von  $\text{Kern}(A)$  fortsetzen. Eine solche Fortsetzung ist  $e_2 + e_1, e_5 + e_3$ . Wir setzen also  $b_9 = e_1 + e_2, b_{10} = e_3 + e_5$ . Dann ist  $b_1, \dots, b_{10}$  eine Basis von  $\mathbb{R}^{10}$ , und es ist

$$b_3 \xrightarrow{A} b_2 \xrightarrow{A} b_1 \xrightarrow{A} 0 \quad b_6 \xrightarrow{A} b_5 \xrightarrow{A} b_4 \xrightarrow{A} 0 \quad b_8 \xrightarrow{A} b_7 \xrightarrow{A} 0 \quad b_9 \xrightarrow{A} 0 \quad b_{10} \xrightarrow{A} 0.$$

Also

$$\mathbb{R}^{10} = Z_F(b_3) \oplus Z_F(b_6) \oplus Z_F(b_8) \oplus Z_F(b_9) \oplus Z_F(b_{10}),$$

und  $A$  operiert auf dieser Basis als die Matrix

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

## 15 Die Jordansche Normalform

*Definition* Sei  $\lambda \in k$  ein Skalar, und  $n \geq 1$ . Sei  $A \in M_n(k)$  die Matrix mit

- $A_{ii} = \lambda$  für jedes  $1 \leq i \leq n$ ;
- $A_{i,i+1} = 1$  für jedes  $1 \leq i \leq n - 1$ ;
- $A_{ij} = 0$  sonst.

Diese Matrix nennt man  $J_n(\lambda)$ , das  $(n \times n)$ -Jordan-Kästchen zum Eigenwert  $\lambda$ . Die Zahl  $n$  werden wir die *Größe* des Jordan-Kästchens nennen.

Hier sind einige Beispiele:

$$J_3(-1) = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix} \quad J_4(1+i) = \begin{pmatrix} 1+i & 1 & 0 & 0 \\ 0 & 1+i & 1 & 0 \\ 0 & 0 & 1+i & 1 \\ 0 & 0 & 0 & 1+i \end{pmatrix}$$

$$J_2(0) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad J_1(3) = (3)$$

*Definition* Eine Matrix  $A \in M_n(k)$  heißt in *Jordansche Normalform*, wenn es Skalare  $\lambda_1, \dots, \lambda_s$  und ganze Zahlen  $m_1, \dots, m_s \geq 1$  gibt, derart,  $A$  die folgende Blockgestalt hat, wobei leere Einträge Null sind:

$$A = \begin{pmatrix} J_{m_1}(\lambda_1) & & & \\ & J_{m_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{m_s}(\lambda_s) \end{pmatrix}.$$

Die Skalare  $\lambda_i$  müssen nicht paarweise verschieden sein. Meistens gruppiert man Jordan-Kästchen mit dem gleichen Eigenwert zusammen, d.h. ist  $\lambda_j = \lambda_i$  für ein  $j > i$ , so ist  $\lambda_\ell = \lambda_i$  für alle  $i \leq \ell \leq j$ .

*Beispiele* Hier sind einige Jordansche Normalformen:

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 7 \end{pmatrix} \quad s = 2, \lambda_1 = 0, \lambda_2 = 7, m_1 = 3, m_2 = 1;$$

$$\begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

**Satz 15.1 (Die Jordansche Normalform)** Sei  $F: V \rightarrow V$  ein triangulierbarer Endomorphismus eines endlich dimensionalen  $k$ -Vektorraums  $V$ . Dann

- a) Es gibt eine Basis  $B$  von  $V$  derart, dass die Matrix  $A = {}_B M_B(F)$  eine Jordansche Normalform ist.
- b) Diese Jordansche Normalform von  $F$  ist im wesentlichen eindeutig: zwei solche Normalformen unterscheiden sich nur in der Reihenfolge der Jordan-Kästchen: pro Eigenwert bleibt die Anzahl der Kästchen einer gegebenen Größe gleich.

Insbesondere gilt: ist  $k = \mathbb{C}$ , so hat jeder Endomorphismus eine Jordansche Normalform.

*Beweis der Existenz.* Ist  $\lambda$  ein Eigenwert von  $F$ , so sei  $W$  der Hauptraum von  $F$  zum Eigenwert  $\lambda$ . Der Hauptraum ist ein invarianter Unterraum, also sei  $G: W \rightarrow W$  der Endomorphismus  $G = F|_W - \lambda \text{Id}_W$ . Nach dem Fitting-Lemma gibt es einen  $m \geq 1$  derart, dass  $W = \text{Kern}(F - \lambda \text{Id})^m$  ist. Somit ist  $G$  nilpotent, denn  $G^m = 0$ . Deshalb gibt es nach Lemma 14.2 eine Zerlegung

$$W = Z_G(u_1) \oplus \cdots \oplus Z_G(u_t).$$

Betrachten wir jetzt einen zyklischen Summanden  $Z_G(u)$ . Sei  $d = \dim Z_G(u)$ , also ist  $G^d(u) = 0$ , und  $u, G(u), \dots, G^{d-1}(u)$  ist eine Basis von  $Z_G(u)$ . Nun, für jedes  $w \in W$  ist  $F(w) = (F - \lambda \text{Id})(w) + \lambda \cdot w = G(w) + \lambda \cdot w$ . Somit operiert  $F$  auf die Basis  $b_1 = G^{d-1}(u), b_2 = G^{d-2}(u), \dots, b_d = u$  als die Matrix  $J_d(\lambda)$ . Wählt man diese Basis für jeden zyklischen Summanden von  $W$ , so ist die Matrix von  $F|_W$  in Jordansche Normalform, wobei alle Jordan-Kästchen den Eigenwert  $\lambda$  haben.

Da  $F$  triangulierbar ist, besagt die kanonische Zerlegung (Lemma 13.5), dass  $V$  eine direkte Summe von Haupträumen ist. Wählt man für jeden Hauptraum  $W$  eine Basis, die  $F|_W$  auf Jordansche Normalform bringt, so erhält man eine Basis von  $V$ , die  $F$  auf Jordansche Normalform bringt.

Zu  $k = \mathbb{C}$ : wir haben in diesem Fall bereits gesehen, dass jeder Endomorphismus triangulierbar ist. ■

*Beispiel* Oben sahen wir, dass die Matrix

$$A = \begin{pmatrix} 1 & 1 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 1 & 0 & 1 & -1 \\ 0 & -1 & 1 & 0 \end{pmatrix} \in M_4(\mathbb{R})$$

die Eigenwerte 0, 1 hat. Eine Basis des Hauptraums zum Eigenwert 0 ist  $b_1 = (1, -1, 0, 0)$ ,  $b_2 = (0, 0, 1, 0)$ ,  $b_3 = (1, 0, 0, 1)$ . Der Vektor  $b_4 = (0, 1, 1, 0)$  ist eine Basis des Hauptraums zum Eigenwert 1.

Es ist  $A \cdot b_3 = b_1$ ,  $A \cdot b_1 = (0, 1, 1, 1) = b_3 + b_2 - b_1$ , und  $A \cdot (0, 1, 1, 1) = 0$ . Wechseln wir also zur Basis  $c_1 = (0, 1, 1, 1)$ ,  $c_2 = (1, -1, 0, 0) = b_1$ ,  $c_3 = (1, 0, 0, 1) = b_3$ ,  $c_4 = (0, 1, 1, 0) = b_4$ . Auf dieser Basis operiert  $A$  als die Matrix

$$J = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} J_3(0) & 0 \\ 0 & J_1(1) \end{pmatrix}.$$

Somit ist  $J$  eine Jordansche Normalform von  $A$ .

**Lemma 15.2** Sei  $A \in M_n(k)$  eine Jordansche Normalform eines triangulierbaren Endomorphismus  $F: V \rightarrow V$ . Sei  $\lambda$  ein Eigenwert von  $\lambda$ . Über die Jordan-Kästchen in  $A$  zum Eigenwert  $\lambda$  gelten folgende Aussagen:

- Die Anzahl der Jordan-Kästchen zum Eigenwert  $\lambda$  stimmt mit der geometrischen Vielfachheit von  $\lambda$  überein.
- Die Summe der Größen der Jordan-Kästchen zum Eigenwert  $\lambda$  stimmt mit der algebraischen Vielfachheit von  $\lambda$  überein.
- Die Größe des größten Jordan-Kästchens zum Eigenwert  $\lambda$  stimmt mit der Vielfachheit von  $\lambda$  als Nullstelle des Minimalpolynoms  $m_F(X)$  überein. Diese Zahl ist zugleich die kleinste Zahl  $d$  derart, dass  $\text{Kern}((F - \lambda \text{Id})^d)$  der Hauptraum zum Eigenwert  $\lambda$  ist.

*Beweis.* Die Matrix  $J = J_n(0)$  bildet  $e_r$  auf  $e_{r-1}$  und  $e_1$  auf 0 ab. Somit ist sie nilpotent, es ist  $J^{n-1} \neq 0$  und  $J^n = 0$ . Bezüglich dieser nilpotenten Matrix ist  $k^n$  zyklisch mit Erzeuger  $e_n$ . Der Nullraum von  $J$  ist eindimensional mit Basis  $e_1$ .

Die Jordansche Normalform von  $F$  entspricht eine Zerlegung von  $V$  als eine direkte Summe  $V = \bigoplus_{i=1}^s V_i$  von invarianten Unterräumen, wobei  $F$  auf  $V_i$  als die Matrix  $J_{m_i}(\lambda_i)$  operiert. Das Minimalpolynom der Einschränkung  $F|_{V_i}$  ist  $(X - \lambda_i)^{m_i}$ .

- Der Eigenraum von  $F$  ist die direkte Summe  $E_\lambda(F) = \bigoplus_{i=1}^s E_\lambda(F|_{V_i})$ , denn jeder Summand ist invariant. Der Eigenraum  $E_\lambda(F|_{V_i})$  ist eindimensional falls  $\lambda_i = \lambda$ , und nulldimensional sonst.
- Da die Jordansche Normalform obere Dreiecksgestalt hat, lässt sich das charakteristische Polynom von den Diagonaleinträgen ablesen. Die Anzahl der Diagonaleinträge, die  $\lambda$  betragen, ist die Summe der Größen der Jordan-Kästchen zu  $\lambda$ .
- Für eine Matrix in Blockdiagonalgestalt  $A = \begin{pmatrix} B & 0 \\ 0 & D \end{pmatrix}$  gilt  $f(A) = \begin{pmatrix} f(B) & 0 \\ 0 & f(D) \end{pmatrix}$  für jedes Polynom  $f$ . Das Minimalpolynom einer Matrix in Jordanscher Normalform ist also das kleinste gemeinsame Vielfache der Minimalpolynome

der Jordan-Kästchen, d.h. das kgV der  $(X - \lambda_i)^{m_i}$ . Hieraus folgt der erste Teil von c).

Zum zweiten Teil: Sei  $V_\lambda$  die direkte Summe der  $V_i$ , für die  $\lambda_i = \lambda$  gilt. Wegen des Minimalpolynoms von  $V_i$  liegt  $V_\lambda$  im Hauptraum zum Eigenwert  $\lambda$ . Wegen b) stimmt  $V_\lambda$  mit dem Hauptraum überein, aus Dimensionsgründen. Da  $d$  die Dimension des größten Kästchen in  $V_\lambda$  ist, ist  $(F - \lambda \text{Id})^d$  die erste Potenz, die auf  $V_\lambda$  verschwindet. ■

*Beispiel* Man kann die Jordansche Normalform auch schneller bestimmen, sofern man nicht eine Basis benötigt, die den Endomorphismus auf Jordansche Normalform bringt. Angenommen, wir haben berechnet, dass  $p_A(X) = X^3(X-1)$  ist für unsere Matrix  $A \in M_4(\mathbb{R})$ , und dass beide Eigenwerte die geometrische Vielfachheit 1 haben. Nach dem Lemma folgt sofort, dass es pro Eigenwert nur ein Jordan-Kästchen gibt, und dass dessen Größe mit der algebraischen Vielfachheit des Eigenwerts übereinstimmt. Somit ist die Jordansche Normalform  $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ .

*Beweis der Eindeutigkeit in Satz 15.1.* Induktion über  $n = \dim V$ . Der Induktionsanfang ist der Fall  $F$  diagonalisierbar, der den Fall  $n = 1$  einbeschließt. In diesem Fall stimmen geometrische und algebraische Vielfachheit überein, für jeden Eigenwert. Nach Lemma 15.2 müssen alle Kästchen die Größe eins haben, und die Anzahl der Kästchen pro Eigenwert ist die Vielfachheit des Eigenwerts. Somit ist die Jordansche Normalform eindeutig, bis auf die Reihenfolge der Kästchen.

Induktionsschritt: Ist  $F$  nicht diagonalisierbar, so sei  $U \subsetneq V$  die direkte Summe aller Eigenräume. Ist  $A = {}_B M_B(F)$  eine Jordansche Normalform, so enthält  $B$  eine Basis des invarianten Unterraums  $U$ . Sei  $\bar{B} = \{b + U \mid b \in B, b \notin U\}$  die entsprechende Basis von  $V/U$ , und sei  $\bar{F}$  der induzierte Endomorphismus  $\bar{F}(v + U) = F(v) + U$  des  $V/U$ . Dann ist die Matrix  $\bar{A} = {}_{\bar{B}} M_{\bar{B}}(\bar{F})$  in Jordanscher Normalform, und zwar besteht  $\bar{A}$  aus einem Kästchen  $J_{d-1}(\lambda)$  pro Kästchen  $J_d(\lambda)$  von  $A$  mit  $d \geq 2$ . Eine weitere Jordansche Normalform  $A'$  von  $F$  muss laut Induktionsannahme zur gleichen Jordanschen Normalform von  $\bar{F}$  führen. Also können  $A$  und  $A'$  sich höchstens in der Anzahl der  $(1 \times 1)$ -Kästchen unterscheiden. Aber deren Anzahl lässt sich wegen Lemma 15.2 aus der algebraischen Vielfachheit des Eigenwerts und der Anzahl und Größen der größeren Kästchen ermitteln. Somit sind  $A$  und  $A'$  gleich, abgesehen von der Reihenfolge der Kästchen. ■

*Bemerkung* Haben also die Matrizen  $A, B$  unterschiedliche Jordansche Normalformen, so sind sie nicht ähnlich.

## 16 Der Dualraum

Zur Erinnerung:  $L(V, W)$  bezeichnet den Vektorraum aller linearen Abbildungen von  $V$  nach  $W$ .

*Definition* Sei  $V$  ein  $k$ -Vektorraum. Den Vektorraum  $L(V, k)$  aller linearen Abbildungen  $\phi: V \rightarrow k$  nennt man den *Dualraum*  $V^*$  von  $V$ . Elemente  $\phi \in V^*$  heißen *Linearformen* auf  $V$ .

*Beispiel* Eine Linearform auf  $\mathbb{R}^3$  ist die Abbildung  $\phi: \mathbb{R}^3 \rightarrow \mathbb{R}$  gegeben durch  $\phi(x, y, z) = 2x - y + 3z$ . Eine weitere ist  $\psi$ , gegeben durch  $\psi(x, y, z) = 4y - z$ . Es ist  $\phi(1, 1, 1) = 4 \neq 3 = \psi(1, 1, 1)$ , weshalb  $\psi \neq \phi$ .

### Die Dimension des Dualraums

Um die Dimension des Dualraums zu bestimmen, benötigen wir das folgende Lemma, das wir eigentlich viel früher hätten beweisen können.

**Lemma 16.1** *Seien  $m, n$  ganze Zahlen  $\geq 1$ . Seien  $V, W$  zwei endlich dimensionalen  $k$ -Vektorräume. Dann*

$$a) \dim M(m \times n, k) = mn \quad b) \dim L(V, W) = \dim(V) \cdot \dim(W).$$

*Beweis.* a) (Vgl. Beweis von Lemma 10.1, ganz am Anfang des Semesters)  
Für  $1 \leq r \leq m$  und  $1 \leq s \leq n$  sei  $E(r, s)$  die  $(m \times n)$ -Matrix gegeben durch  $E(r, s)_{ij} = \delta_{ir} \delta_{js}$ , d.h.  $E(r, s)$  hat eine 1 an der  $(r, s)$ -Stelle, und alle anderen Einträge sind Null. Die  $mn$  Matrizen dieser Art sind linear unabhängig (Vergleich der  $(r, s)$ -Einträge), und wegen  $A = \sum_{r=1}^m \sum_{s=1}^n A_{rs} E(r, s)$  spannen Sie  $M(m \times n, k)$  auch auf. Somit bilden sie eine Basis.

b) Sei  $n = \dim(V)$ ,  $m = \dim(W)$ . Sei  $B$  bzw.  $C$  eine Basis von  $V$  bzw.  $W$ . Im Wintersemester zeigten wir (in Lemma 4.6), dass die Abbildung  $L(V, W) \rightarrow M(m \times n, k)$ ,  $f \mapsto {}_B M_C(f)$  ein Isomorphismus ist. ■

**Korollar 16.2** *Sei  $V$  ein endlich dimensionaler  $k$ -Vektorraum. Dann  $\dim V^* = \dim V$ . Somit sind die Vektorräume  $V$  und  $V^*$  isomorph.*

*Beweis.* Es ist  $\dim V^* = \dim V \cdot \dim k = \dim V$ . Im Wintersemester zeigten wir, dass zwei endlich dimensionale Vektorräume isomorph sind, wenn sie die gleiche Dimension haben (Korollar 4.4). ■

*Bemerkung* Der Beweis von Korollar 4.4 geht so: man wählt Basen von  $V$  und von  $V^*$ . Diese sind gleich groß, also kann man die eine bijektiv auf die andere abbilden. Durch lineare Fortsetzung erhält man so einen Isomorphismus.

Im allgemeinen Fall (d.h. für unendlich dimensionale Räume) hat  $V^*$  größere Dimension als  $V$ .

## Das Doppeldual

Ist  $V$  ein endlich dimensionaler  $k$ -Vektorraum, so haben  $V, V^*$  die gleiche Dimension. Nach dem gleichen Argument haben auch  $V^*$  und das Doppeldual  $V^{**} := (V^*)^*$  die gleiche Dimension. Also sind auch  $V, V^{**}$  isomorph. Bei  $V, V^*$  musste man Basen wählen, um einen Isomorphismus zu erhalten. Dagegen sind  $V, V^{**}$  *natürlich* isomorph, wie wir jetzt sehen werden. Dies bedeutet, dass man keine Wahlen treffen muss, um den Isomorphismus angeben zu können.

*Alternative Schreibweise* Für  $v \in V$  und  $\phi \in V^*$  schreiben wir  $\langle \phi, v \rangle$  für das Skalar  $\phi(v) \in k$ . Somit erhalten wir eine Abbildung  $\langle, \rangle: V^* \times V \rightarrow k$ .

**Lemma 16.3** Sei  $V$  ein endlich dimensionaler  $k$ -Vektorraum.

- a) Für festes  $\phi \in V^*$  ist die Abbildung  $V \rightarrow k, v \mapsto \langle \phi, v \rangle$  linear.
- b) Für festes  $v \in V$  ist die Abbildung  $V^* \rightarrow k, \phi \mapsto \langle \phi, v \rangle$  linear.
- c) Ist  $\phi \in V^*$  derart, dass  $\langle \phi, v \rangle = 0$  gilt für jedes  $v \in V$ , dann ist  $\phi = 0$ .
- d) Ist  $v \in V$  derart, dass  $\langle \phi, v \rangle = 0$  gilt für jedes  $\phi \in V^*$ , dann ist  $v = 0$ .

*Beweis.* a) Laut Voraussetzung ist die Abbildung  $\phi$  linear.

b) So sind Addition und Skalarmultiplikation auf  $V^*$  definiert.

c) So ist die Nullabbildung definiert.

d) Ist  $v \neq 0$ , so können wir  $v$  zu einer Basis  $b_1 = v, b_2, \dots, b_n$  des  $V$  fortsetzen. Nach dem Satz von der linearen Fortsetzung gibt es eine lineare Abbildung  $\phi: V \rightarrow k$  mit  $\phi(b_i) = 1$  für jedes  $i$ . Also  $\phi(v) = 1 \neq 0$ . ■

*Bemerkung* Wegen a), b) sagt man, dass die Paarung  $\langle, \rangle$  *bilinear* ist. Aufgrund von c), d) heißt diese bilineare Paarung *nicht ausgeartet*.

*Bezeichnung* Sei  $v$  ein Element des  $k$ -Vektorraums  $V$ . Sei  $e_v: V^* \rightarrow k$  die Abbildung  $e_v(\phi) = \langle \phi, v \rangle$ . Nach Teil b) des Lemmas ist  $e_v$  eine Linearform auf  $V^*$ , d.h.  $e_v \in V^{**}$ . Die Abbildung  $e: V \rightarrow V^{**}, v \mapsto e_v$  nennt man die *Auswerteabbildung*.

**Lemma 16.4** Für einen endlich dimensionalen  $k$ -Vektorraum  $V$  ist die Auswerteabbildung  $e: V \rightarrow V^{**}$  ein Isomorphismus.

*Beweis.* Wegen  $\dim V^{**} = \dim V$  reicht es zu zeigen, dass  $e$  linear ist, und dass  $\text{Kern}(e) = 0$  ist. Teil d) aus Lemma 16.3 besagt, dass  $\text{Kern}(e) = 0$  ist. Linearität: es ist

$$\begin{aligned} e_{\lambda v + \mu w}(\phi) &= \langle \phi, \lambda v + \mu w \rangle = \lambda \langle \phi, v \rangle + \mu \langle \phi, w \rangle \\ &= \lambda e_v(\phi) + \mu e_w(\phi) = (\lambda e_v + \mu e_w)(\phi). \end{aligned} \quad \blacksquare$$



## Die Dualbasis

*Definition* Sei  $B : b_1, \dots, b_n$  eine Basis des  $k$ -Vektorraums  $V$ . Sei  $1 \leq i \leq n$ . Nach der Satz von der linearen Fortsetzung gibt es genau eine Linearform  $b_i^* : V \rightarrow k$  derart, dass

$$b_i^*(b_j) = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & \text{sonst} \end{cases}$$

gilt für jedes  $j$ . Man nennt  $b_1^*, \dots, b_n^*$  die *Dualbasis* von  $V^*$ .

*Bemerkung* Für ein beliebiges Element  $v = \sum_{j=1}^n \lambda_j b_j$  des  $V$  hat man also  $b_i^*(v) = \lambda_i$ .

In der Alternativ-Bezeichnung hat man  $\langle b_i^*, b_j \rangle = \delta_{ij}$ .

*Beispiel* Oben betrachteten wir die Linearformen  $\phi(x, y, z) = 2x - y + 3z$  und  $\psi(x, y, z) = 4y - z$  auf  $\mathbb{R}^3$ . Es ist  $\phi = 2e_1^* - e_2^* + 3e_3^*$  sowie  $\psi = 4e_2^* - e_3^*$ .

**Lemma 16.5** *Die Dualbasis ist tatsächlich eine Basis des Dualraums.*

*Beweis.* Linear unabhängig: Ist  $\sum_{i=1}^n \lambda_i b_i^* = 0$ , so ist  $\sum_{i=1}^n \lambda_i b_i^*(b_j) = 0$  für jedes  $j$ , und deshalb  $\lambda_j = 0$  für alle  $j$ .

Erzeugendensystem: Ist  $\phi \in V^*$ , so ist  $\phi = \sum_{i=1}^n \phi(b_i) b_i^*$  nach dem Satz von der linearen Fortsetzung, denn beide Linearformen nehmen in jedem  $b_j$  den Wert  $\phi(b_j)$  an. ■

## Der Annulator

*Definition* Sei  $V$  ein  $k$ -Vektorraum und  $T \subseteq V$  eine Teilmenge. Den *Annulator*  $T^\circ$  von  $T$  definiert man als

$$T^\circ = \{ \phi \in V^* \mid \langle \phi, v \rangle = 0 \text{ für jedes } v \in T \},$$

eine Teilmenge des Dualraums  $V^*$ .

**Lemma 16.6** *Sei  $T$  eine Teilmenge des endlich dimensionalen Vektorraums  $V$ .*

- $T^\circ$  ist ein Unterraum von  $V^*$ .
- Ist  $T_1 \subseteq T_2$ , dann  $T_2^\circ \subseteq T_1^\circ$ .
- Es ist  $T^\circ = W^\circ$  für  $W = \text{Spann}(T)$ , der durch  $T$  aufgespannte Unterraum.
- Für einen Unterraum  $W$  ist  $\dim W + \dim W^\circ = \dim V$ .
- Für Unterräume  $U, W$  von  $V$  gelten

$$(U + W)^\circ = U^\circ \cap W^\circ \qquad (U \cap W)^\circ = U^\circ + W^\circ.$$

*Beweis.* a) Folgt aus der Bilinearität von  $\langle, \rangle$ .

b) Ist  $\langle \phi, v \rangle = 0$  für alle  $v \in T_2$ , dann insbesondere auch für alle  $v \in T_1$ .

c) Zu zeigen ist  $T^\circ \subseteq W^\circ$ . Sei  $\phi \in T^\circ$  und  $v \in W$ . Wegen  $W = \text{Spann}(T)$  gibt es  $t_1, \dots, t_r \in T$  und  $\lambda_1, \dots, \lambda_r \in k$  mit  $v = \sum_{i=1}^r \lambda_i t_i$ . Also

$$\langle \phi, v \rangle = \left\langle \phi, \sum_{i=1}^r \lambda_i t_i \right\rangle = \sum_{i=1}^r \lambda_i \langle \phi, t_i \rangle = \sum_{i=1}^r \lambda_i \cdot 0 = 0.$$

d) Man wähle eine Basis  $b_1, \dots, b_r$  von  $V$  und setze sie zu einer Basis  $b_1, \dots, b_n$  von  $V$  fort. Sei  $\beta_1, \dots, \beta_n$  die Dualbasis von  $V^*$ . Wir zeigen:  $\beta_{r+1}, \dots, \beta_n$  ist eine Basis von  $W^\circ$ . Es ist  $\beta_i \in W^\circ$  für  $i \geq r+1$ , denn  $\beta_i(b_j) = 0$  für alle solche  $i$  und für alle  $j \leq r$ . Ist umgekehrt  $\phi = \sum_{i=1}^n \lambda_i \beta_i \in W^\circ$ , so ist  $\phi(b_j) = 0$  für alle  $j \leq r$ : aber  $\phi(b_j) = \lambda_j$ . Also  $\phi = \sum_{i=r+1}^n \lambda_i \beta_i$ .

e) Es ist  $U + W = \text{Spann}(U \cup W)$ . Aus der Definition von  $T^\circ$  folgt, dass  $(U \cup W)^\circ = U^\circ \cap W^\circ$  ist. Wegen c) folgt also  $(U + W)^\circ = U^\circ \cap W^\circ$ .

Aus der Definition folgt  $U^\circ + W^\circ \subseteq (U \cap W)^\circ$ . Es reicht also zu zeigen, dass beide Unterräume die gleiche Dimension haben. Es ist

$$\begin{aligned} \dim(U^\circ + W^\circ) &= \dim U^\circ + \dim W^\circ - \dim U^\circ \cap W^\circ \\ &= \dim U^\circ + \dim W^\circ - \dim(U + W)^\circ. \end{aligned}$$

Wendet man jetzt Teil d) an, so erhält man

$$\begin{aligned} \dim(U^\circ + W^\circ) &= \dim V - \dim U - \dim W + \dim(U + W) \\ &= \dim V - \dim(U \cap W) = \dim(U \cap W)^\circ. \quad \blacksquare \end{aligned}$$

*Beispiel* Sei  $v \in \mathbb{R}^3$  der Vektor  $v = (1, 1, 2)$ . Dann bilden  $e_1^* - e_2^*, 2e_1^* - e_3^*$  eine Basis des Annulators  $v^\circ$ . Alles, was  $v$  annulliert, annulliert auch den ganzen eindimensionalen Unterraum  $U = \text{Spann}(v) = \{(x, y, z) \mid x = y, z = 2x\}$ .

**Lemma 16.7** Sei  $W$  ein Unterraum des endlich dimensionalen Vektorraums  $V$ . Es gibt natürliche Isomorphismen

a) von  $(V/W)^*$  nach  $W^\circ$ ; und

b) von  $V^*/W^\circ$  nach  $W^*$ .

*Beweis.* Es ist  $\dim(V/W)^* = \dim W^\circ$  und  $\dim(V^*/W^\circ) = \dim W^*$ . In beiden Fällen reicht es also, eine injektive lineare Abbildung zu konstruieren.

- a) Für  $\phi \in (V/W)^*$  sei  $\bar{\phi}: V \rightarrow k$  die Abbildung  $\bar{\phi}(v) = \phi(v + W)$ : die Verknüpfung von  $\phi$  mit der kanonischen Projektion  $V \rightarrow V/W$ , und somit linear. Außerdem ist  $\bar{\phi}(w) = \phi(w + W) = \phi(0 + W) = 0$  für jedes  $w \in W$ . Wir haben also eine Abbildung  $(V/W)^* \rightarrow W^\circ$ ,  $\phi \mapsto \bar{\phi}$ . Diese Abbildung ist linear:

$$\overline{\lambda\phi + \mu\psi}(v) = \lambda\phi(v + W) + \mu\psi(v + W) = (\lambda\bar{\phi} + \mu\bar{\psi})(v).$$

Ist  $\phi \neq 0$ , so gibt es ein  $v + W$  mit  $\phi(v + W) \neq 0$ , also  $\bar{\phi}(v) \neq 0$ , also  $\bar{\phi} \neq 0$ .

- b) Für  $\psi \in V^*$  setzen wir  $F(\psi + W^\circ) = \psi|_W$ , die Einschränkung von  $\psi$  auf  $W$ . Ist  $\psi + W^\circ = \phi + W^\circ$ , so ist  $\phi = \psi + \chi$  für ein  $\chi \in W^\circ$ . Wegen  $\chi \in W^\circ$  gilt  $\chi|_W = 0$  und deshalb  $\phi|_W = \psi|_W$ . Somit ist die Abbildung  $F: V^*/W^\circ \rightarrow W^*$ ,  $\psi + W^\circ \mapsto \psi|_W$  wohldefiniert. Sie ist auch linear:  $(\lambda\phi + \mu\psi)|_W = \lambda\phi|_W + \mu\psi|_W$ . Ferner ist sie injektiv: denn ist  $\psi|_W = 0$ , so liegt  $\psi$  in  $W^\circ$ , weshalb  $\psi + W^\circ = 0 + W^\circ$ . ■

*Beispiel* Eine Linearform auf  $\mathbb{R}^3$ , die auf  $e_1$  verschwindet, ist das gleiche wie eine Linearform auf der  $(y, z)$ -Ebene: denn ist  $\phi(1, 0, 0) = 0$ , dann  $\phi(x, y, z) = \psi(y, z)$ , wobei  $\psi(y, z) := \phi(0, y, z)$ .

Jede Linearform auf der  $(x, z)$ -Ebene lässt sich zu einer auf ganz  $\mathbb{R}^3$  fortsetzen. Zum Beispiel lässt sich  $\phi(x, z) = 3x - 2z$  zu  $\psi(x, y, z) = 3x + y - 2z$  oder auch zu  $\chi(x, y, z) = 3x - 4y - 2z$  fortsetzen. Die Differenz zwischen  $\psi$  und  $\chi$  ist  $(x, y, z) \mapsto 5y$ , was auf der  $(x, z)$ -Ebene verschwindet.

## Duale Abbildungen

Ist  $\phi \in W^*$  eine Linearform auf  $W$  und  $f: V \rightarrow W$  eine lineare Abbildung, so ist auch die Verknüpfung  $\phi \circ f: V \rightarrow k$  eine lineare Abbildung, d.h.  $\phi \circ f$  ist eine Linearform auf  $V$  und deshalb ein Element von  $V^*$ .

*Definition* Sei  $f: V \rightarrow W$  eine lineare Abbildung. Die durch  $f^*(\phi) := \phi \circ f$  definierte Abbildung  $f^*: W^* \rightarrow V^*$  nennt man die *duale* Abbildung zu  $f$ .

*Beispiel* Sei  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  die lineare Abbildung

$$f(x, y) = (x + y, 3x - 2y, y - 2x).$$

Die duale Abbildung  $f^*$  ist eine Abbildung von  $(\mathbb{R}^3)^*$  nach  $(\mathbb{R}^2)^*$ . Für die Linearform  $\phi(x, y, z) = z - 2y - x$  auf  $\mathbb{R}^3$  wollen wir jetzt die Linearform  $f^*(\phi)$  auf  $\mathbb{R}^2$  berechnen. Es ist

$$f^*(\phi)(x, y) = \phi(x + y, 3x - 2y, y - 2x) = (y - 2x) - 2(3x - 2y) - (x + y) = 3y - 9x.$$

**Lemma 16.8** Sei  $f: V \rightarrow W$  eine lineare Abbildung zwischen endlich dimensionalen Unterräumen. Dann:

- a) Die Abbildung  $f^*: W^* \rightarrow V^*$  ist linear.
- b) Sei  $B$  bzw.  $C$  eine Basis von  $V$  bzw. von  $W$ , und sei  $B^*$  bzw.  $C^*$  die duale Basis des Dualraums  $V^*$  bzw.  $W^*$ . Ist  $A = {}_B M_C(f)$  die Matrix von  $f$  bezüglich den Basen  $B, C$ , so ist  ${}_{C^*} M_{B^*}(f^*) = A^T$ .
- c) Identifiziert  $V$  mit  $V^{**}$ , so ist  $f^{**} = f$ .
- d) Es ist  $\text{Kern}(f^*) = \text{Bild}(f)^\circ$  und  $\text{Bild}(f^*) = \text{Kern}(f)^\circ$ .

*Beweis.* a) Es ist

$$\begin{aligned} f^*(\lambda\phi + \mu\psi)(v) &= (\lambda\phi + \mu\psi)(f(v)) \\ &= \lambda\phi \circ f(v) + \mu\psi \circ f(v) = \lambda f^*(\phi)(v) + \mu f^*(\psi)(v). \end{aligned}$$

- b) Sei  $B = b_1, \dots, b_n$ , sei  $C = c_1, \dots, c_m$ , und sei  $D = {}_{C^*} M_{B^*}(f^*)$ . Es ist

$$\sum_{i=1}^n D_{ij} b_i^* = f^*(c_j^*) = c_j^* \circ f.$$

$$\text{Also } D_{ij} = c_j^*(f(b_i)) = c_j^*\left(\sum_{\ell=1}^m A_{\ell i} c_\ell\right) = A_{ji}.$$

- c) Folgt aus b), denn  $(A^T)^T = A$ .
- d) Ist  $\phi \in W^*$  ein Element aus  $(\text{Bild } f)^\circ$ , so ist  $\phi(f(v)) = 0$  für jedes  $v \in V$  und deshalb  $f^*(\phi) = 0$ . Somit ist  $(\text{Bild } f)^\circ \subseteq \text{Kern}(f^*)$ . Nach b) gilt  $\dim \text{Bild}(f^*) = \dim \text{Bild}(f)$ . Also

$$\begin{aligned} \dim \text{Kern}(f^*) &= \dim W - \dim \text{Bild}(f^*) \\ &= \dim W - \dim \text{Bild}(f) = \dim (\text{Bild}(f))^\circ. \end{aligned}$$

Ist  $\phi \in V^*$  der Gestalt  $\phi = f^*(\psi)$  für ein  $\psi \in W^*$ , so ist  $\phi(v) = \psi(f(v))$  für jedes  $v \in V$ . Ist  $f(v) = 0$ , dann  $\phi(v) = 0$ . Also  $\text{Bild}(f^*) \subseteq \text{Kern}(f)^\circ$ . Aber  $\dim \text{Bild}(f^*) = \dim \text{Bild}(f) = \dim V - \dim \text{Kern}(f) = \dim \text{Kern}(f)^\circ$ . ■

*Beispiel* Im obigen Beispiel hat  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  gilt:

$$\text{Matrix von } f = \begin{pmatrix} 1 & 1 \\ 3 & -2 \\ -2 & 1 \end{pmatrix}; \quad \text{Matrix von } f^* = \begin{pmatrix} 1 & 3 & -2 \\ 1 & -2 & 1 \end{pmatrix}.$$

Da  $f$  injektiv ist (seine Matrix hat ja Rang 2), ist  $f^*$  surjektiv. Die Linearform  $\chi(x, y, z) = x + y + z$  nimmt den Wert  $1 + 3 - 2 = 2$  auf  $(1, 3, -2) = f(e_1)$ . Somit liegt  $\chi$  nicht in  $\text{Kern}(f^*)$ , denn  $\chi$  annulliert  $\text{Bild}(f)$  nicht. Tatsächlich ist

$$f^*(\chi)(x, y) = \chi(x + y, 3x - 2y, y - 2x) = (x + y) + (3x - 2y) + (y - 2x) = 2x,$$

d.h.  $f^*(\chi) = 2e_1^* = 2e_1^* + 0e_2^*$ . Dies entspricht der Tatsache, dass

$$\begin{pmatrix} 1 & 3 & -2 \\ 1 & -2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \text{ ist.}$$

# 17 Bilinearformen

Sei  $V$  ein endlich dimensionaler  $k$ -Vektorraum.

## Grundbegriffe

*Definition* Eine *Bilinearform* auf  $V$  ist eine Abbildung  $b: V \times V \rightarrow k$ , die in beiden Argumenten linear ist. Das heißt, für alle  $x, x', y, y' \in V$ ,  $\lambda, \mu \in k$  gelten

$$\begin{aligned} b(\lambda x + \mu x', y) &= \lambda b(x, y) + \mu b(x', y) \quad \text{und} \\ b(x, \lambda y + \mu y') &= \lambda b(x, y) + \mu b(x, y'). \end{aligned}$$

Eine Bilinearform heißt

- *symmetrisch*, falls  $b(x, y) = b(y, x)$  gilt für alle  $x, y$ ;
- *schiefsymmetrisch*, falls  $b(x, y) = -b(y, x)$  gilt für alle  $x, y$ ; und
- *alternierend*, falls  $b(x, x) = 0$  gilt für alle  $x$ .

Durch  $(b+b')(x, y) = b(x, y) + b'(x, y)$ ,  $(\lambda b)(x, y) = \lambda \cdot b(x, y)$  erhält die Menge  $\text{Bil}^2(V)$  aller Bilinearformen auf  $V$  die Struktur eines  $k$ -Vektorraums. Die symmetrischen bzw. alternierenden Bilinearformen bilden einen Unterraum  $\text{Sym}^2(V)$  bzw.  $\text{Alt}^2(V)$  von  $\text{Bil}^2(V)$ .

*Beispiele* Das Standardskalarprodukt auf  $\mathbb{R}^n$  ist eine symmetrische Bilinearform. Das Standardskalarprodukt auf  $\mathbb{C}^n$  ist keine Bilinearform, sondern eine sogenannte *Sesquilinearform*, denn  $\langle z, iw \rangle = -i\langle z, w \rangle$ .

Die Bilinearform  $b((x_1, x_2), (y_1, y_2)) = x_1y_2 + x_2y_1$  auf  $k^2$  ist symmetrisch. Die Bilinearform  $b'((x_1, x_2), (y_1, y_2)) = x_1y_2 - x_2y_1$  auf  $k^2$  ist alternierend und schiefsymmetrisch. Für  $k = \mathbb{F}_2$  stimmen diese beiden Bilinearformen überein.

Körper mit  $1+1 = 0$  haben eine Sonderrolle bei Bilinearformen, wie dieses Beispiel bereits erahnen lässt. Häufig beschränkt man sich auf Körper mit  $1 + 1 \neq 0$ , d.h. mit  $\text{char}(k) \neq 2$ .

*Definition* Die *Charakteristik*  $\text{char}(k)$  eines Körpers  $k$  ist die kleinste Zahl  $n \geq 1$  derart, dass  $n = 0$  in  $k$  gilt, d.h.  $\underbrace{1 + 1 + \dots + 1}_{n \text{ fach}} = 0$ . Gibt es kein solches  $n$ , so setzt man  $\text{char}(k) = 0$ .

**Hilfssatz 2** Die Charakteristik  $\text{char}(k)$  ist entweder 0 oder eine Primzahl. Jeder dieser Werte kommt auch vor.

*Beweis.* In einem Körper ist  $1 \neq 0$ , also  $\text{char}(k) \neq 1$ . Also  $\text{char}(k) \geq 2$ ; oder  $\text{char}(k) = 0$ , was z.B. für  $k = \mathbb{Q}, \mathbb{R}, \mathbb{C}$  der Fall ist. Ist  $\text{char}(k) = n \geq 2$ , so muss  $n$  eine Primzahl sein: ist  $n = ab$  mit  $2 \leq a, b < n$ , dann ist  $ab = 0$  in  $k$ , obwohl  $a, b \neq 0$ , ein Widerspruch. Es ist  $\text{char}(\mathbb{F}_p) = p$ . Insbesondere ist  $\text{char}(\mathbb{F}_2) = 2$ . ■

**Lemma 17.1** a) Jede alternierende Bilinearform ist schiefsymmetrisch. Ist  $\text{char}(k) \neq 2$ , so ist jede schiefsymmetrische Bilinearform auch alternierend.

b) Ist  $\text{char}(k) \neq 2$ , so ist  $\text{Bil}^2(V) = \text{Sym}^2(V) \oplus \text{Alt}^2(V)$ . Etwas genauer: für jedes  $b \in \text{Bil}^2(V)$  ist  $b = b_s + b_a$ , wobei

$$b_s(x, y) = \frac{1}{2} (b(x, y) + b(y, x)) \quad b_a(x, y) = \frac{1}{2} (b(x, y) - b(y, x)),$$

mit  $b_s$  symmetrisch und  $b_a$  alternierend.

*Beweis.* Ist  $\text{char}(k) \neq 2$ , so gilt  $x = -x$  nur für  $x = 0$ . Ist  $\text{char}(k) = 2$ , so gilt  $x = -x$  für jedes  $x \in k$ .

a) Ist  $b$  alternierend, so ist

$$\begin{aligned} 0 &= b(x + y, x + y) = b(x, x + y) + b(y, x + y) \\ &= b(x, x) + b(x, y) + b(y, x) + b(y, y) = b(x, y) + b(y, x), \end{aligned}$$

weshalb  $b$  auch schiefsymmetrisch ist. Ist  $b$  schiefsymmetrisch und  $\text{char}(k) \neq 2$ , so ist  $b(x, x) = -b(x, x)$  und deshalb  $b(x, x) = 0$ .

b) Es ist  $\text{Bil}^2(V) = \text{Sym}^2(V) + \text{Alt}^2(V)$ , denn offensichtlich ist  $b_s$  symmetrisch und  $b_a$  alternierend. Ist  $b \in \text{Sym}^2(V) \cap \text{Alt}^2(V)$ , so ist  $b$  symmetrisch und schiefsymmetrisch, also  $b(x, y) = b(y, x) = -b(x, y)$ , also  $b(x, y) = 0$ , also  $b = 0$ . ■

## Die Matrix einer Bilinearform

*Bezeichnung* Sei  $v_1, \dots, v_n$  eine Basis von  $V$ . Die Matrix  $A \in M_n(k)$  gegeben durch  $A_{ij} = b(v_i, v_j)$  heißt die Matrix der Bilinearform  $b$  bezüglich der Basis  $v_1, \dots, v_n$  von  $V$ .

Nun sei  $x = \sum_{i=1}^n \lambda_i v_i$ ,  $y = \sum_{i=1}^n \mu_i v_i$ . Dann

$$b(x, y) = \sum_{i,j=1}^n b(\lambda_i v_i, \mu_j v_j) = \sum_{i,j=1}^n \lambda_i A_{ij} \mu_j.$$

Also

$$b(x, y) = \underline{\lambda}^T \cdot A \cdot \underline{\mu}, \quad (17.2)$$

wobei  $\underline{\lambda} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$  bzw.  $\underline{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$  der Koordinatenvektor von  $x$  bzw.  $y$  bezüglich der Basis  $v_1, \dots, v_n$  ist.

*Beispiel* Die alternierende Bilinearform  $b((x_1, x_2), (y_1, y_2)) = x_1y_2 - x_2y_1$  hat Matrix  $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$  bezüglich der Standardbasis des  $k^2$ . Es ist

$$b((x_1, x_2), (y_1, y_2)) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

## Bilinearformen und der Dualraum

**Lemma 17.3** Sei  $b$  eine Bilinearform auf  $V$ .

- a) Durch  $b_\ell(x) = b(x, \_)$  und  $b_r(y) = b(\_, y)$  – d.h. durch  $b_\ell(x)(y) = b(x, y)$  und  $b_r(y)(x) = b(x, y)$  – erhält man lineare Abbildungen  $b_\ell, b_r: V \rightarrow V^*$ .
- b) Sei  $A \in M_n(k)$  die Matrix von  $b$  bezüglich einer Basis  $C: c_1, \dots, c_n$  von  $V$ . Dann ist  $A^T$  die Matrix von  $b_\ell$  und  $A$  ist die Matrix von  $b_r$  bezüglich der Basis  $C$  von  $V$  und der Dualbasis  $C^*: c_1^*, \dots, c_n^*$  von  $V^*$ .

*Beweis.* a) Es ist  $b_\ell(x) \in V^*$ , denn wegen Bilinearität von  $b$  ist  $y \mapsto b(x, y)$  linear. Die Zuordnung  $y \mapsto b_r(y)$  ist linear, denn

$$\begin{aligned} b_r(\lambda y + \mu y')(x) &= b(x, \lambda y + \mu y') \\ &= \lambda b(x, y) + \mu b(x, y') = (\lambda b_r(y) + \mu b_r(y'))(x). \end{aligned}$$

- b) Es ist  $b_\ell(c_i)(c_j) = A_{ij}$  und  $b_r(c_i)(c_j) = A_{ji}$ , also

$$b_\ell(c_i) = \sum_{j=1}^n A_{ij} c_j^* = \sum_j A_{ji}^T c_j^*$$

$$\text{und } b_r(c_i) = \sum_{j=1}^n A_{ji} c_j^*. \quad \blacksquare$$

**Korollar 17.4** Für eine Bilinearform  $b$  auf  $V$  sind die folgenden drei Aussagen äquivalent.

- a) Zu jedem  $0 \neq x \in V$  gibt es ein  $y \in V$  mit  $b(x, y) \neq 0$ .
- b) Zu jedem  $0 \neq y \in V$  gibt es ein  $x \in V$  mit  $b(x, y) \neq 0$ .
- c)  $b_\ell, b_r: V \rightarrow V^*$  sind Isomorphismen.

Gelten diese Aussagen, so heißt die Bilinearform  $b$  nicht ausgeartet.

*Beweis.* Die erste Aussage besagt, dass  $b_\ell$  injektiv ist; die zweite besagt, dass  $b_r$  injektiv ist. Da  $\dim V = \dim V^*$  ist eine lineare Abbildung  $V \rightarrow V^*$  genau dann injektiv, wenn sie ein Isomorphismus ist. Da die Matrix von  $b_\ell$  die Transponierte der Matrix von  $b_r$  ist, haben beide lineare Abbildungen den gleichen Rang, d.h. entweder sind beide Isomorphismen, oder keine ist.  $\blacksquare$

*Beispiel* Die Bilinearform  $b((x_1, x_2), (y_1, y_2))$  ist nicht ausgeartet, denn die Matrix  $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$  hat Rang 2. Dagegen ist die Bilineaform  $b((x_1, x_2), (y_1, y_2)) = x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2$  ausgeartet, denn  $b((1, -1), (y_1, y_2)) = 0$  für alle  $y_1, y_2$ . Entsprechend ist die Matrix  $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$  vom Rang  $1 < 2$ .

## Symmetrische und alternierende Bilinearformen

Wir werden uns hauptsächlich mit Bilinearformen beschäftigen, die entweder symmetrisch oder alternierend sind.

**Lemma 17.5** *Ist  $n = \dim(V)$ , so gelten*

$$\dim \text{Bil}^2(V) = n^2 \quad \dim \text{Sym}^2(V) = \frac{n(n+1)}{2} \quad \dim \text{Alt}^2(V) = \frac{n(n-1)}{2}.$$

*Beweis.* Sei  $v_1, \dots, v_n$  eine Basis von  $V$ . Jede Bilinearform  $b \in \text{Bil}^2(V)$  hat eine Matrix  $A \in M_n(k)$  bezüglich dieser Basis. Umgekehrt erhält man aus einer Matrix  $A$  eine Bilinearform  $b$  gegeben durch

$$b\left(\sum_i \lambda_i v_i, \sum_j \mu_j v_j\right) = \underline{\lambda}^T \cdot A \cdot \underline{\mu}.$$

Die Vektorräume  $\text{Bil}^2(V)$  und  $M_n(k)$  sind also isomorph, und  $\dim \text{Bil}^2(V) = n^2$ .

Die Bilinearform  $b$  ist genau dann symmetrisch, wenn die zugehörige Matrix  $A$  symmetrisch ist, d.h.  $A^T = A$ . Eine Basis der symmetrischen Matrizen bilden die  $\frac{n(n-1)}{2}$  Matrizen  $S(r, s)$  für  $1 \leq r < s \leq n$  zusammen mit den  $n$  Matrizen  $E(r)$  für  $1 \leq r \leq n$ ; wobei  $S(r, s)_{rs} = S(r, s)_{sr} = 1$  und  $S(r, s)_{ij} = 0$  sonst, sowie  $E(r)_{rr} = 1$  und  $E(r)_{ij} = 0$  sonst.

Eine Bilinearform  $b$  ist genau dann alternierend, wenn die zugehörige Matrix  $A$  schief-symmetrisch ist ( $A^T = -A$ ) und die Diagonaleinträge alle 0 sind. Eine Basis der Matrizen dieser Art bilden die  $\frac{n(n-1)}{2}$  Matrizen  $A(r, s)$  für  $1 \leq r < s \leq n$ , wobei  $A(r, s)_{rs} = 1$ ,  $A(r, s)_{sr} = -1$  und  $A(r, s)_{ij} = 0$  sonst. ■

*Bemerkung* Hier sind einige dieser Matrizen für  $n = 3$ :

$$S(1, 3) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad E(2) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad A(2, 3) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

*Bezeichnung* Sei  $b$  eine Bilinearform auf  $V$ , und  $F: V \rightarrow V$  ein Endomorphismus. Gilt  $b(F(x), F(y)) = b(x, y)$  für alle  $x, y \in V$ , so heißt der Endomorphismus  $F$  *orthogonal* bezüglich  $b$ .

**Lemma 17.6** *Sei  $b$  eine nicht ausgeartete Bilinearform auf  $V$ . Dann ist jeder bezüglich  $b$  orthogonale Endomorphismus invertierbar, und die orthogonalen Endomorphismen bilden eine Gruppe  $O(V, b)$ , die Orthogonalgruppe der Bilinearform  $b$ .*



*Beweis.* Sei  $F$  ein orthogonaler Endomorphismus. Für Invertierbarkeit reicht es zu zeigen, dass  $\text{Kern}(F) = 0$  ist. Sei also  $0 \neq x \in V$ . Da  $b$  nicht ausgeartet ist, gibt es ein  $y \in V$  mit  $b(x, y) \neq 0$ . Also  $b(F(x), F(y)) = b(x, y) \neq 0$ , weshalb  $F(x) \neq 0$ . Also ist  $F$  ein Isomorphismus.

Sei  $u = F^{-1}(x)$ ,  $v = F^{-1}(y)$ . Dann  $b(u, v) = b(F(u), F(v)) = b(x, y)$ , d.h.  $b(F^{-1}(x), F^{-1}(y)) = b(x, y)$  und  $F^{-1}$  ist orthogonal. Auch die Verknüpfung zweier orthogonaler Endomorphismen ist orthogonal. ■

## Kongruente Matrizen und den Rang einer Bilinearform

Wann gehören zwei Matrizen zur gleichen Bilinearform? Wenn sie *kongruent* sind.

*Definition* Zwei Matrizen  $A, B \in M_n(k)$  heißen *kongruent*, wenn es eine invertierbare  $(n \times n)$ -Matrix  $S$  gibt derart, dass  $B = S^T A S$  gilt.

**Lemma 17.7** a) *Kongruenz ist eine Äquivalenzrelation auf  $M_n(k)$ .*

b) *Sind  $A, B$  die Matrizen einer Bilinearform  $b \in \text{Bil}^2(V)$  bezüglich zwei verschiedener Basen, so sind  $A, B$  kongruent. Umgekehrt gilt: ist  $A$  die Matrix von  $b$  bezüglich einer Basis, und ist  $B$  zu  $A$  kongruent, so ist  $B$  die Matrix von  $b$  bezüglich einer zweiten Basis von  $V$ .*

c) *Kongruente Matrizen haben den gleichen Rang.*

*Beweis.* a) Transitivität: ist  $B = S^T A S$  und  $C = R^T B R$ , so ist  $C = R^T S^T A S R = (S R)^T A (S R)$ .

b) Ist  $\underline{\lambda}$  bzw.  $\underline{\mu}$  der Koordinatenvektor von  $v$  bzw.  $w$  bezüglich der ersten Basis, so gibt es eine (invertierbare) Basiswechsellmatrix  $S$  derart, dass  $S \underline{\lambda}$  bzw.  $S \underline{\mu}$  der Koordinatenvektor von  $v$  bzw.  $w$  bzgl. der zweiten Basis ist. Nach Gleichung (17.2) ist

$$b(v, w) = \underline{\lambda}^T A \underline{\mu} = \underline{\lambda}^T S^T B S \underline{\mu}$$

für alle  $v, w$ . Also  $A = S^T B S$ . Da jede invertierbare Matrix als Basiswechsellmatrix vorkommt, gilt auch der zweite Teil.

c) Da  $S$  invertierbar ist, gilt:

$$\text{Nullraum}(S^T B S) = \{S^{-1} \cdot v \mid B \cdot v = 0\}.$$

Dies ist isomorph zum Nullraum von  $B$  und hat somit die gleiche Dimension. Aber  $\text{Rang} = n - \dim \text{Nullraum}$ . ■

**Korollar 17.8** *Der Rang  $\text{Rang}(b)$  der Matrix einer Bilinearform  $b \in \text{Bil}^2(V)$  hängt nur von  $b$  ab, nicht von der gewählten Basis von  $V$ .* ■

## Senkrechte Vektoren

**Hilfssatz 3** Sei  $b \in \text{Bil}^2(V)$ . Die Relation  $x \perp y$  auf  $V$  gegeben durch

$$x \perp y \quad \text{falls} \quad b(x, y) = 0$$

ist genau dann symmetrisch ( $x \perp y \Leftrightarrow y \perp x$ ), wenn  $b$  entweder symmetrisch oder alternierend ist.

*Beweis.* Alternierende Formen sind schiefsymmetrisch. Ist  $b$  symmetrisch oder schiefsymmetrisch, so ist die Relation symmetrisch. Für die Umkehrrichtung nehmen wir an, dass  $b$  weder symmetrisch noch alternierend ist, und wollen zeigen, dass  $x \perp y$  keine symmetrische Relation ist.

Da  $b$  nicht symmetrisch ist, gibt es  $r, s$  mit  $b(r, s) \neq b(s, r)$ . Da  $b$  nicht alternierend ist, gibt es  $t$  mit  $b(t, t) \neq 0$ .

Gibt es  $a, c$  mit  $b(a, c) \neq b(c, a)$  und  $b(a, a) \neq 0$ , so sind wir fertig: es ist  $b(x, y) = 0$  für  $x = a, y = c - \frac{b(a, c)}{b(a, a)}a$ , aber  $b(y, x) = b(c, a) - b(a, c) \neq 0$ .

Ist  $\text{char}(k) = 2$ , so ist  $b(r, s) + b(s, r) \neq 0$ , also muss  $b(a, a) \neq 0$  sein für mindestens ein  $a \in \{r, s, r+s\}$ : denn ist  $b(r, r) = b(s, s) = 0$ , so ist  $b(r+s, r+s) = b(r, s) + b(s, r) \neq 0$ . Für jedes solche  $a$  finden wir ein  $b \in \{r, s\}$  mit  $b(a, c) \neq b(c, a)$ , also sind wir fertig.

Ist  $\text{char}(k) \neq 2$ , so ist ohne Einschränkung  $b(r, r) = b(s, s) = 0, b(r, t) = b(t, r)$  und  $b(s, t) = b(t, s)$ , sonst wären wir wie oben fertig. Ist  $b(t, r) = 0$ , so setzen wir  $a = t+r, c = s$ : dann  $b(a, a) = b(t, t) \neq 0$  und  $b(a, c) - b(c, a) = b(r, s) - b(s, r) \neq 0$ . Ist  $b(t, r) \neq 0$ , so ist  $b(x, y) = 0$  für  $x = r, y = s - \frac{b(r, s)}{b(r, t)}t$ , aber  $b(y, x) = b(s, r) - b(r, s) \neq 0$ . ■

Von nun an also sei  $b \in \text{Sym}^2(V) \cup \text{Alt}^2(V)$  eine Bilinearform, die entweder symmetrisch oder alternierend ist, weshalb  $x \perp y$  eine symmetrische Relation ist.

*Bezeichnung* Sei  $b \in \text{Sym}^2(V) \cup \text{Alt}^2(V)$  und  $T \subseteq V$  eine Teilmenge. Dann wird  $T^\perp \subseteq V$  definiert durch

$$T^\perp = \{x \in V \mid x \perp y \text{ für jedes } y \in T\}.$$

Falls  $T = \{x\}$  aus nur einem Element besteht, schreibt man  $x^\perp$  für  $T^\perp$ .

### Lemma 17.9

- a)  $T^\perp$  ist ein Unterraum von  $V$ .
- b) Ist  $T_1 \subseteq T_2$ , so ist  $T_1^\perp \supseteq T_2^\perp$ .
- c) Es ist  $T^\perp = W^\perp$  für  $W = \text{Spann}(T)$ .
- d)  $b$  ist genau dann nicht ausgeartet, wenn  $V^\perp = 0$  ist.

Ist  $b$  nicht ausgeartet, so gelten außerdem:

e)  $b_\ell$  und  $b_r$  sind Isomorphismen von  $T^\perp$  nach  $T^\circ \subseteq V^*$ .

f) Es ist  $\dim W + \dim W^\perp = \dim V$  sowie  $(V^\perp)^\perp = V$ .

*Beweis.* Wegen  $b_\ell(x)(y) = b(x, y)$  folgt:

$$T^\perp = \{x \mid b_\ell(x) \in T^\circ\} = b_\ell^{-1}(T^\circ). \quad (*)$$

Ganz allgemein gilt: ist  $f: V \rightarrow W$  linear und  $U \subseteq W$  ein Unterraum, so ist auch  $f^{-1}(U) \subseteq V$  ein Unterraum. Also ist  $T^\perp \subseteq V$  ein Unterraum. Wegen (\*) und den Eigenschaften von  $T^\circ$  (Lemma 16.6) folgen b) und c). Es ist

$$V^\perp = \{x \in V \mid b(x, y) = 0 \text{ für alle } y \in V\},$$

also folgt d) aus der Definition von „nicht ausgeartet“. Ist  $b$  nicht ausgeartet, so sind  $b_\ell, b_r$  Isomorphismen von  $V$  nach  $V^*$ , und  $T^\perp = b_\ell^{-1}(T^\circ) = b_r^{-1}(T^\circ)$ . Allgemein gilt: ist  $f: V \rightarrow W$  ein Isomorphismus und  $U \subseteq W$  ein Unterraum, so ist  $f$  auch ein Isomorphismus von  $f^{-1}(U)$  nach  $U$ . Also gilt e). Der erste Teil von f) folgt aus der Dimensionsformel für  $W^\circ$ ; für den zweiten Teil stellt man fest, dass  $W \subseteq (W^\perp)^\perp$ , und dass beide Unterräume die gleiche Dimension haben. ■

*Beispiel* Betrachten wir die nicht ausgearteten Bilinearformen  $b, b'$  auf  $\mathbb{R}^2$  gegeben durch

$$\begin{aligned} b((x_1, x_2), (y_1, y_2)) &= x_1y_1 + x_2y_2 && \text{(symmetrisch)} \\ b'((x_1, x_2), (y_1, y_2)) &= x_1y_2 - x_2y_1 && \text{(alternierend)}. \end{aligned}$$

Es ist dann Bezüglich  $b$  ist  $(x\text{-Achse})^\perp = y\text{-Achse}$ ; bezüglich  $b'$  ist  $(x\text{-Achse})^\perp = x\text{-Achse}$ .

**Lemma 17.10** Ist  $b \in \text{Sym}^2(V) \cup \text{Alt}^2(V)$ , so ist  $\dim V^\perp = \dim(V) - \text{Rang}(b)$ .

*Beweis.* Es ist  $V^\perp = \text{Kern}(b_r)$ . Die Matrix von  $b$  ist gleichzeitig die Matrix von  $b_r$ . Somit ist  $\text{Rang}(b_r) = \text{Rang}(b)$ . Das Ergebnis folgt aus der Dimensionsformel für die lineare Abbildung  $b_r$ . ■

## Symmetrische Bilinearformen

**Satz 17.11** Sei  $b$  eine symmetrische Bilinearform auf  $V$ . Es gelte  $\text{char}(k) \neq 2$ . Dann gibt es eine orthogonale Basis für  $V$ , d.h. eine Basis  $v_1, \dots, v_n$  derart, dass  $b(v_i, v_j) = 0$  ist für alle  $j \neq i$ .

*Beweis.* Induktion über  $n = \dim(V)$ . Den Induktionsanfang stellt der Fall  $b = 0$  dar, was insbesondere für  $n = 0$  der Fall ist. In diesem Fall ist *jede* Basis  $v_1, \dots, v_n$  von  $V$  orthogonal bzgl.  $b$ .

Für  $\text{char}(k) \neq 2$  ist  $\text{Sym}^2(V) \cap \text{Alt}^2(V) = \{0\}$ . Ist also  $b \neq 0$ , so gibt es ein  $v_1 \in V$  mit  $b(v_1, v_1) \neq 0$ . Mit  $W = v_1^\perp$  ist also  $v_1 \notin W$  und deshalb  $V = \text{Spann}(v_1) \oplus W$ . Auch als Bilinearform auf  $W$  ist  $b$  symmetrisch, also gibt es nach Induktionsannahme eine orthogonale Basis  $v_2, \dots, v_n$  für  $W$ . Also ist  $v_1, \dots, v_n$  eine orthogonale Basis für  $V$ . ■

*Beispiel* Betrachten wir die symmetrische Bilinearform auf  $\mathbb{R}^3$  gegeben durch die Matrix  $\begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & 1 \\ 0 & 1 & -1 \end{pmatrix}$ . Es ist  $b(e_1, e_1) = 1$ , also dürfen wir  $v_1 = e_1$  wählen. Es ist  $b(e_1, e_2) = 2$  und  $b(e_1, e_3) = 0$ , also ist  $e_2 - 2e_1, e_3$  eine Basis für  $v_1^\perp$ . Es ist  $b(e_3, e_3) = -1$ , also dürfen wir  $v_2 = e_3$  wählen. Es ist  $b(e_3, e_2 - 2e_1) = 1$ , also ist  $e_2 + e_3 - 2e_1$  eine Basis für  $\{v_1, v_2\}^\perp$ . Somit bilden  $v_1 = e_1, v_2 = e_3$  und  $v_3 = e_2 + e_3 - 2e_1$  eine orthogonale Basis von  $\mathbb{R}^3$ . Wir mussten nicht einmal  $b(v_3, v_3)$  berechnen: es ist aber

$$\begin{aligned} b(v_3, v_3) &= b(e_2, e_2) + b(e_3, e_3) + 4b(e_1, e_1) + 2b(e_2, e_3) - 4b(e_1, e_3) - 4b(e_1, e_2) \\ &= 3 - 1 + 4 + 2 - 4 \cdot 0 - 8 = 0. \end{aligned}$$

Also  $v_3 \in V^\perp$ . Ist dagegen  $w = \lambda v_1 + \mu v_2 + \nu v_3 \in V^\perp$ , so ist  $0 = b(w, v_1) = \lambda$  und  $0 = b(w, v_2) = -\mu$ , also  $w \in \text{Spann}(v_3)$ , weshalb  $v_3$  eine Basis von  $V^\perp$  ist. Dies entspricht der Tatsache, dass der Rang der Matrix  $2 = 3 - 1$  beträgt.

Bezüglich dieser Orthogonalbasis ist  $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$  die Matrix der Bilinearform.

## Reell symmetrische Bilinearformen: der Trägheitssatz

*Bezeichnung (nicht standard)* Sei  $C: c_1, \dots, c_n$  eine Orthogonalbasis des  $\mathbb{R}$ -Vektorraums  $V$  bezüglich der symmetrischen Bilinearform  $b$ . Für jedes  $1 \leq i \leq n$  gilt genau eins aus  $b(c_i, c_i) > 0$ ,  $b(c_i, c_i) < 0$ ,  $b(c_i, c_i) = 0$ . Setzen wir

$$\begin{aligned} \text{Pos}(C) &= \{c_i \mid b(c_i, c_i) > 0\} & a_+(b) &= |\text{Pos}(C)| & (\text{d.h. Anzahl der Elemente}) \\ \text{Neg}(C) &= \{c_i \mid b(c_i, c_i) < 0\} & a_-(b) &= |\text{Neg}(C)| \\ \text{Null}(C) &= \{c_i \mid b(c_i, c_i) = 0\} & a_0(b) &= |\text{Null}(C)|. \end{aligned}$$

*Definition* Die Signatur einer reell symmetrischen Bilinearform  $b$  wird definiert durch  $\text{Signatur}(b) = a_+(b) - a_-(b)$ .

*Beispiel* Hat  $b \in \text{Sym}^2(\mathbb{R}^4)$  die Matrix  $\text{diag}(1, 0, -2, 3)$  bezüglich der Orthogonalbasis  $C: c_1, c_2, c_3, c_4$ , so ist  $\text{Pos}(C) = \{c_1, c_4\}$ ,  $\text{Neg}(C) = \{c_3\}$  und  $\text{Null}(C) = \{c_2\}$ . Also  $a_+(b) = 2$ ,  $a_-(b) = a_0(b) = 1$ ,  $\text{Rang}(b) = 3$  und  $\text{Signatur}(b) = 2 - 1 = 1$ .

**Der Trägheitssatz von Sylvester** Sei  $b \in \text{Sym}^2(V)$  eine reell symmetrische Bilinearform.

- a) Die Zahlen  $a_+$ ,  $a_-$ ,  $a_0$  sowie die Signatur von  $b$  hängen nur von  $b$  ab, und nicht von der Wahl der Orthogonalbasis. Insbesondere ist  $a_0(b) = \dim V^\perp$  und  $a_+(b) + a_-(b) = \text{Rang}(b)$ .
- b) Jede symmetrische Matrix  $A \in M_n(\mathbb{R})$  ist kongruent zu einer Blockmatrix der Gestalt  $\begin{pmatrix} E_x & 0 & 0 \\ 0 & -E_y & 0 \\ 0 & 0 & 0 \end{pmatrix}$ , für  $x = \frac{r+s}{2}$  und  $y = \frac{r-s}{2}$ , wobei  $r$  bzw.  $s$  der Rang bzw. die Signatur der entsprechende Bilinearform ist.
- c) Zwei symmetrische Matrizen  $A, B \in M_n(\mathbb{R})$  sind genau dann kongruent, wenn die zwei zugehörigen Bilinearformen den gleichen Rang und die gleiche Signatur haben.

*Beweis.* Zuerst halten wir fest: Da  $c_1, \dots, c_n$  eine Orthogonalbasis ist, ist

$$b(c_i, v) = \lambda_i b(c_i, c_i) \qquad b(v, v) = \sum_{i=1}^n \lambda_i^2 b(c_i, c_i) \qquad (*)$$

für alle  $i$  und für alle  $v = \sum_{j=1}^n \lambda_j c_j \in V$ .

Ist also  $b(c_i, c_i) = 0$ , so ist  $b(c_i, v) = 0$  für alle  $v \in V$  und deshalb  $c_i \in V^\perp$ . Ist dagegen  $v = \sum_{j=1}^n \lambda_j c_j \in V^\perp$ , so ist  $\lambda_i = 0$  für alle  $i$  mit  $b(c_i, c_i) \neq 0$ . Also  $V^\perp = \text{Spann Null}(C)$  und  $a_0 = \dim V^\perp$ , woraus folgt  $a_+ + a_- = \dim V - \dim V^\perp = \text{Rang}(b)$ .

Angenommen  $D: d_1, \dots, d_n$  ist eine weitere Orthogonalbasis von  $V$ . Setzen wir

$$\begin{aligned} V_+ &= \text{Spann Pos}(C) & V_- &= \text{Spann}(\text{Neg}(C) \cup \text{Null}(C)) \\ W_+ &= \text{Spann Pos}(D) & W_- &= \text{Spann}(\text{Neg}(D) \cup \text{Null}(D)). \end{aligned}$$

Es ist also  $V = V_+ \oplus V_- = W_+ \oplus W_-$ . Wegen (\*) ist  $b(v, v) > 0$  für jedes  $0 \neq v \in V_+ \cup W_+$  sowie  $b(v, v) \leq 0$  für jedes  $v \in V_- \cup W_-$ . Für a) ist noch zu zeigen, dass  $\dim W_+ = \dim V_+$ . Ohne Einschränkung ist  $\dim W_+ \geq \dim V_+$  und deshalb  $\dim W_- \leq \dim V_-$ . Ist  $\dim W_+ > \dim V_+$ , so ist  $\dim W_+ + \dim V_- > \dim V$  und deshalb  $W_+ \cap V_- \neq \emptyset$ . Ist aber  $0 \neq v \in W_+ \cap V_-$ , so ist  $b(v, v) > 0$  und  $b(v, v) \leq 0$ , ein Widerspruch. Also  $\dim W_+ = \dim V_+$ , und a) ist bewiesen.

Zu b): Sei  $b$  die Bilinearform auf  $\mathbb{R}^n$  mit Matrix  $A$ . Aus a) folgt  $x = a_+(b)$  und  $y = a_-(b)$ . Nachdem man die Reihenfolge der Elemente der Orthogonalbasis  $v_1, \dots, v_n$  geändert hat, ist  $b(v_i, v_i) > 0$  für  $1 \leq i \leq x$ ;  $b(v_i, v_i) < 0$  für  $x+1 \leq i \leq x+y$ ; und  $b(v_i, v_i) = 0$  für  $i > x+y$ . Ersetzt man  $v_i$  durch  $v_i/\sqrt{|b(v_i, v_i)|}$  für  $1 \leq i \leq x+y$ , so hat die Matrix von  $b$  die erwünschte Gestalt.

Zu c): Wegen a) müssen kongruente symmetrische Matrizen die gleiche Signatur und den gleichen Rang haben. Wegen b) sind symmetrische Matrizen mit der gleichen Signatur und dem gleichen Rang kongruent. ■

*Beispiel* Sei  $b \in \text{Sym}^2(\mathbb{R}^2)$  gegeben durch die Matrix  $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ , dann ist  $e_1, e_2$  bereits eine Orthogonalbasis. Der Rang ist 2, und die Signatur ist  $1 - 1 = 0$ . Für  $v = (1, 1)$  ist  $b(v, v) = 0$ , aber  $v \neq (\mathbb{R}^2)^\perp$ , denn  $b(e_1, v) = 1$ . Aufgrund der Gleichungen (\*) im Beweis des Trägheitssatzes kann also  $v$  keine Orthogonalbasis von  $\mathbb{R}^2$  angehören.

*Beispiel aus der Physik* In der speziellen Relativitätslehre beschäftigt man sich mit einem „Skalarprodukt“ auf der Raumzeit  $\mathbb{R}^4$ , das aber kein Skalarprodukt ist, sondern eine symmetrische Bilinearform mit Rang 4 und Signatur 2: die

$$\text{Matrix ist } \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -c^2 \end{pmatrix}.$$

*Beispiel* Wir berechnen Rang, Signatur und eine Orthogonalbasis von  $b \in \text{Sym}^2(\mathbb{R}^4)$  mit Matrix  $\begin{pmatrix} -3 & 2 & 2 & 1 \\ 2 & 0 & 0 & 0 \\ 3 & 0 & 3 & 2 \\ 1 & 0 & 2 & 1 \end{pmatrix}$ .

Durch Gaußsche Elimination erkennt man, dass die Matrix den Rang 4 hat, also  $\text{Rang}(b) = 4$ . Wir konstruieren jetzt eine Orthogonalbasis  $v_1, v_2, v_3, v_4$ . Es ist  $b(e_4, e_4) = 1$ , also wählen wir  $v_1 = e_4$ . Eine Basis für  $v_1^\perp$  ist  $e_1 - e_4, e_2, e_3 - 2e_4$ . Wir suchen ein  $v_2 \in v_1^\perp$  mit  $b(v_2, v_2) \neq 0$ . Wegen

$$b(e_3 - 2e_4, e_3 - 2e_4) = b(e_3, e_3) - 4b(e_3, e_4) + 4b(e_4, e_4) = 3 - 4 \cdot 2 + 4 \cdot 1 = -1$$

wählen wir  $v_2 = e_3 - 2e_4$ . Bereits  $e_1 - e_4, e_2$  bilden eine Basis für  $\{v_1, v_2\}^\perp = \{v \in v_1^\perp \mid v \perp v_2\}$ . Wir dürfen nicht  $v_3 = e_2$  wählen, denn  $b(e_2, e_2) = 0$  obwohl  $e_2 \notin V^\perp$ . Aber  $v_3 = e_1 - e_4$  geht schon, denn  $b(e_1 - e_4, e_1 - e_4) = -4$ . Als  $v_4$  dürfen wir ein beliebiges Basiselement von  $\{v_1, v_2, v_3\}^\perp = \{v \in \{v_1, v_2\}^\perp \mid v \perp v_3\}$  nehmen. Wegen  $b(e_2, v_3) = 2$  und  $b(v_3, v_3) = -4$  wählen wir  $v_4 = 2e_2 + v_3 = e_1 + 2e_2 - e_4$ . Dann  $b(v_4, v_4) = 4$ . Eine Orthogonalbasis ist also

$$v_1 = e_4 \quad v_2 = e_3 - 2e_4 \quad v_3 = e_1 - e_4 \quad v_4 = e_1 + 2e_2 - e_4.$$

Bezüglich dieser Basis hat diese Bilinearform die Matrix  $\text{diag}(1, -1, -4, 4)$ . Die Signatur beträgt also 0. Bezüglich der Basis  $e_4, \frac{1}{2}(e_1 + 2e_2 - e_4), e_3 - 2e_4, \frac{1}{2}(e_1 - e_4)$  beträgt die Matrix  $\text{diag}(1, 1, -1, -1) = \begin{pmatrix} E_2 & 0 \\ 0 & -E_2 \end{pmatrix}$ .

## Quadratische Formen

*Definition* Sei  $k$  ein Körper und  $V$  ein endlich dimensionaler  $k$ -Vektorraum. Eine Abbildung  $q: V \rightarrow k$  heißt eine *quadratische Form* auf  $V$ , falls gelten:

- $q(\lambda v) = \lambda^2 q(v)$  für alle  $\lambda \in k, v \in V$ ;
- $\beta(v, w) := q(v + w) - q(v) - q(w)$  definiert eine symmetrische Bilinearform auf  $V$ .

Beachten Sie, dass quadratische Formen *nicht* linear sind.

*Beispiel*  $q(x_1, x_2) = x_1^2 + x_1x_2$  ist eine quadratische Form auf  $\mathbb{R}^2$ ; die Bilinearform  $\beta$  ist  $\beta((x_1, x_2), (y_1, y_2)) = 2x_1y_1 + x_1y_2 + x_2y_1$ .

**Lemma 17.12** *Ist  $b \in \text{Sym}^2(V)$ , so definiert  $q(v) = b(v, v)$  eine quadratische Form auf  $V$ . Ist  $\text{char}(k) \neq 2$ , so ist die Abbildung*

$$\begin{aligned} \text{Sym}^2(V) &\longrightarrow \{\text{Quadratische Formen auf } V\} \\ b &\longmapsto q \end{aligned}$$

eine Bijektion.

*Beweis.* Es ist  $q(\lambda v) = b(\lambda v, \lambda v) = \lambda^2 b(v, v) = \lambda^2 q(v)$ , und  $q(v+w) - q(v) - q(w) = 2b(v, w)$ , eine symmetrische Bilinearform.

Ist  $\text{char}(k) \neq 2$ , so gilt

$$b(v, w) = \frac{1}{2} (q(v+w) - q(v) - q(w)) \quad (\text{„Polarisierung“})$$

für die quadratische Form  $q$  definiert durch  $q(v) = b(v, v)$ . Genauso gilt  $q(v) = \frac{1}{2} \beta(v, v)$  für jede quadratische Form  $q$ , wobei  $\beta$  die durch  $q$  definierte Bilinearform ist. ■

*Beispiel* Zwei verschiedene symmetrische Bilinearformen auf  $(\mathbb{F}_2)^2$  sind

$$q_1(x_1, x_2) = x_1x_2 \qquad q_2(x_1, x_2) = x_1^2 + x_1x_2 + x_2^2.$$

In beiden Fällen ist die zugehörige Bilinearform  $\beta((x_1, x_2), (y_1, y_2)) = x_1y_2 + x_2y_1$ .

*Bemerkung* Man kann also den Rang und die Signatur einer reellen quadratischen Form definieren, als die entsprechenden Invarianten der assoziierten Bilinearform.

*Wie man eine quadratische Form diagonalisiert* Nach einer geeigneten Basiswahl wird die quadratische Form  $q(x_1, x_2) = x_1^2 + x_1x_2$  auf  $\mathbb{R}^2$  zu  $q(y_1, y_2) = y_1^2 - y_2^2$ , und zwar für  $y_1 = x_1 + \frac{1}{2}x_2$ ,  $y_2 = \frac{1}{2}x_2$ . Der Rang beträgt also 2, die Signatur ist 0.

Für  $\text{char}(k) \neq 2$  kann man jede quadratische Form diagonalisieren, d.h. in einer Form bringen, wo nur Quadrate  $y_i^2$  und keine gemischten Terme  $y_iy_j$  vorkommen. Vorgehensweise:

Input: Eine quadratische Form  $q$  auf  $k^n$ .

- 1) Man bestimme die Matrix der Bilinearform  $b \in \text{Sym}^2(V)$  gegeben durch  $b(v, v) = q(v)$ , d.h.  $2b(v, w) = q(v+w) - q(v) - q(w)$ .

- 2) Man finde eine Orthogonalbasis  $C: c_1, \dots, c_n$  für  $b$ .
- 3) Für  $v = \sum_{i=1}^n y_i c_i$  ist dann  $q(v) = \sum_{i=1}^n b(c_i, c_i) y_i^2$ . Anders gesagt:  $q = \sum_{i=1}^n b(c_i, c_i) c_i^{*2}$ .
- 4) Ist  $k = \mathbb{R}$ , so kann man wahlweise die Basis so ändern, dass die Matrix von  $b$  die Diagonalblockgestalt  $\text{diag}(E_\ell, -E_m, 0)$  annimmt.
- 5) Man drückt jedes  $c_i^*$  als eine Linearkombination der Standarddualbasis  $e_1^*, \dots, e_n^*$  aus. Ist  $S \in M_n(k)$  die Matrix, deren  $i$ ten Spalte  $c_i \in k^n$  ist, so ist die  $i$ te Zeile von  $S^{-1}$  der Koordinatenvektor von  $c_i^*$  bezüglich der Basis  $e_1^*, \dots, e_n^*$ , d.h.  $c_i^* = \sum_{j=1}^n (S^{-1})_{ij} e_j^*$ .
- 6) Jetzt kann man das Ergebnis hinschreiben.

Begründung für Schritt 5: Ist  $v = \sum_{i=1}^n y_i c_i = \sum_{j=1}^n x_j e_j$ , so ist  $c_i^*(v) = y_i$ ,  $e_j^*(v) = x_j$ . Die Matrix  $S$  erfüllt  $\underline{x} = S \cdot \underline{y}$ , also  $\underline{y} = S^{-1} \cdot \underline{x}$  und deshalb  $y_i = \sum_{j=1}^n (S^{-1})_{ij} x_j$ .

*Beispiel* Wir diagonalisieren die quadratische Form  $q$  auf  $\mathbb{R}^3$ , gegeben durch:

$$q(x_1, x_2, x_3) = x_1^2 + 2x_2x_3 - x_1x_3.$$

- 1) Die Matrix von  $b$  ist  $\begin{pmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 \\ -\frac{1}{2} & 1 & 0 \end{pmatrix}$ .
- 2) Eine Orthogonalbasis ist  $c_1 = e_1$ ,  $c_2 = e_1 + 2e_3$ ,  $c_3 = 2e_1 + e_2 + 4e_3$ . Bezüglich dieser Basis ist  $\text{diag}(1, -1, 4)$  die Matrix der Bilinearform. Wir erkennen, dass  $q$  vom Rang 3 und Signatur 1 ist.
- 3) Es ist also  $q(y_1c_1 + y_2c_2 + y_3c_3) = y_1^2 - y_2^2 + 4y_3^2$ .
- 4) Mit  $d_1 = c_1 = e_1$ ,  $d_2 = \frac{1}{2}c_3 = e_1 + \frac{1}{2}e_2 + 2e_3$  und  $d_3 = c_2 = e_1 + 2e_3$  erhalten wir eine weitere Orthogonalbasis. Die Matrix ist jetzt  $\text{diag}(1, 1, -1)$ . Es ist  $q(y_1d_1 + y_2d_2 + y_3d_3) = y_1^2 + y_2^2 - y_3^2$ .
- 5) Für die Basis  $d_1, d_2, d_3$  sind die Matrizen  $S, S^{-1}$  gegeben durch

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

Also

$$d_1^* = e_1^* - \frac{1}{2}e_3^* \quad d_2^* = 2e_2^* \quad d_3^* = \frac{1}{2}e_3^* - 2e_2^*.$$



6) Es ist also

$$q(x_1, x_2, x_3) = (x_1 - \frac{1}{2}x_3)^2 + 4x_2^2 - (\frac{1}{2}x_3 - 2x_2)^2.$$

Eine Kontrolle ist ratsam; sie wird hier bestanden.

## Alternierende Formen

**Satz 17.13** Sei  $b \in \text{Alt}^2(V)$  eine alternierende Bilinearform auf einem  $n$ -dimensionalen Vektorraum  $V$ . Dann gibt es eine ganze Zahl  $0 \leq r \leq \frac{n}{2}$  und eine Basis von  $V$  derart, dass die Matrix von  $b$  die Blockmatrix  $\begin{pmatrix} 0 & E_r & 0 \\ -E_r & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$  ist.

Folglich ist  $r = \frac{1}{2} \text{Rang}(b)$ , und  $b$  kann nur dann nicht ausgeartet sein, wenn  $n = \dim V$  eine gerade Zahl ist.

*Beispiel* Für  $n = 5$  gibt es drei Fälle: Die Nullmatrix ( $r = 0$ ) sowie

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (r = 1) \qquad \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (r = 2).$$

*Beweis.* Letzter Teil: Die Blockmatrix hat Rang  $2r$ . Ist  $b$  nicht ausgeartet, so muss der Rang  $n$  sein.

Hauptaussage: Eine andere Formulierung ist, dass es eine Basis  $x_1, \dots, x_r, y_1, \dots, y_r, z_1, \dots, z_{n-2r}$  gibt derart, dass  $b(x_i, y_i) = 1$ ,  $b(y_i, x_i) = -1$ , und  $b(c, d) = 0$  in allen anderen Fällen mit  $c, d$  Basiselemente. Dies zeigen wir per Induktion über  $n$ . Ist  $b = 0$ , so sind wir fertig mit  $r = 0$ . Induktionsanfang: Da  $b$  alternierend ist, ist  $b(v, v) = 0$  für alle  $v \in V$ , und deshalb  $b = 0$  falls  $n = 0, 1$ .

Induktionsschritt: ist  $b \neq 0$ , so gibt es  $u, v \in V$  mit  $b(u, v) \neq 0$ . Mit  $x_1 = u$ ,  $y_1 = \frac{1}{b(u, v)}v$  ist also  $b(x_1, y_1) = 1$ . Da  $b$  schiefsymmetrisch ist, ist  $b(y_1, x_1) = -1$ . Da  $b$  alternierend ist, sind  $x_1, y_1$  linear unabhängig wegen  $b(x_1, y_1) = 1 \neq 0$ . Sei  $U = \text{Spann}(x_1, y_1)$  und  $W = U^\perp$ , also  $\dim W = n - 2$ . Ist  $w \in U \cap W$ , so ist  $w = \lambda x_1 + \mu y_1$  und  $b(w, x_1) = b(w, y_1) = 0$ , also  $-\mu = \lambda = 0$ . Das heißt,  $U \cap W = \{0\}$  und  $V = U \oplus U^\perp$ . Auch eingeschränkt auf  $U^\perp$  ist  $b$  eine alternierende Form, also nach Induktionsannahme gibt es eine ganze Zahl  $0 \leq r' \leq \frac{n-2}{2}$  und für  $r = r' + 1$  eine Basis  $x_2, \dots, x_r, y_2, \dots, y_r, z_1, \dots, z_{n-2r}$  von  $W$ , die die erwünschte Eigenschaften hat. Somit ist  $x_1, \dots, x_r, y_1, \dots, y_r, z_1, \dots, z_{n-2r}$  die erwünschte Basis von  $V$ . ■

*Beispiel* Die Matrix  $A = \begin{pmatrix} 0 & 2 & -1 & 3 \\ -2 & 0 & 4 & 1 \\ 1 & -4 & 0 & -2 \\ -3 & -1 & 2 & 0 \end{pmatrix}$  ist schiefsymmetrisch, stellt

also eine alternierende Bilinearform  $b \in \text{Alt}^2(\mathbb{R}^4)$  dar. Es ist  $b(e_3, e_1) = 1$ , also wählen wir  $x_1 = e_3, y_1 = e_1$  und suchen eine Basis für  $U^\perp$ , wobei  $U =$

Spann( $e_1, e_3$ ) ist. Nun, die 1. bzw. die 3. Zeile von  $A$  stellt  $b(e_1, \cdot)$  bzw.  $b(e_3, \cdot)$  dar, also ist  $U^\perp$  der Lösungsraum des Gleichungssystems, das durch die Matrix  $\begin{pmatrix} 0 & 2 & -1 & 3 \\ 1 & -4 & 0 & -2 \end{pmatrix}$  gegeben wird. Somit bilden  $u = 8e_1 + 3e_2 - 2e_4$  und  $v = 4e_1 + e_2 + 2e_3$  eine Basis von  $U^\perp$ . Es ist

$$b(u, v) = (8 \ 3 \ 0 \ -2) \begin{pmatrix} 0 & 2 & -1 & 3 \\ -2 & 0 & 4 & 1 \\ 1 & -4 & 0 & -2 \\ -3 & -1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 4 \\ 1 \\ 2 \\ 0 \end{pmatrix} = (8 \ 3 \ 0 \ -2) \begin{pmatrix} 0 \\ 0 \\ 0 \\ -9 \end{pmatrix},$$

d.h.  $b(u, v) = 18$ . Also mit

$$\begin{array}{ll} x_1 = e_3 & y_1 = e_1 \\ x_2 = u = 8e_1 + 3e_2 - 2e_4 & y_2 = \frac{1}{18}v = \frac{4}{18}e_1 + \frac{1}{18}e_2 + \frac{2}{18}e_3 \end{array}$$

hat  $b$  die Matrix  $\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$  bezüglich der Basis  $x_1, x_2, y_1, y_2$ .

## Hermitesche Formen

Sei  $V$  eine endlich dimensionale  $\mathbb{C}$ -Vektorraum.

*Definition* Eine Abbildung  $b: V \times V \rightarrow \mathbb{C}$  heißt eine *Sesquilinearform* auf  $V$ , falls  $b(x, y)$  linear in  $x$  und antilinear in  $y$  ist, d.h.

$$\begin{array}{ll} b(u + v, w) = b(u, w) + b(v, w) & b(u, v + w) = b(u, v) + b(u, w) \\ b(\lambda u, v) = \lambda b(u, v) & b(u, \lambda v) = \bar{\lambda} b(u, v) \end{array}$$

gelten für alle  $u, v, w \in V$  und alle  $\lambda \in \mathbb{C}$ . Gilt außerdem  $b(v, u) = \overline{b(u, v)}$ , so heißt die Sesquilinearform  $b$  eine *hermitesche* Form auf  $V$ .

*Beispiel* So ist etwa  $b((z_1, z_2), (w_1, w_2)) = 3z_1\bar{w}_1 + (1 - i)z_1\bar{w}_2 + (1 + i)z_2\bar{w}_1 - z_2\bar{w}_2$  eine hermitesche Form auf  $\mathbb{C}^2$ .

*Sesquilinearformen und Matrizen* Ist  $v_1, \dots, v_n$  eine Basis von  $V$ , so kann man einer Sesquilinearform  $b$  eine Matrix  $A \in M_n(\mathbb{C})$  zuordnen. Für Bilinearformen setzen wir  $A_{ij} = b(v_i, v_j)$ . Für Sesquilinearformen ist es aber etwas besser, wenn man  $A_{ij} = b(v_j, v_i)$  setzt. Zusammenfassend gilt für eine Sesquilinearform  $b$ :

$$A_{ij} = b(v_j, v_i) \quad v = \sum_{i=1}^n \lambda_i v_i \quad w = \sum_{i=1}^n \mu_i v_i \quad b(v, w) = \underline{w}^H \cdot A \cdot \underline{v}, \quad (17.14)$$

wobei man  $C^H \in M(n \times m, \mathbb{C})$  definiert für  $C \in M(m \times n, \mathbb{C})$  durch  $(C^H)_{ij} = \overline{C_{ji}}$ . Die Sesquilinearform  $b$  ist genau dann hermitesch, wenn  $A^H = A$  gilt für die Matrix  $A$  von  $b$ .

Ändert man die Basis von  $V$ , so ersetzt man  $A$  durch  $S^H \cdot AS$ , wobei die Basiswechselmatrix  $S \in M_n(\mathbb{C})$  invertierbar ist. Umgekehrt stellen zwei Matrizen die gleiche Sesquilinearform dar, wenn  $A' = S^H AS$  ist mit  $S$  invertierbar.

**Lemma 17.15** *Für eine hermitesche Form  $b$  auf  $V$  sei  $q: V \rightarrow \mathbb{R}$  die Abbildung  $q(v) = b(v, v)$ . Dann*

a) *Für jedes  $v \in V$  ist  $b(v, v) \in \mathbb{R}$ , d.h.  $q$  nimmt tatsächlich Werte in  $\mathbb{R}$ . Außerdem ist  $q(\lambda v) = |\lambda|^2 q(v)$  für alle  $\lambda \in \mathbb{C}$ .*

b) *Es ist*

$$q(v+w) - q(v) - q(w) = 2\Re b(v, w) \quad q(v+iw) - q(v) - q(iw) = 2\Im b(v, w).$$

c) *Also gilt die Polarisierungsformel*

$$b(v, w) = \frac{1}{2} (b(v+w, v+w) - b(v, v) - b(w, w)) \\ + \frac{i}{2} (b(v+iw, v+iw) - b(v, v) - b(iw, iw)).$$

*Folglich gilt: ist  $b(v, v) = 0$  für alle  $v$ , dann  $b(v, w) = 0$  für alle  $v, w$ .*

*Beweis.* a) Es ist  $b(v, v) = \overline{b(v, v)}$  und  $b(\lambda v, \lambda v) = \lambda \bar{\lambda} b(v, v)$ .

b) Es ist

$$b(v+w, v+w) - b(v, v) - b(w, w) = b(v, w) + b(w, v) \\ = b(v, w) + \overline{b(v, w)} = 2\Re b(v, w).$$

Der zweite Teil folgt, denn  $\Re b(v, iw) = \Im b(v, w)$ .

c) Die Polarisierungsformel folgt aus b). ■

*Hermitesche Formen haben Orthogonalbasen* Wegen der Bedingung  $b(v, u) = \overline{b(u, v)}$  gilt  $b(u, v) = 0$  genau dann wenn  $b(v, u) = 0$  gilt, also auch hier ist die Relation  $u \perp v$  symmetrisch. Zusammen mit dem Lemma zeigt diese Überlegung, dass der Orthogonalisierungssatz für symmetrische Bilinearformen (Satz 17.11) mitsamt seiner konstruktiven Beweismethode auch für hermitesche Formen gilt.

*Beispiel* Wir konstruieren eine Orthogonalbasis für die hermitesche Form  $b((z_1, z_2), (w_1, w_2)) = 3z_1\bar{w}_1 + (1-i)z_1\bar{w}_2 + (1+i)z_2\bar{w}_1 - z_2\bar{w}_2$  auf  $\mathbb{C}^2$ . Es ist  $b(e_1, e_1) = 3$ , also nehmen wir  $v_1 = e_1$ . Es ist  $b(e_1, e_2) = (1-i)$ , also ist  $b(e_1, v_2) = 0$  für  $v_2 = (1+i)e_1 - 3e_2$ . Es ist

$$\begin{aligned} b(v_2, v_2) &= (1+i)\overline{(1+i)}b(e_1, e_1) - 3(1+i)b(e_1, e_2) - 3\overline{(1+i)}b(e_2, e_1) + 9b(e_2, e_2) \\ &= 2 \cdot 3 - 3(1+i)(1-i) - 3(1-i)(1+i) - 9 = 6 - 6 - 6 - 9 = -15. \end{aligned}$$

Bezüglich der Basis  $e_1, (1+i)e_1 - 3e_2$  hat  $b$  also die Matrix  $\begin{pmatrix} 3 & 0 \\ 0 & -15 \end{pmatrix}$ .

*Hermitesche Formen und der Trägheitssatz* Wegen Lemma 17.15 kann man auch für eine hermitesche Form eine Orthogonalbasis  $C: c_1, \dots, c_n$  in drei Teilmengen  $\text{Pos}(C)$ ,  $\text{Neg}(C)$  und  $\text{Null}(C)$  aufteilen, mit  $c_i \in \text{Pos}(C)$  genau dann, wenn  $b(c_i, c_i) > 0$ , usw. Der Beweis des Trägheitssatzes funktioniert auch für hermitesche Formen, also haben auch solche Formen eine Signatur.

Im obigen Beispiel ist der Rang 2 und die Signatur 0.

## Das Hurwitz-Kriterium

Sei  $b$  eine reell symmetrische Bilinearform auf  $V$ , bzw. eine hermitesche Form auf  $V$ . Dann ist  $b$  genau dann ein Skalarprodukt (vgl. LAAG1), wenn es *positiv definit* ist, d.h.  $b(v, v) > 0$  für alle  $0 \neq v \in V$ . Dies setzt insbesondere voraus, dass  $b$  nicht ausgeartet ist.

**Satz 17.16** *Sei  $A$  die Matrix von  $b$ , eine symmetrische Bilinearform auf  $\mathbb{R}^n$  bzw. eine hermitesche Form auf  $\mathbb{C}^n$ . Genau dann ist  $b$  positiv definit, wenn  $\det A(r)$  eine reelle Zahl  $> 0$  ist für alle  $1 \leq r \leq n$ , wobei  $A(r)$  die  $(r \times r)$ -Matrix links oben in  $A$  ist, d.h.  $A(r)_{ij} = A_{ij}$  für alle  $1 \leq i, j \leq r$ .*

*Beweis.* Wir behandeln nur den hermiteschen Fall, der reell symmetrischen Fall ist ähnlich.

Da  $A$  die Matrix einer hermiteschen Form ist, ist  $A^T = \bar{A}$  und deshalb  $\det(A) = \det(A^T) = \overline{\det(A)}$ , d.h.  $\det(A)$  ist reell. Ist  $A'$  die Matrix von  $b$  bezüglich einer anderen Bilinearform, so ist  $A' = S^H A S$  für eine invertierbare Matrix  $S$ , also ist  $\det A' = \det(S^H) \det(A) \det(S) = |\det S|^2 \cdot \det A$  mit  $\det(S) \neq 0$ , d.h. für alle Matrizen von  $b$  hat  $\det(A) \in \mathbb{R}$  das gleiche Vorzeichen.

Angenommen  $b$  ist positiv definit. Es gibt eine Orthogonalbasis  $v_1, \dots, v_n$ , und es ist  $b(v_i, v_i) > 0$  für alle  $i$  ( $b$  ist positiv definit). Die Matrix  $A'$  von  $b$  bezüglich dieser Basis ist diagonal, und die Diagonaleinträge sind reelle Zahlen  $> 0$ . Also  $\det(A') > 0$  und deshalb  $\det(A) > 0$ . Nun, für jedes  $1 \leq r \leq n$  ist  $b$  auch positiv definit als hermitesche Form auf  $\text{Spann}(e_1, \dots, e_r)$ . Also  $\det A(r) > 0$ , denn  $A(r)$  ist die Matrix von  $b$  auf  $\text{Spann}(e_1, \dots, e_r)$ .

Umgekehrt nehmen wir an, dass  $\det A(r) > 0$  für alle  $1 \leq r \leq n$ . Wir zeigen per Induktion über  $n$ , dass  $b$  positiv definit ist. Nach Induktionsannahme ist

$b$  positiv definit als eine hermitesche Form auf  $U = \text{Spann}(e_1, \dots, e_{n-1})$ , denn  $\det A(r) > 0$  für alle  $r \leq n-1$ . Ist also  $u_1, \dots, u_{n-1}$  eine Orthogonalbasis von  $U$ , so ist  $b(u_i, u_i) > 0$  für alle  $1 \leq i \leq n-1$ . Nach der Beweismethode des Orthogonalisierungssatzes (17.11) können wir diese Basis von  $U$  zu einer Orthogonalbasis  $u_1, \dots, u_n$  von  $V$  fortsetzen. Sei  $A'$  die Matrix von  $b$  bezüglich dieser Basis, dann ist  $A$  diagonal und  $A'_{ii} = b(u_i, u_i)$ . Wegen  $\det(A) > 0$  ist auch  $\det(A') > 0$ . Da  $b(u_i, u_i) > 0$  ist für alle  $i \leq n-1$ , muss auch  $b(u_n, u_n) > 0$  sein. Da es eine Orthogonalbasis gibt derart, dass  $b(u_i, u_i) > 0$  ist für alle  $i$ , folgt es, dass  $b$  positiv definit ist. ■

*Bemerkung* Beispiele sollten aus der Analysis bekannt sein.

## Der adjungierte Endomorphismus

**Lemma 17.17** *Sei  $b$  eine nicht ausgeartete hermitesche Form auf  $V$ . Dann gibt es zu jedem Endomorphismus  $F$  von  $V$  genau einen Endomorphismus  $F^*$  mit der Eigenschaft, dass*

$$b(F(v), w) = b(v, F^*(w))$$

*gilt für alle  $v, w \in V$ . Man nennt  $F^*$  den zu  $F$  adjungierten Endomorphismus.*

*Beweis.* Wie für eine Bilinearform gibt es eine Bijektion  $b_r: V \rightarrow V^*$ ,  $b_r: w \rightarrow b(\cdot, w)$ , nur diesmal ist  $b_r$  antilinear, d.h.  $b_r(\lambda w) = \bar{\lambda}b_r(w)$ . Für festes  $w$  ist  $v \mapsto b(F(v), w)$  eine Linearform auf  $V$ , also gibt es genau ein Element  $F^*(w) \in V$  mit  $b(F(v), w) = b(v, F^*(w))$  für alle  $v$ . Wegen  $b(F(u), \lambda v + \mu w) = b(u, F^*(\lambda v + \mu w))$  einerseits und

$$\begin{aligned} b(F(u), \lambda v + \mu w) &= \bar{\lambda}b(F(u), v) + \bar{\mu}b(F(u), w) \\ &= \bar{\lambda}b(u, F^*(v)) + \bar{\mu}b(u, F^*(w)) = b(u, \lambda F^*(v) + \mu F^*(w)) \end{aligned}$$

andererseits, gilt  $F^*(\lambda v + \mu w) = \lambda F^*(v) + \mu F^*(w)$  aufgrund der Eindeutigkeit von  $F^*(w)$ . Also ist  $F^*$  linear. ■

*Bemerkung* Genauso gut kann man die adjungierte Abbildung einer nicht ausgearteten Bilinearform konstruieren. In der Quantenmechanik etwa spielen Endomorphismen mit  $F^* = F$  eine wichtige Rolle, vgl. die *selbstadjungierten* Endomorphismen aus der LAAG1.

## 18 Affine Unterräume 2

Affine Räume lernten wir im Kapitel 9 der LAAG1 kennen. Jede Gerade in  $\mathbb{R}^2$  ist ein affiner Unterraum, aber nur die Geraden durch den Ursprung sind Untervektorräume.

*Definition* Sei  $V$  ein  $k$ -Vektorraum. Sei  $v_0, \dots, v_m$  Elemente von  $V$ . Ein Vektor  $v \in V$  heißt eine *affine Kombination* des Systems  $v_0, \dots, v_m$ , falls es Skalare  $\lambda_0, \dots, \lambda_m \in k$  gibt mit  $v = \sum_{i=0}^m \lambda_i v_i$  und  $\sum_{i=0}^m \lambda_i = 1$ .

Die Ergebnisse von Kapitel 9 fassen wir jetzt zusammen.

**Satz 18.1** a) Für eine Teilmenge  $A \subseteq V$  eines  $k$ -Vektorraums  $V$  sind die folgenden Aussagen äquivalent:

- Es ist  $A \neq \emptyset$ , und jede affine Kombination von Elementen aus  $A$  liegt ebenfalls in  $A$ .
- Es gibt einen Untervektorraum  $U \subseteq V$ , derart dass  $A$  die Restklasse  $A = a + U$  ist für ein (und deshalb für alle)  $a \in A$ .

Ein solches  $A$  nennt man einen affinen Unterraum von  $V$ . Man definiert  $\dim A$  durch  $\dim A := \dim U$ . Es ist  $U = \{b - a \mid a, b \in A\}$ , also ist  $U$  eindeutig durch  $A$  definiert.

b) Seien  $A = a + U \subseteq V$  und  $B = b + T \subseteq W$  zwei affine Unterräume. Für eine Abbildung  $f: A \rightarrow B$  sind folgende Aussagen äquivalent:

- $f$  verträgt sich mit affinen Kombinationen: sind  $a_0, \dots, a_m \in A$  und  $\lambda_0, \dots, \lambda_m \in k$  mit  $\sum_{i=0}^m \lambda_i = 1$ , so ist  $f(\sum_{i=0}^m \lambda_i a_i) = \sum_{i=0}^m \lambda_i f(a_i)$ .
- Es gibt eine lineare Abbildung  $g: U \rightarrow T$  derart, dass  $f(a + u) = f(a) + g(u)$  gilt für alle  $u \in U$ .

Eine solche Abbildung nennt man eine *affine Abbildung*. Die Abbildung  $g$  ist eindeutig durch  $f$  bestimmt, und es gilt  $f(a + u) = f(a) + g(u)$  für alle  $a \in A$ .

*Beweis.* Teil a) folgt aus den Lemmas 9.1 und 9.2, sowie die Anmerkung nach dem Beweis von Lemma 9.2. Teil b) ist Lemma 9.4. ■

*Beispiel* Sei  $A \subseteq \mathbb{R}^2$  die Gerade  $x = -1$  und  $B \subseteq \mathbb{R}^2$  die Gerade  $x + y = 1$ . Sei  $f: A \rightarrow B$  die Abbildung  $f(-1, y) = (3 - y, y - 2)$ . Dann  $A = (-1, 0) + \text{Spann}(e_2)$ ,  $B = (1, 0) + \text{Spann}(e_1 - e_2)$ , und  $g(\lambda e_2) = -\lambda(e_1 - e_2)$ .

*Bemerkung* Die Eindeutigkeit von  $g$  folgt daraus, dass

$$g(a'' - a') = f(a'') - f(a')$$

gilt für alle  $a', a'' \in A$ .

**Lemma 18.2** Sei  $f: A \rightarrow B$  eine affine Abbildung, mit  $\dim(A), \dim(B) < \infty$ . Dann sind folgende drei Aussagen äquivalent:

- a)  $f$  ist bijektiv;
- b) die assoziierte lineare Abbildung  $g: U \rightarrow T$  ist ein Isomorphismus;
- c) es gibt eine affine Abbildung  $h: B \rightarrow A$  mit  $h \circ f = \text{Id}_A$  und  $f \circ h = \text{Id}_B$ .

Eine solche Abbildung  $f$  nennt man eine Affinität von  $A$  nach  $B$ .

*Beweis.* a) folgt unmittelbar aus c). Sei  $A = a + U$ ,  $B = b + T$ , wobei wir  $b = f(a)$  wählen, es ist also  $f(a + u) = b + g(u)$ . Ist  $g(u) = g(u')$ , so ist  $f(a + u) = f(a + u')$ . Ist  $b + t \in \text{Bild}(f)$ , so gibt es ein  $u$  mit  $f(a + u) = b + t$ , d.h.  $g(u) = t$ . Folgerung: ist  $f$  bijektiv, dann auch  $g$ . Aus a) folgt also b). Gilt b), dann auch c) mit  $h(b + t) = a + g^{-1}(t)$ . ■

*Beispiel* Das obige Beispiel war eine Affinität von der Gerade  $x = -1$  zur Gerade  $x + y = 1$ .

## Affine Unabhängigkeit

*Definition* Sei  $V$  ein  $k$ -Vektorraum. Sei  $v_0, \dots, v_m$  Elemente von  $V$ . Ein System  $v_0, \dots, v_m$  von Elementen aus  $V$  heißt *affin unabhängig*, wenn gilt: sind  $\lambda_0, \dots, \lambda_m \in k$  Skalare mit  $\sum_{i=0}^m \lambda_i v_i = 0$  in  $V$  und  $\sum_{i=0}^m \lambda_i = 0$  in  $k$ , so gilt  $\lambda_i = 0$  für alle  $i$ .

**Lemma 18.3** Für ein System  $v_0, \dots, v_m$  von Elementen eines Vektorraums  $V$  sind folgende Aussagen äquivalent:

- a) Das System ist affin unabhängig.
- b) Die Vektoren  $v_1 - v_0, v_2 - v_0, \dots, v_m - v_0$  sind linear unabhängig.

Ist  $V = k^n$ , so kommt eine dritte äquivalente Bedingung hinzu:

- c) Ist  $v_i = (v_{i1}, \dots, v_{in}) \in k^n$  für  $0 \leq i \leq m$ , so sind die Vektoren  $v_0^+, \dots, v_m^+$  linear unabhängig in  $k^{n+1}$ , wobei  $v_i^+ = (1, v_{i1}, v_{i2}, \dots, v_{in})$ .

*Beweis.* Ist  $\sum_{i=1}^m \lambda_i(v_i - v_0) = 0$ , so ist  $\sum_{i=0}^m \lambda_i v_i = 0$  für  $\lambda_0 = -\sum_{i=1}^m \lambda_i$ . Somit ist  $\sum_{i=0}^m \lambda_i = 0$ . Aus a) folgt also b). Ist dagegen  $\sum_{i=0}^m \lambda_i v_i = 0$  und  $\sum_{i=0}^m \lambda_i = 0$ , so ist  $\lambda_0 = -\sum_{i=1}^m \lambda_i$  und deswegen

$$0 = \sum_{i=1}^m \lambda_i v_i - \sum_{i=1}^m \lambda_i v_0 = \sum_{i=1}^m \lambda_i (v_i - v_0).$$

Gilt b), so ist  $\lambda_i = 0$  für alle  $i \geq 1$  und deswegen  $\lambda_0 = 0$  auch.

Ist  $V = k^n$ , so ist c) eine Umformulierung von den beiden Bedingungen in a). ■

*Beispiel* Um festzustellen, dass die vier Vektoren  $v_0 = (1, 1, 1)$ ,  $v_1 = (1, 2, 3)$ ,  $v_2 = (1, 0, 1)$  und  $v_3 = (3, 1, 1)$  affin unabhängig in  $\mathbb{R}^3$  sind, reicht es nachzuweisen, dass  $(0, 1, 2)$ ,  $(0, -1, 0)$  und  $(2, 0, 0)$  linear unabhängig in  $\mathbb{R}^3$  sind. Man könnte stattdessen nachweisen, dass  $(1, 1, 1, 1)$ ,  $(1, 1, 2, 3)$ ,  $(1, 1, 0, 1)$  und  $(1, 3, 1, 1)$  linear unabhängig in  $\mathbb{R}^4$  sind.

*Beispiel* Drei Punkte sind affin unabhängig, wenn sie nicht kollinear sind sondern eine Ebene aufspannen.

*Definition* Ist  $\dim V = n$ , so können höchstens  $n + 1$  Elemente affin unabhängig sein. Ein System  $P_0, \dots, P_N$  von Elementen aus  $V$  heißt *in allgemeiner Lage*, wenn jedes Teilsystem mit höchstens  $n + 1$  Elementen affin unabhängig ist.

**Satz 18.4** In  $\mathbb{R}^n$  kann man für beliebig großes  $N$  ein System  $P_0, \dots, P_N$  von Punkten in allgemeiner Lage finden.

*Bemerkung* Für  $n = 2$  ist dies die Aussage, dass man beliebig viele Punkte in der Ebene wählen kann derart, dass keine drei Punkte kollinear sind.

Für den Beweis, benötigen wir die Formel für die so genannte *Vandermonde-Determinante*.

**Lemma 18.5** Skalare  $x_1, \dots, x_n \in k$  definieren eine Vandermonde-Matrix in  $M_n(k)$  mit  $r, s$ -Eintrag  $x_s^{r-1}$ . Es ist

$$\det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ x_1^3 & x_2^3 & \cdots & x_n^3 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{pmatrix} = \prod_{1 \leq i < j \leq n} (x_j - x_i).$$

Für  $n = 3$  ist also zum Beispiel

$$\det \begin{pmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ x_1^2 & x_2^2 & x_3^2 \end{pmatrix} = (x_2 - x_1)(x_3 - x_1)(x_3 - x_2).$$



*Beweis.* Induktion über  $n$ . Induktionsanfang  $n = 1$ :  $\det(1) = 1$ . Sei also  $n \geq 2$ . Zieht man für  $1 \leq r \leq n-1$  das  $x_1$ -fache von der  $r$ ten Zeile von der  $(r+1)$ ten Zeile ab, so beträgt diese Zeile  $(0 \quad (x_2 - x_1)x_2^{r-1} \quad (x_3 - x_1)x_3^{r-1} \quad \cdots \quad (x_n - x_1)x_n^{r-1})$ , die Determinante ändert sich dabei nicht. Zieht man also das  $x_1$ -fache der  $(n-1)$ ten Zeile von der  $n$ ten Zeile ab, dann das  $x_1$ -fache der  $(n-2)$ ten Zeile von der  $(n-1)$ ten Zeile, und so weiter bis man das  $x_1$ -fache der ersten Zeile von der zweiten Zeile abzieht, so erhält man

$$\det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ x_1^3 & x_2^3 & \cdots & x_n^3 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{pmatrix} = \det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & x_2 - x_1 & \cdots & x_n - x_1 \\ 0 & (x_2 - x_1)x_2 & \cdots & (x_n - x_1)x_n \\ 0 & (x_2 - x_1)x_2^2 & \cdots & (x_n - x_1)x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & (x_2 - x_1)x_2^{n-2} & \cdots & (x_n - x_1)x_n^{n-2} \end{pmatrix}.$$

Laplace-Entwicklung nach der ersten Spalte:

$$= \det \begin{pmatrix} x_2 - x_1 & x_3 - x_1 & \cdots & x_n - x_1 \\ (x_2 - x_1)x_2 & (x_3 - x_1)x_3 & \cdots & (x_n - x_1)x_n \\ (x_2 - x_1)x_2^2 & (x_3 - x_1)x_3^2 & \cdots & (x_n - x_1)x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ (x_2 - x_1)x_2^{n-2} & (x_3 - x_1)x_3^{n-2} & \cdots & (x_n - x_1)x_n^{n-2} \end{pmatrix}.$$

Faktor  $(x_s - x_1)$  aus der  $(s-1)$ ten Spalte ausziehen ( $2 \leq s \leq n$ ):

$$= (x_2 - x_1)(x_3 - x_1) \cdots (x_n - x_1) \det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_2 & x_3 & \cdots & x_n \\ x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_2^{n-2} & x_3^{n-2} & \cdots & x_n^{n-2} \end{pmatrix}.$$

Dank der Induktionsannahme sind wir jetzt fertig. ■

*Beweis des Satzes.* Es reicht, den Fall  $N \geq n$  zu behandeln. Wählen wir paarweise verschiedene Skalare  $\lambda_0, \dots, \lambda_N \in \mathbb{R}$ , zum Beispiel  $\lambda_r = r$ . Für  $0 \leq r \leq N$  sei  $P_r = (\lambda_r, \lambda_r^2, \dots, \lambda_r^n) \in \mathbb{R}^n$  und  $Q_r = (1, \lambda_r, \lambda_r^2, \dots, \lambda_r^n) \in \mathbb{R}^{n+1}$ . Nach Lemma 18.5 ist jede  $(n+1)$ -elementige Teilmenge der  $Q_r$  linear unabhängig, denn die Vandermonde-Determinante verschwindet nicht. Nach Lemma 18.3 ist also jede  $(n+1)$ -elementige Teilmenge der  $P_r$  affin unabhängig. ■

*Anwendung: geschlossene kombinatorische Flächen* Die Oberfläche eines Kugels ist eine geschlossene Fläche: zweidimensional, kompakt und ohne Rand. Die Oberfläche eines Oktaeders ist eine geschlossene kombinatorische Fläche. Unter

anderem bedeutet dies, dass sie sich aus endlich vielen Dreiecken zusammensetzt, dass an jede Kante sich zwei Dreiecke treffen, und dass der Schnitt von zwei Dreiecke entweder eine gemeinsame Kante oder eine gemeinsame Ecke ist. In der Computergraphik werden kombinatorische Flächen manchmal benutzt, um Flächen darzustellen.

Das Möbiusband verfügt über nur einen Rand. Klebt man zwei Möbiusbänder Rand am Rand zusammen, so erhält man die sogenannte Kleinsche Flasche. Im dreidimensionalen Raum ist bei dieser Konstruktion eine Selbstdurchdringung unvermeidbar.

Dagegen lässt sich jede geschlossene kombinatorische Fläche ohne Selbstdurchdringungen im  $\mathbb{R}^5$  konstruieren<sup>1</sup>. Gegeben sei also eine Konfiguration von endlich vielen Dreiecken mit der folgenden Einschränkung:

Der Schnitt zweier Dreiecke ist entweder leer oder eine gemeinsame Ecke oder eine (ganze) gemeinsame Kante<sup>2</sup>.

Das heißt, uns liege eine Bauanleitung vor, wie wir  $N$  Dreiecke zusammenkleben sollen, Kante an Kante. Wir können die Kantenlängen der Dreiecke uns aussuchen, dürfen aber die Dreiecke nicht biegen. Wir wollen zeigen, dass jedes solche Bauanleitung sich im  $\mathbb{R}^5$  umsetzen lässt.

Zuerst studiert man die Bauanleitung, um festzustellen, wie viele Ecken das zusammengebautes Modell haben sollte. Für diese Ecken wählt man dann Punkte des  $\mathbb{R}^5$  in allgemeiner Lage. Betrachten wir ein Dreieck mit Eckpunkten  $P, Q, R$ . Jeder Punkt des Dreiecks ist eine affine Kombination von  $P, Q, R$ : es ist sogar

$$\text{Dreieck}(P, Q, R) = \{\lambda P + \mu Q + \nu R \mid \lambda + \mu + \nu = 1; \lambda, \mu, \nu \geq 0\}.$$

Schneiden sich die Dreiecke  $(P, Q, R)$  und  $(S, T, U)$ , so ist jeder Schnittpunkt eine affine Kombination von  $P, Q, R$  und von  $S, T, U$ . Da die Ecken aus einer Menge in allgemeiner Lage in  $\mathbb{R}^5$  stammen, ist jede affine Kombination der Menge  $\{P, Q, R, S, T, U\}$  auf genau einer Weise eine solche Kombination. Die Lösung dieses Gegensatzes ist, dass die beiden Dreiecke mindestens eine Ecke teilen. Jeder Punkt im Schnitt ist dann eine affine Kombination der gemeinsamen Ecken. Sind alle Ecken gemeinsame Ecken, so sind die beiden Dreiecke gleich. Sind zwei Ecken gemeinsam, so ist die Schnittmenge die Kante zwischen diesen beiden Ecken. Ist nur eine Ecke gemeinsam, so besteht der Schnitt lediglich aus dieser einen Ecke.

---

<sup>1</sup>Und sogar auch im  $\mathbb{R}^4$ , was wir allerdings noch nicht zeigen können.

<sup>2</sup>Diese Bedingung lässt jede geschlossene kombinatorische Fläche zu, und auch einiges mehr.

## 19 Affine Quadriken

In diesem Kapitel ist  $k$  stets ein Körper der Charakteristik  $\neq 2$ .

*Beispiel* Die Kurven  $y = x^2$  und  $xy = 1$  im  $\mathbb{R}^2$  sowie die Fläche  $x^2 + y^2 = z^2$  im  $\mathbb{R}^3$  sind Beispielen von Quadriken.

*Definition* Sei  $q$  eine quadratische Form auf  $k^n$ , sei  $\phi \in (k^n)^*$  eine Linearform auf  $k^n$ , und sei  $c \in k$  ein Skalar. Ist  $q \neq 0$ , so heißt die Lösungsmenge

$$Q := \{v \in k^n \mid q(v) + \phi(v) + c = 0\}$$

eine (affine) *Quadrik* in  $k^n$ .

*Beispiel* Die Parabel  $x^2 - y = 0$  ist eine Quadrik: es ist  $q(x, y) = x^2$ ,  $\phi(x, y) = -y$  und  $c = 0$ . Auch die Hyperbel  $xy - 1 = 0$  ist eine Quadrik mit  $q(x, y) = xy$ ,  $\phi(x, y) = 0$  und  $c = -1$ .

**Lemma 19.1** *Ist  $Q \subseteq k^n$  eine Quadrik und  $f: k^n \rightarrow k^n$  eine Affinität, so ist auch das Bild  $f(Q)$  eine Quadrik.*

Um dieses Lemma zu zeigen, führen wir eine Bezeichnung ein, die es uns erlaubt, affine Abbildungen und Quadriken jeweils durch eine Matrix darzustellen.

*Bezeichnung* Sei  $v = (v_1, \dots, v_n)$  ein Vektor aus  $k^n$ . Wir werden mit  $v'$  den Vektor  $v' = (1, v_1, \dots, v_n)$  des  $k^{n+1}$  bezeichnen. Also  $v' = (v'_0, \dots, v'_n)$  mit  $v'_0 = 1$  und  $v'_r = v_r$  für  $r \geq 1$ . Auch bei Matrizen, die auf  $v'$  operieren, werden wir die Einträge ab 0 nummerieren.

**Lemma 19.2** *a) Zu jeder affinen Abbildung  $f: k^n \rightarrow k^n$  gibt es eine Matrix  $C \in M_{n+1}(k)$  mit  $C_{00} = 1$  und  $C_{0r} = 0$  für alle  $r \geq 1$  derart, dass  $f(v)' = C \cdot v'$  gilt für jedes  $v \in k^n$ . Zusatz: Ist  $f(v) = b + S \cdot v$ , so ist  $C = \begin{pmatrix} 1 & 0 \\ b & S \end{pmatrix}$ .*

*b) Zu jeder Quadrik  $Q \subseteq k^n$  gibt es eine symmetrische Matrix  $A \in M_{n+1}(k)$  derart, dass*

$$Q = \{v \in k^n \mid v'^T \cdot A \cdot v' = 0\}$$

*gilt. Umgekehrt gilt: ist  $A = \begin{pmatrix} A_{00} & A_{01} & \cdots & A_{0n} \\ A_{10} & & & \\ \vdots & & B & \\ A_{n0} & & & \end{pmatrix}$  eine symmetrische*

*Matrix derart, dass  $B \neq 0 \in M_n(k)$  ist, so ist  $Q = \{v \in k^n \mid v'^T \cdot A \cdot v' = 0\}$  eine Quadrik in  $k^n$ .*

*Beweis.* a) Der Zusatz ist der Beweis.

b) Sei  $B \neq 0$  die Matrix der quadratischen Form  $q$ . Sei  $a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$  der Vektor gegeben durch  $\phi(v) = \sum_{i=1}^n 2a_i v_i$ . Dann mit  $A = \begin{pmatrix} c & a^T \\ a & B \end{pmatrix}$  gilt  $Q = \{v \in k^n \mid v'^T \cdot A \cdot v' = 0\}$ . Für die Umkehrung definiert man  $q$  durch  $B$  usw., statt umgekehrt. ■

*Beweis von Lemma 19.1.* Die Matrix  $C = \begin{pmatrix} 1 & 0 \\ b & S \end{pmatrix}$  einer affinen Abbildung ist genau dann invertierbar, wenn  $S$  invertierbar ist, d.h. wenn es sich um eine Affinität handelt. Sei nun  $C = \begin{pmatrix} 1 & 0 \\ b & S \end{pmatrix}$  die Matrix der Affinität  $f^{-1}$ , und  $A = \begin{pmatrix} c & a^T \\ a & B \end{pmatrix}$  die definierende Matrix für  $Q$ . Es ist  $v \in f(Q)$  genau dann, wenn  $f^{-1}(v) \in Q$  ist. Das heißt, wenn  $w'^T A w' = 0$  gilt für  $w' = C v'$ . Das heißt, wenn  $v'^T C^T A C v' = 0$  gilt. Es reicht also zu zeigen, dass die Matrix  $C^T A C$  eine Quadrik definiert. Diese Matrix ist symmetrisch, und es ist

$$C^T A C = \begin{pmatrix} 1 & b^T \\ 0 & S^T \end{pmatrix} \begin{pmatrix} c & a^T \\ a & B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ b & S \end{pmatrix} = \begin{pmatrix} ? & ? \\ ? & S^T B S \end{pmatrix}.$$

Da  $S$  invertierbar und  $B \neq 0$  ist, ist auch  $S^T B S \neq 0$ . Somit definiert  $C^T A C$  eine Quadrik. ■

## Bewegungen

*Definition* Im Fall eines euklidischen Raums<sup>3</sup>  $V$  nennt man eine Affinität  $f: V \rightarrow V$  eine *Bewegung* von  $V$ , wenn der assoziierte Endomorphismus  $g: V \rightarrow V$  orthogonal ist.

Für  $\mathbb{R}^n$  mit dem Standardskalarprodukt bedeutet dies, dass die Bewegungen genau die Affinitäten der Art  $f(v) = b + S \cdot v$  mit  $S \in O(n)$  sind, d.h.  $S^T \cdot S = S \cdot S^T = E_n$ .

*Bemerkung* Die Bewegungen sind genau die Affinitäten, die *abstandstreu* sind, d.h. die alle Abstände zwischen Punkte erhalten. Man überlegt, dass die Bewegungen eine Gruppe bilden.

*Beispiel* Eine Bewegung des  $\mathbb{R}^2$  ist  $f(x, y) = (1 + \frac{1}{\sqrt{2}}x - \frac{1}{\sqrt{2}}y, \frac{1}{\sqrt{2}}x + \frac{1}{\sqrt{2}}y - 2)$ , mit  $b = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$  und  $S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ .

**Lemma 19.3** „Bewegungen und Streckungen ergeben alle Affinitäten“

Sei  $A \in M_n(\mathbb{R})$  eine invertierbare Matrix. Dann gibt es eine Diagonalmatrix  $D \in M_n(\mathbb{R})$  und zwei orthogonale Matrizen  $S, T \in O(n)$  derart, dass  $A = S D T$  gilt.

<sup>3</sup>Das heißt ein (endlich dimensionaler) reeller Vektorraum mit Skalarprodukt, vgl. LAAG1 Kapitel 8.

*Beweis.* Die Matrix  $A^T \cdot A$  ist reell symmetrisch, daher gibt es nach der Hauptachsentransformation (Satz 8.7) eine orthogonale Matrix  $B \in O(n)$  derart, dass  $B^T A^T A B$  eine Diagonalmatrix ist:

$$B^T A^T A B = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Nun ist  $A^T A$  positiv definit, denn für jedes  $0 \neq v \in \mathbb{R}^n$  ist  $Av \neq 0$  und deshalb  $v^T A^T A v = \langle Av, Av \rangle > 0$ . Somit ist  $\lambda_i > 0$  für jedes  $i$ . Sei also  $D = \text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_n})$ . Dann  $B^T A^T A B = D^2$ , und  $D$  ist symmetrisch und invertierbar. Also  $(ABD^{-1})^T (ABD^{-1}) = D^{-1} B^T A^T A B D^{-1} = E_n$ , weshalb  $S := ABD^{-1}$  orthogonal ist. Mit  $T = B^{-1}$  ist  $A = SDT$ . ■

## Hauptachsentransformation für affine Quadriken

*Beispiel* Wechselt man zum Koordinatensystem  $x', y'$  gegeben durch  $x' = \frac{x+y}{2}$ ,  $y' = \frac{x-y}{2}$ , so wird die Hyperbel  $xy = 1$  jetzt durch die Gleichung  $x'^2 - y'^2 = 1$  definiert. Dies ist ein Beispiel der sog. Hauptachsentransformation für affine Quadriken.

**Satz 19.4** a) Sei  $Q = \{v \in k^n \mid q(v) + \phi(v) + c = 0\}$  eine Quadrik in  $k^n$ . Dann gibt es eine Affinität  $f: k^n \rightarrow k^n$ , derart, dass die definierende Gleichung  $q(f^{-1}(v)) + \phi(f^{-1}(v)) + c = 0$  der Quadrik  $f(Q)$  eine der folgenden drei Standardformen annimmt:

- $\lambda_1 x_1^2 + \lambda_2 x_2^2 + \dots + \lambda_m x_m^2 = 0$ ;
- $\lambda_1 x_1^2 + \lambda_2 x_2^2 + \dots + \lambda_m x_m^2 + 2x_{m+1} = 0$ ;
- $\lambda_1 x_1^2 + \lambda_2 x_2^2 + \dots + \lambda_m x_m^2 + \gamma = 0$ .

In allen Fällen ist  $m \geq 1$  und  $\lambda_1, \dots, \lambda_m \neq 0$ . Im letzten Fall ist  $\gamma \neq 0$ . Im zweiten Fall ist  $m \leq n - 1$ , sonst ist  $m \leq n$ .

- b) Welche Standardform angenommen wird, wird durch die definierende Gleichung von  $Q$  eindeutig bestimmt, ebenso den Wert von  $m$ .
- c) Im Fall  $k = \mathbb{R}$  kann man verlangen, dass die Affinität eine Bewegung ist. Allerdings nimmt dann die zweite Standardform die Gestalt  $\lambda_1 x_1^2 + \lambda_2 x_2^2 + \dots + \lambda_m x_m^2 + 2\gamma x_{m+1} = 0$  an, mit  $\gamma \neq 0$ .
- d) Lässt man dagegen im Fall  $k = \mathbb{R}$  jede Affinität zu, so kann man sicherstellen, dass  $\lambda_i = \pm 1$  ist.

*Bemerkung* Es wird nicht behauptet, dass die definierende Gleichung eindeutig durch die Quadrik definiert wird. Multipliziert man die Gleichung mit einer invertierbaren Skalar, so bleibt die Quadrik offensichtlich gleich. Im  $\mathbb{R}^3$  definieren die Gleichungen  $x_1^2 + 3 = 0$  und  $x_1^2 + x_2^2 + 1 = 0$  die gleiche, leere Quadrik.

*Beweis.* a), c) Sei  $A = \begin{pmatrix} c & a^T \\ a & B \end{pmatrix} \in M_{n+1}(k)$  die Matrix, die die definierende Gleichung von  $Q$  darstellt. Die Matrix  $B \in M_n(k)$  ist symmetrisch. Nach dem Orthogonalisierungssatz 17.11 für symmetrische Bilinearformen gibt es eine invertierbare Matrix  $S \in M_n(k)$  derart, dass die Matrix  $D = S^T B S$  Diagonalgestalt hat. Nach der Hauptachsentransformation (Satz 8.7) kann man im euklidischen Fall eine solche Matrix  $S \in SO(n)$  finden. Wie im Beweis von Lemma 19.1 folgt es, dass die Affinität bzw. Bewegung  $f_1$  mit Matrix  $\begin{pmatrix} 1 & 0 \\ 0 & S \end{pmatrix}^{-1}$  die Matrix  $A$  in  $\begin{pmatrix} 1 & 0 \\ 0 & S \end{pmatrix}^T \cdot A \cdot \begin{pmatrix} 1 & 0 \\ 0 & S \end{pmatrix} = \begin{pmatrix} ? & ? \\ ? & D \end{pmatrix}$  ändert. Das heißt, die definierende Gleichung von  $f_1(Q)$  ist der Gestalt

$$\sum_{i=1}^m \lambda_i x_i^2 + \sum_{j=1}^n 2a_j x_j + c = 0$$

für  $1 \leq m \leq n$  und  $\lambda_i \neq 0$  für alle  $i$ . Indem man jetzt für  $1 \leq i \leq m$

$$x_i^{\text{neu}} = x_i^{\text{alt}} - \frac{a_i}{\lambda_i}$$

setzt – was eine Affinität bzw. Bewegung der Art  $f_2(v) = b + v = b + E_n \cdot v$ , eine sogenannte *Verschiebung*, entspricht –, erreicht die definierende Gleichung die Gestalt

$$\sum_{i=1}^m \lambda_i x_i^2 + \sum_{j=m+1}^n 2a_j x_j + c = 0.$$

Ist  $a_j = 0$  für alle  $m+1 \leq j \leq n$ , haben wir die 1. bzw. die 3. Standardform erreicht, jenachdem ob  $c = 0$  oder  $c \neq 0$ . Sind die  $a_j$  nicht alle 0, so kann man durch einen Variablenwechsel erreichen, dass

$$x_{m+1}^{\text{neu}} = \sum_{j=m+1}^n a_j x_j^{\text{alt}} + \frac{c}{2}$$

ist, und erhält somit die 2. Standardform. Im euklidischen Fall muss man die rechte Seite dieser Gleichung durch  $\sqrt{a_{m+1}^2 + \dots + a_n^2}$  teilen und erhält somit die geänderte 2. Standardform, da orthogonale Endomorphismen die Länge eines Vektors nicht ändern.

- b) Die Matrix  $A = \begin{pmatrix} c & a^T \\ a & B \end{pmatrix}$  stellt die definierende Gleichung dar. Affinitäten ändern weder den Rang von  $A$  noch – wie man aus dem Beweis von Lemma 19.2 b) erkennt – den Rang von  $B$ . Bei der ersten Standardform ist  $\text{Rang}(A) = \text{Rang}(B) = m$ . Bei der zweiten ist  $\text{Rang}(B) = m$ ,  $\text{Rang}(A) = m + 2$ ; und bei der dritten ist  $\text{Rang}(B) = m$ ,  $\text{Rang}(A) = m + 1$ . Für  $n = 4$  ist z.B. die Matrix der Standardform  $x_1^2 - 2x_2^2 + 5x_3^2 + 2x_4 = 0$  gegeben

durch

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix};$$

es ist  $\text{Rang}(A) = 5$ ,  $\text{Rang}(B) = 3 = m$ .

d) Man ersetzt  $x_i$  durch  $\frac{1}{\sqrt{|\lambda_i|}}x_i$  für  $i \leq m$ . ■

*Beispiel* Wir betrachten die Quadrik  $Q \subseteq \mathbb{R}^4$  gegeben durch

$$x_1^2 - 2x_2^2 + x_3^2 + x_4^2 + 2x_1x_3 - 2x_1x_4 - 2x_3x_4 + 4x_1 + 2x_2 + 2x_3 + 1 = 0.$$

Es ist hier

$$A = \begin{pmatrix} 1 & 2 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 & -1 \\ 1 & 0 & -2 & 0 & 0 \\ 1 & 1 & 0 & 1 & -1 \\ 0 & -1 & 0 & -1 & 1 \end{pmatrix} = \begin{pmatrix} c & a^T \\ a & B \end{pmatrix}$$

für

$$c = 1 \quad a = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & -2 & 0 & 0 \\ 1 & 0 & 1 & -1 \\ -1 & 0 & -1 & 1 \end{pmatrix}.$$

Zunächst diagonalisieren wir den quadratischen Teil  $x_1^2 - 2x_2^2 + x_3^2 + x_4^2 + 2x_1x_3 - 2x_1x_4 - 2x_3x_4$ . Mit  $y_1 = x_1 + x_3 - x_4$  ist

$$x_1^2 - 2x_2^2 + x_3^2 + x_4^2 + 2x_1x_3 - 2x_1x_4 - 2x_3x_4 = y_1^2 - 2x_2^2.$$

Wegen  $x_1 = y_1 - x_3 + x_4$  lautet die definierende Gleichung also

$$y_1^2 - 2x_2^2 + 4y_1 + 2x_2 - 2x_3 + 4x_4 + 1 = 0.$$

Setzen wir also  $z_1 = y_1 + 2$ ,  $z_2 = x_2 - \frac{1}{2}$ , so ist  $z_1^2 - 2z_2^2 - 2x_3 + 4x_4 - \frac{5}{2} = 0$ .  
Setzen wir jetzt  $z_3 = -x_3 + 2x_4 - \frac{5}{4}$ , so ist  $z_1^2 - 2z_2^2 + 2z_3 = 0$ .

Das heißt, die definierende Gleichung ist  $z_1^2 - 2z_2^2 + 2z_3 = 0$  (2. Standardform)  
für  $z_1 = x_1 + x_3 - x_4 + 2$ ,  $z_2 = x_2 - \frac{1}{2}$ ,  $z_3 = 2x_4 - x_3 - \frac{5}{4}$  und  $z_4 = x_4$ .

*Fortsetzung des Beispiels* Weiterhin sei  $Q \subseteq \mathbb{R}^4$  die Quadrik gegeben durch

$$x_1^2 - 2x_2^2 + x_3^2 + x_4^2 + 2x_1x_3 - 2x_1x_4 - 2x_3x_4 + 4x_1 + 2x_2 + 2x_3 + 1 = 0.$$

Sei  $Q_1 \subseteq \mathbb{R}^4$  die Standardform-Quadrik

$$Q_1 = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_1^2 - 2x_2^2 + 2x_3 = 0\}.$$

Wir haben gesehen, dass nach dem Variablenwechsel

$$z_1 = x_1 + x_3 - x_4 + 2 \quad z_2 = x_2 - \frac{1}{2} \quad z_3 = 2x_4 - x_3 - \frac{5}{4} \quad z_4 = x_4$$

die definierende Gleichung von  $Q$  die Gestalt  $z_1^2 - 2z_2^2 + 2z_3 = 0$  annimmt. Es muss also eine Affinität  $f: \mathbb{R}^4 \rightarrow \mathbb{R}^4$  geben mit  $f(Q) = Q_1$ . Diese Affinität  $f$  ist bereits in dem Variablenwechsel enthalten, es ist

$$\begin{aligned} f(x_1, x_2, x_3, x_4) &= (x_1 + x_3 - x_4 + 2, x_2 - \frac{1}{2}, 2x_4 - x_3 - \frac{5}{4}, x_4) \\ &= \begin{pmatrix} 2 \\ -\frac{1}{2} \\ -\frac{5}{4} \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}. \end{aligned}$$

Kontrolle: Sei  $A$  bzw.  $A_1$  bzw.  $C \in M_5(\mathbb{R})$  die (erweiterte) Matrix von  $Q$  bzw.  $Q_1$  bzw.  $f$ . Es ist also

$$\begin{aligned} A &= \begin{pmatrix} 1 & 2 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 & -1 \\ 1 & 0 & -2 & 0 & 0 \\ 1 & 1 & 0 & 1 & -1 \\ 0 & -1 & 0 & -1 & 1 \end{pmatrix} & A_1 &= \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\ C &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 1 & -1 \\ -\frac{1}{2} & 0 & 1 & 0 & 0 \\ -\frac{5}{4} & 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

Nach dem Beweis von Lemma 19.1 ist  $A_1$  die Matrix von  $f(Q)$ , wenn  $A_1 = (C^{-1})^T A C^{-1}$  gilt, d.h. wenn  $A = C^T A_1 C$  gilt. Es empfiehlt sich daher nachzurechnen, dass  $A = C^T A_1 C$  tatsächlich gilt. Ich habe es nachgerechnet, es stimmt auch.



## 20 Projektive Räume: Eine kurze Einführung

*Dieses Kapitel ist unvollständig, außerdem fehlen die Bilder.*

In der Ebene  $\mathbb{R}^2$  gelten die folgenden Aussagen:

- A) Sind  $P_1 \neq P_2$  zwei Punkte, so gibt es genau eine Gerade  $L$ , die beide Punkte enthält.
- B) Sind  $L_1 \neq L_2$  zwei Geraden, dann *entweder* gibt es genau einen Punkt  $P$ , der auf beiden Geraden liegt, *oder* die Geraden sind parallel.

Es wäre interessant, wenn man die Fallunterscheidung in B) vermeiden könnte. Dann hätte man folgende Aussage, die leider aber für die Ebene  $\mathbb{R}^2$  falsch ist:

- B') Sind  $L_1 \neq L_2$  zwei Geraden, so gibt es genau einen Punkt  $P$ , der auf beiden Geraden liegt.

Die *projektive Ebene*  $\mathbb{RP}^2 = P_2(\mathbb{R}) = P(\mathbb{R}^3)$  entsteht, wenn man der Ebene zusätzliche, „unendlich weite“ Punkte hinzufügt, damit A) und B') gleichzeitig gelten.

**Eine schnurgerade Straße auf flaches Land** Ich stehe mitten auf einer unendlich lange Straße, die schnurgerade über eine flache Ebene läuft. Die beiden Ränder der Straße sind parallele Geraden und treffen sich somit nie. Trotzdem sehe ich, wenn ich Richtung Horizont schaue, zwei Geraden, die sich immer weiter annähern und schließlich sich doch unendlich weit weg zu treffen scheinen. Wenn ich mich umdrehe und in der anderen Richtung den Straßenverlauf anschau, so sehe ich das gleiche.

Konsequent ausgelegt bedeutet B'), dass diese beiden imaginären, unendlich weiten Punkten zu den Punkten gehören, die man der Ebene hinzufügen muss, um die projektive Ebene zu erhalten. Genauer genommen, stellen diese beiden unendlich weiten Punkte den *gleichen* Punkt der projektiven Ebene dar.

*Bemerkung* Um selber den topologischen Typ der projektiven Ebene zu basteln, nimmt man eine Scheibe und klebt für jeden Randpunkt  $P$  die Punkte  $P, P'$  zusammen, wobei  $P'$  der gegenüberliegende Randpunkt ist<sup>4</sup>. Dies ist leider nicht im  $\mathbb{R}^3$  ohne Selbstdurchdringungen machbar, geht aber schon im  $\mathbb{R}^5$  – vgl. Kapitel 18.

---

<sup>4</sup>Bei einer Scheibe mit Radius  $r$  wird also für jeden Winkel  $\theta$  die beiden Punkten  $(r, \theta)$  und  $(r, \theta + \pi)$  miteinander identifiziert (Polarkoordinaten). So werden etwa Nord und Süd miteinander identifiziert, auch werden Ost und West miteinander identifiziert.

**Vorsicht:** Die Randpunkte werden nur *paarweise* miteinander identifiziert; zum Beispiel sind NordSüd und OstWest zwei verschiedene Punkte des  $P_2(\mathbb{R})$ .

**Ein anderes Modell** Wir stellen uns die Ebene als die Ebene  $z = 1$  im  $\mathbb{R}^3$  vor. Für jeden Punkt  $P$  dieser Ebene gibt es genau eine Gerade  $L_P$  im  $\mathbb{R}^3$ , die den Ursprung des  $\mathbb{R}^3$  und den Punkt  $P$  enthält. Außerdem ist  $L_P = L_Q$  nur dann, wenn  $P = Q$  ist. Da der Ursprung in  $L_P$  enthalten ist, ist  $L_P$  ein eindimensionaler Untervektorraum des  $\mathbb{R}^3$ .

Wir haben also die Ebene  $z = 1$  mit einer Teilmenge der Menge aller eindimensionalen Untervektorräume des  $\mathbb{R}^3$  identifiziert: für ein Punkt  $P$  ist  $L_P$  der Unterraum, der  $P$  als eine Basis hat; und ist  $L$  ein Unterraum, so ist  $P_L$  der Schnittpunkt von  $L$  mit der Ebene  $z = 1$ .

Allerdings gibt es diesen Schnittpunkt für manche Unterräume nicht: dies sind die Unterräume, die in der Ebene  $z = 0$  liegen. Diese Unterräume entsprechen Richtungen, die es auf der Ebene  $z = 1$  gibt. Zwei parallele Geraden in der Ebene  $z = 1$  haben die gleiche Richtung. Es liegt nahe, den eindimensionalen Unterraum, der dieser gemeinsamen Richtung entspricht, zum unendlich weiten Schnittpunkt dieser beiden Geraden zu erklären.

*Definition* Sei  $V$  ein  $k$ -Vektorraum. Der *projektive Raum*  $P(V)$  auf  $V$  wird definiert durch

$$P(V) = \{U \subseteq V \mid U \text{ ist ein eindimensionaler Untervektorraum}\}.$$

Insbesondere setzt man  $P_n(k) = P(k^{n+1})$ . Man schreibt  $\dim P(V) = \dim(V) - 1$ , also  $\dim P_n(k) = n$ .

## Homogene Koordinaten auf projektiver Raum

Der projektive Raum  $P(V)$  ist die Menge aller eindimensionalen Unterräume von  $V$ . Jeder solche Unterraum hat eine Basis, die aus einem Vektor  $v \neq 0$  besteht. Umgekehrt ist jedes  $0 \neq v \in V$  Basis eines eindimensionalen Unterrums; und  $v, w \in W$  sind Basen des gleichen Unterrums genau dann, wenn es ein (zwangsweise invertierbares)  $\lambda \in k$  gibt mit  $w = \lambda v$ . Eine solche Äquivalenzklasse  $[v]$  von Vektoren entspricht also ein Element des  $P(V)$ , und auch umgekehrt. Es ist also

$$P(V) = \{[v] \mid 0 \neq v \in V, \text{ und } [v] = [w] \Leftrightarrow \exists \lambda \in k \ w = \lambda v\}.$$

*Bezeichnung* Für  $v = (x_0, x_1, \dots, x_n) \in k^{n+1}$  mit  $x_i$  nicht alle  $= 0$  bezeichnen wir das Element  $[v]$  des  $P_n(k)$  mit  $(x_0 : x_1 : \dots : x_n)$ . Es ist also

$$P_n(k) = \{(x_0 : x_1 : \dots : x_n) \mid x_i \in k, \text{ nicht alle } = 0\},$$

und  $(x_0 : x_1 : \dots : x_n) = (y_0 : y_1 : \dots : y_n)$  genau dann, wenn es ein  $\lambda \in k$  gibt derart, dass  $y_i = \lambda x_i$  gilt für alle  $i$ .

**Vorsicht:**  $(0 : 0 : \dots : 0)$  gibt es nicht!

*Beispiel* In  $P_2(\mathbb{R})$  ist  $(1 : 2 : -1) = (-2 : -4 : 2) \neq (1 : 0 : -1)$ .

*Bemerkung* Die Abbildung  $\mathbb{R}^2 \rightarrow P_2(\mathbb{R})$ ,  $(x, y) \mapsto (1 : x : y)$  ist injektiv, sie bettet die Ebene in der projektiven Ebene ein, und zwar als der sogenannte „affine Teil“  $x_0 \neq 0$ . Das Komplement des Bildes ist die „unendlich weite Gerade“  $x_0 = 0$ . Es ist

$$P_2(\mathbb{R}) = \{(1 : x : y) \mid x, y \in \mathbb{R}^2\} \uplus \{(0 : 1 : x) \mid x \in \mathbb{R}\} \uplus \{(0 : 0 : 1)\}.$$

**Lemma 20.1** In  $P_n(\mathbb{R})$  und in  $P_n(\mathbb{C})$  hat jede Folge eine konvergente Teilfolge.

*Beweis.* Sei  $k = \mathbb{R}$  oder  $k = \mathbb{C}$ . Für  $0 \leq r \leq n$  setzen wir

$$A_r = \{(x_0 : x_1 : \dots : x_n) \in P_2(k) \mid |x_r| = \max(|x_0|, \dots, |x_n|)\}.$$

Dann  $P_n(k) = A_0 \cup A_1 \cup \dots \cup A_n$ . Sei  $B \subseteq k^n$  die Menge  $B = \{(x_1, \dots, x_n) \mid |x_i| \leq 1 \text{ für jedes } i\}$ . Dann ist  $B$  beschränkt und abgeschlossen, also hat jede Folge in  $B$  eine konvergente Teilfolge. Aber jedes  $A_r$  ist eine Kopie von  $B$ : es ist

$$A_r = \{(x_1 : \dots : x_r : 1 : x_{r+1} : \dots : x_n) \mid (x_1, \dots, x_n) \in B\},$$

wobei die 1 an der  $r$ ten Stelle vorkommt (gezählt ab der 0ten Stelle).

Also: jede Folge in  $P_n(k)$  hat mindestens eine Teilfolge, die in einer der  $n + 1$  Mengen  $A_r$  liegt. Aber jede Folge in  $A_r$  hat eine konvergente Teilfolge, denn  $A_r$  ist eine Kopie von  $B$ . ■

## Lineare Unterräume

*Definition* Ist  $W \subseteq V$  ein Untervektorraum, so ist  $P(W)$  eine Teilmenge von  $P(V)$ . Eine solche Teilmenge nennt man einen *linearen* Unterraum des  $P(V)$ . Es ist  $\dim P(W) = \dim W - 1$ . Ist  $W$  eindimensional, so besteht  $P(W)$  aus einem Punkt des  $P(V)$ . Ist  $W$  zweidimensional, so heißt  $P(W)$  eine Gerade in  $P(V)$ . Ist  $\dim(W) = \dim(V) - 1$ , so heißt  $P(V)$  eine Hyperebene in  $P(V)$ .

**Lemma 20.2** a) Sind  $P_1 \neq P_2$  zwei Punkte des  $P(V)$ , so gibt es genau eine Gerade  $L$  in  $P(V)$ , die  $P_1$  und  $P_2$  enthält.

b) Sind  $L_1 \neq L_2$  zwei Geraden der projektiven Ebene  $P_2(k)$ , so besteht deren Schnittmenge aus einem Punkt  $P$ .

*Beweis.* a) Sei  $P_i = [v_i]$  für  $i = 1, 2$ . Wegen  $P_1 \neq P_2$  sind  $v_1, v_2$  linear unabhängig. Der Untervektorraum  $W = \text{Spann}(v_1, v_2)$  ist also zweidimensional, somit ist  $P(W)$  eine Gerade, die  $L_1, L_2$  enthält. Ist  $P(U)$  ein weiterer linearer Unterraum, der  $P_1, P_2$  enthält, so liegen  $v_1, v_2$  und deshalb auch  $W$  in  $U$ . Also entweder  $U = W$  und  $P(U) = P(W)$ , oder  $U \supsetneq W$  und  $P(U)$  ist keine Gerade.

- b) Sei  $L_i = P(W_i)$ ,  $i = 1, 2$ . Diese Untervektorräume  $W_1, W_2$  sind zweidimensional und liegen in  $k^3$ . Wegen  $L_1 \neq L_2$  ist  $W_1 \neq W_2$ . Also ist  $W_1 + W_2 = k^3$ . Nach der Dimensionsformel folgt  $\dim U = 1$  für  $U = W_1 \cap W_2$ . Also ist  $P(U)$  ein Punkt, der auf beiden Geraden liegt. Ist umgekehrt  $[v]$  ein Schnittpunkt, so ist  $v \in W_1$  und  $v \in W_2$ , also  $v \in U = W_1 \cap W_2$ , also ist  $[v]$  der Punkt  $P(U)$ . ■

## Klassifikation der Geraden in der projektiven Ebene

Sei  $L = P(W)$  eine Gerade in der projektiven Ebene ...

## 21 Zwei Anwendungen

*Dieses Kapitel ist unvollständig, außerdem fehlen die Bilder.*

### 21.1 Die Dreifärbungszahl in der Knotentheorie

*Vorüberlegung* In dem Körper  $\mathbb{F}_3$  mit drei Elementen  $0, 1, -1$  ist der Lösungsraum der Gleichung  $x + y + z = 0$  zweidimensional; eine Basis stellen die Lösungen  $(x, y, z) = (1, 1, 1)$  und  $(x, y, z) = (0, 1, -1)$  dar<sup>5</sup>. Der Lösungsraum enthält insgesamt neun Elemente.

Hier ist eine andere Beschreibung des Lösungsraums: es ist  $x + y + z = 0$  genau dann, wenn entweder  $x = y = z$  gilt (3 Lösungen), oder  $x, y, z$  paarweise verschieden sind (6 Lösungen).

Jeder Knoten in einem Stück Schnur lässt sich aufpicken, indem man ein Ende rauszieht. Hält man dagegen beide Enden fest, so lässt sich ein echter Knoten nicht lösen. Das gleiche Effekt erreicht man, indem man beide Enden der Schnur zusammen bindet. Dagegen lässt sich ein Knoten mit Schleife – wie etwa in meinem Schnürsenkel – auch dann lösen, wenn beide Enden zusammengebunden sind.

Wie soll man entscheiden, ob ein Knoten echt verknotet ist? Zwei Knoten (mit Enden zusammengebunden) gelten als äquivalent, wenn man den einen durch zupfen, zerren usw. in den anderen überführen kann. Ein Knoten ist dann echt verknotet, wenn er nicht zu einem unverknoteten Stück Schnur äquivalent ist.

Äquivalenz von Knoten ist eine Äquivalenzrelation. Sind zwei Knoten äquivalent, so findet man meistens durch ausprobieren einen Weg, den einen in den anderen zu überführen. Nichtäquivalenz ist schwieriger nachzuweisen: auch wenn ich es nicht schaffe, eine Umformung zu finden, vielleicht schafft es ein Einstein oder ein Houdini schon. Um wasserdichte Nichtäquivalenz-Beweise führen zu können, benutzt man sogenannte *Invarianten*. Eine der zugänglichsten Knoteninvarianten ist die /Dreifärbungszahl.

Um mit Knoten zu rechnen, bildet man die auf der Ebene ab, vermerkt aber bei jeder Kreuzung, welcher Teil der Schnur oben und welcher Teil unten liegt. So erhält man ein *Knotendiagramm*, die aus Kreuzungen und Bögen besteht: ein Bogen ist der Teil der Schnur zwischen zwei Unterkreuzungen. An jeder Kreuzung treffen sich drei Bögen.

Was passiert, wenn ich durch zupfen, zerren usw. ein Knotendiagramm in ein äquivalentes überführe? K. Reidemeister zeigte, dass jede solche Überführung sich als eine lange Kette von Einzelschritten zerlegen lässt, wobei nur drei Arten von Einzelschritten benutzt werden: die drei *Reidemeister-Züge*.

Eine Auswirkung von Reidemeisters Ergebnis lautet so: ordne ich jedem Knotendiagramm eine Zahl zu, und zwar so, dass die drei Reidemeister-Züge die

---

<sup>5</sup>In  $\mathbb{F}_3$  gilt  $1 + 1 = -1$  und deshalb auch  $1 + 1 + 1 = 0$ .

Zahl unverändert lassen, so ist die Zahl eine Knoten-Invariante, d.h. äquivalente Knotendiagramme haben immer die gleiche Zahl. Folgerung: habe ich zwei Knotendiagramme mit unterschiedlichen Zahlen, so sind die entsprechenden Knoten nicht äquivalent.

**Definition der Dreifärbungszahl** Malt man jeden Bogen eines Knotendiagramms in einer der drei Farben rot, blau oder grün an, so liegt eine *Dreifärbung* des Diagramms vor. Die Dreifärbung ist *zulässig*, wenn an jeder Kreuzung gilt:

Entweder haben alle drei Bögen die gleiche Farbe, oder jede der drei Farben kommt einmal vor.

Die *Dreifärbungszahl* ist die Anzahl der zulässigen Dreifärbungen.

**Als Lineare Algebra aufgefasst** Wir wählen eine Bijektion zwischen  $\mathbb{F}_3$  und der Menge der drei Farben rot, grün, blau. Jeder Bogen stellt ein unbekanntes Element des  $\mathbb{F}_3$  dar, jede Kreuzung stellt eine Gleichung dar: sind  $x, y, z$ , die Unbekannten, die die drei Bögen an der Kreuzung entsprechen, so lautet die Gleichung  $x + y + z = 0$ , vgl. Vorüberlegung oben. Die Menge der zulässigen Dreifärbungen ist der Lösungsraum des Gleichungssystem; dies ist ein  $\mathbb{F}_3$ -Vektorraum, die Dreifärbungszahl ist also eine Potenz von 3.

**Invarianz der Dreifärbungszahl** Der Lineare-Algebra-Zugang hilft, einen kurzen Beweis dafür zu geben, dass die Reidemeister-Züge die Anzahl der Lösungen nicht ändern. Somit ist die Dreifärbungszahl eine Invariante.

*Beispiel* Der Kleeblattknoten: Drei Bögen, drei Kreuzungen, an jeder Kreuzung treffen sich alle drei Bögen. Das Gleichungssystem ist also dreimal  $x + y + z = 0$ , der Lösungsraum hat neun Elemente. Ein unverknotetes Stück Schnur hat einen Bogen und keine Kreuzungen, die Dreifärbungszahl ist also drei. Da die Dreifärbungszahlen unterschiedlich sind, ist der Knoten wirklich verknotet.

*Bemerkung* Die Dreifärbungszahl erkennt nicht, dass der Achterknoten und der Palstek verknotet sind. Hier benötigt man die Etikettierungszahl modulo einer Primzahl  $p$ : jeder Bogen wird mit einem Element aus  $\mathbb{F}_p$  etikettiert; an jeder Kreuzung muss gelten  $2x = y + z$ , wo  $x$  die Etikette des Ober-Bogens ist und  $y, z$  die Etiketten der Unter-Bögen sind. Auch hier ändern die Reidemeister-Züge die Etikettierungszahl nicht. Mit  $p = 5$  bzw.  $p = 13$  sieht man, dass der Achterknoten bzw. der Palstek verknotet ist.

## 21.2 Lineare Codes

Es soll eine Nachricht von einem Sender nach einem Empfänger geschickt werden. Hier stellen sich zwei Probleme:

- Unterwegs könnte die Nachricht von einem Unbefugten abgefangen werden. In der *Kryptographie* sucht man nach Lösungen.
- Unterwegs könnte die Nachricht durch Rauschen zufällig gestört werden. In der *Codierungstheorie* will man sicherstellen, dass die richtige Nachricht trotzdem ankommt.

Hier geht es um die Codierungstheorie. Zwei konkrete Fälle sind:

- Ein Raumsonder (Sender) funkt Bilder des Saturn zur Erde (Empfänger). Kosmische Strahlen, Staub usw. können das Signal stören.
- Ein HiFi-System (Empfänger) versucht, eine Audio-CD (Sender) zu lesen. Die CD ist aber gekratzt und trägt fettige Fingerabdrücke.

Zwei Ziele sind:

- Erkennen, ob die Nachricht gestört wurde.
- Aus einer gestörten Nachricht die eigentliche Nachricht rekonstruieren.

*Beispiel* Die Nachricht besteht aus zwei Bits<sup>6</sup>. Es sollte so übermittelt werden, dass ein korrumpiertes Bit korrigiert werden kann.

Eine Möglichkeit ist, jedes Bit dreimal zu wiederholen: ist die Nachricht 01, so wird 000111 gefunkt. Kommt z.B. 000101 an, so geht der Empfänger davon aus, dass die 0 an der 5ten Stelle ein Fehler ist, und dass die eigentlich Nachricht 01 war. Man muss aber davon ausgehen, dass die Fehler unabhängig voneinander auftreten, und deshalb die Wahrscheinlichkeit von zwei Fehlern deutlich kleiner ist als die Wahrscheinlichkeit von einem: denn 000101 ist auch 000000 mit zwei Fehlern.

Dieser Wiederholungscode ist allerdings ineffizient: um zwei Bits zu übermitteln, werden sechs Bits gefunkt. Es geht aber auch mit fünf Bits:

Nachricht	Codewort
00	00000
01	01101
10	10110
11	11011

Um eine 2-Bit-Nachricht zu übermitteln, wird das entsprechende 5-Bit-Codewort gefunkt. Das ist effizienter. Und da zwei Codewörter sich immer an mindestens drei Stellen unterschiedlicher Einträge haben, kann auch hier ein falsches Bit erkannt und korrigiert werden: kommt etwa 11110 an, und so wurde vermutlich 10110 gefunkt, d.h. die Nachricht ist 10.

---

<sup>6</sup>Ein Bit ist entweder 0 oder 1.

Das 5-Bit-Beispiel ist ein *linearer Code*, da die Menge der Codewörter ein Untervektorraum des  $k^5$  ist für  $k = \mathbb{F}_2$ .

Lineare Codes sind nicht genau so effizient wie allgemeine Codes, werden aber häufig benutzt, da ihre Handhabung sich deutlich einfacher gestaltet.

*Definition* Sei  $k$  ein Körper mit endlich vielen Elementen; sei  $n \geq 1$ ; und sei  $C$  eine Teilmenge des  $k^n$ .

- Für  $v, w \in k^n$  definiert man den *Hamming-Abstand*  $d(v, w)$  durch

$$d(v, w) = |\{i \mid v_i \neq w_i\}|.$$

Es ist z.B.  $d(11010, 10001) = 3$ .

- Der *Minimalabstand*  $d(C)$  von  $C$  ist

$$d(C) = \min\{d(v, w) \mid v \neq w, v, w \in C\}.$$

- Ist  $C \subseteq k^n$  ein  $m$ -dimensionaler Untervektorraum, so heißt  $C$  ein *linearer Code* vom Typ  $[n, m, d]$ , für  $d = d(C)$ .

**Lemma 21.1** Für alle  $u, v, w \in k^n$  gelten

- $d(v, w) \geq 0$  ist eine ganze Zahl;  $d(v, w) = 0$  genau dann, wenn  $v = w$ ;  $d(v, w) = d(w, v)$ .
- Dreiecksungleichung:*  $d(u, w) \leq d(u, v) + d(v, w)$ .
- Translationsinvarianz:*  $d(v + u, w + u) = d(v, w)$ .
- Ist  $C$  ein linearer Code, so ist

$$d(C) = \min\{d(0, v) \mid 0 \neq v \in C\}.$$

*Beweis.* a) gilt sofort. Zu b): ist  $u_i \neq w_i$ , so muss mindestens eins aus  $u_i \neq v_i$ ,  $v_i \neq w_i$  gelten. Zu c):  $d(v, w)$  hängt nur von  $v - w$  ab. Zu d): Es ist  $d(v, w) = d(0, w - v)$ . Sind  $v, w \in C$ , dann  $0, w - v$  auch. ■

Sei  $k$  ein Körper mit  $q$  Elementen und  $C$  ein linearer Code vom Type  $[n, m, d]$ . Sei  $e \geq 0$  mit  $2e + 1 \leq d$ . Dann ist  $q^m$  die Anzahl von Elementen von  $C$ . Also kann man  $C$  benutzen, um  $q^m$  verschiedene Nachrichten als  $n$  Elemente von  $k$  zu funken, und zwar so, dass wenn höchstens  $e$  Elemente von  $k$  falsch ankommen, man die gemeinte Nachricht richtig erkennen kann.

Denn wird  $c$  gefunkt und kommt  $v$  an mit  $d(v, c) \leq e$ , so ist – dank der Dreiecksungleichung und  $2e + 1 \leq d$  – der Abstand  $d(v, c') \geq e + 1$  für jeden weiteren Codewort  $c'$ .



*Beispiel* Das obige Beispiel ist ein  $[5, 2, 3]$ -Code über  $\mathbb{F}_2$ . Es ist  $e = 1$ .

**Lemma 21.2** (Die Hamming-Schranke) *Ist  $C$  ein linearer Code über  $k$  vom Typ  $[n, m, d]$ , dann gilt*

$$q^n \geq q^m \sum_{j=0}^e \binom{n}{j} (q-1)^j.$$

*Hier ist  $q = |k|$  und  $e$  ist die größte ganze Zahl mit  $2e + 1 \leq d$ .*

*Gilt Gleichheit, so heißt  $C$  ein perfekter Code.*

*Beweis.*  $k^n$  hat  $q^n$  Elemente; auf der rechten Seite werden die aufgezählt, deren Abstand zu einem Codewort höchstens  $e$  beträgt; keine werden doppelt gezählt, da der Abstand zwischen zwei Codewörtern mindestens  $2e + 1$  beträgt. ■

Mehr über Codes in Huppert und Willems, „Lineare Algebra“.