Lineare Algebra

Martin Ziegler

Freiburg WS 11/12, SS 12[1] [2]

[1]version15-1-g93e8913, Mon Oct 15 22:27:44 2018 +0200

[2]Ich danke Heike Mildenberger, Enrique Casanovas und Christina Pflanz für eine kritische Durchsicht.

Inhaltsverzeichnis

Kapitel 1

Der n-dimensionale euklidische Raum

1.1 Lineare Gleichungen

Eine *lineare Gleichung* (über $\mathbb{R}$) ist ein Ausdruck der Form

$$\alpha_1 x_1 + \ldots + \alpha_n x_n = \beta,$$

für reelle Zahlen α_i und β. Eine *Lösung* ist ein n-Tupel

$$(\xi_1, \ldots, \xi_n)$$

von reellen Zahlen, das die Gleichung erfüllt.

Ein *lineares Gleichungssystem* G (in n Variablen) ist ein System

$$\begin{array}{ccccccc}
\alpha_{11}x_1 & + & \cdots & + & \alpha_{1n}x_n & = & \beta_1 \\
 & & \vdots & & & & \\
 & & \vdots & & & & \\
\alpha_{m1}x_1 & + & \cdots & + & \alpha_{mn}x_n & = & \beta_m
\end{array}$$

von linearen Gleichungen. In Kurzform

$$\sum_{j=1}^{n} \alpha_{ij} x_j = \beta_i \quad (i = 1, \ldots, m).$$

Die *Lösungsmenge* von G ist

$$\mathrm{L}(G) = \Big\{ (\xi_1, \ldots, \xi_n) \;\Big|\; \xi_1, \ldots \xi_n \in \mathbb{R},\ \sum_{j=1}^{n} \alpha_{ij} \xi_j = \beta_i \quad (i = 1, \ldots, m) \Big\}$$

,

Das System A der $\alpha_{i,j}$ heißt die *Matrix* von G Die Spalte der β_i ist die *rechte Seite* von G. m ist die Zahl der *Zeilen* von A, n die Zahl der *Spalten.* [1]

Ein Gleichungssystem heißt *homogen*, wenn seine rechte Seite Null ist und *quadratisch*, wenn seine Matrix quadratisch ist, also wenn $m = n$.

Ein Gleichungssystem G ist in *Normalform*, wenn es die Gestalt

$$\begin{array}{cccccccccc}
x_1 & & & & + & \alpha_{1,k+1}x_{k+1} & + & \cdots & \alpha_{1,n}x_n & = & \beta_1 \\
& x_2 & & & + & \alpha_{2,k+1}x_{k+1} & + & \cdots & \alpha_{2,n}x_n & = & \beta_2 \\
& & \ddots & & & \vdots & & & \vdots & & \vdots \\
& & & x_k & + & \alpha_{k,k+1}x_{k+1} & + & \cdots & \alpha_{k,n}x_n & = & \beta_k \\
& & & & & & & & 0 & = & \beta_{k+1} \\
& & & & & & & & \vdots & & \vdots \\
& & & & & & & & 0 & = & \beta_m
\end{array}$$

hat. k heißt der *Rang* von G. Beachte, daß $0 \leq k \leq \min(m, n)$. In Normalform zu sein und der Rang sind Eigenschaften der Matrix von G.

G ist genau dann lösbar, wenn $\beta_{k+1} = \ldots = \beta_m = 0$. Um die Lösungsgesamtheit zu bestimmen, gibt man sich $\xi_{k+1}\ , \ldots ,\ \xi_n$ beliebig vor und wählt die $\xi_1, \ldots, \xi_k$ so, daß

$$\xi_i + \sum_{j=k+1}^{n} \alpha_{i,j}\xi_j = \beta_i$$

für $i = 1, \ldots, k$.

Die Lösungsmenge

$$\left\{ \left(\beta_1 - \sum_{j=k+1}^{n} \alpha_{1,j}\xi_j, \ldots, \beta_k - \sum_{j=k+1}^{n} \alpha_{k,j}\xi_j,\ \xi_{k+1}\ , \ldots ,\ \xi_n\right) \middle|\ \xi_{k+1}\ , \ldots ,\ \xi_n \in \mathbb{R} \right\}$$

ist also $(n - k)$-*parametrig*.

Lemma 1.1.1. *Sei A eine Matrix in Normalform mit m Zeilen, n Spalten und Rang k. Dann ist $k = n$ genau dann, wenn alle Gleichungssysteme mit Matrix A höchsten eine Lösung haben. Und es ist $k = m$ genau dann, wenn alle Gleichungssystem mit Matrix A lösbar sind.*

Beweis. Klar. □

Eine *Zeilenoperation* macht aus G ein neues Gleichungssystem durch Multiplikation der i–ten Zeile mit einer Zahl $\lambda \neq 0$ oder durch Addieren des λ–fachen

[1] i ist der *Zeilenindex*, j der *Spaltenindex*[2]. Wenn keine Mißverständnisse entstehen können, lassen wir das Komma weg: α_{ij}.

[2] In einer früheren Version wurde das verwechselt. Ich danke S. Stroppel für den Hinweis.

der i–ten Zeile zur i'–ten Zeile ($i \neq i'$). Wir bezeichnen diese Operation mit Z_i^λ bzw. $\mathrm{Z}_{i,i'}^\lambda$. Zeilenoperationen sind umkehrbar: $\mathrm{Z}_i^{\lambda^{-1}}$ ist die Umkehrung von Z_i^λ und $\mathrm{Z}_{i,i'}^{-\lambda}$ die Umkehrung von $\mathrm{Z}_{i,i'}^\lambda$.

Lemma 1.1.2. *Ein Gleichungssystem G', das aus G durch Zeilenoperationen hervorgeht, hat die gleichen Lösungen wie G.*

Beweis. Für die Operation Z_i^λ sieht man das zum Beispiel so: Sei

$$\alpha_1 x_1 + \ldots + \alpha_n x_n = b$$

die i–te Zeile des Gleichungssystems. $(\xi_1, \ldots, \xi_n)$ ist genau dann eine Lösung dieser Gleichung, wenn $\alpha_1\xi_1 + \ldots + \alpha_n\xi_n = b$. Das ist genau dann der Fall, wenn

$$\lambda(\alpha_1\xi_1 + \ldots + \alpha_n\xi_n) = (\lambda\alpha_1)\xi_1 + \ldots + (\lambda\alpha_n)\xi_n = \lambda b,$$

das heißt, wenn $(\xi_1, \ldots, \xi_n)$ eine Lösung der Gleichung

$$(\lambda\alpha_1)x_1 + \ldots + (\lambda\alpha_n)x_n = \lambda b$$

ist. □

Satz 1.1.3. *Jedes lineare Gleichungssystem läßt sich durch Zeilenoperationen und Vertauschung von Variablen (d.h. von Spalten der Matrix) in Normalform bringen.*

Die Matrix A' der Normalform hängt nur von der Matrix A des Gleichungssystems ab. Wir nennen A' eine Normalform von A.

Beweis: Wir überlegen uns zuerst, daß man durch vier Zeilenoperationen zwei Zeilen vertauschen kann. Denn sei zum Beispiel a die erste Zeile und b die zweite Zeile des Gleichungssystems. Nach Anwendung der Operationen $\mathrm{Z}_{1,2}^1$, $\mathrm{Z}_{2,1}^{-1}$, $\mathrm{Z}_{1,2}^1$, Z_1^{-1} sind a und b vertauscht:

$$\begin{matrix} a \\ b \end{matrix} \longrightarrow \begin{matrix} a \\ a+b \end{matrix} \longrightarrow \begin{matrix} -b \\ a+b \end{matrix} \longrightarrow \begin{matrix} -b \\ a \end{matrix} \longrightarrow \begin{matrix} b \\ a \end{matrix}.$$

Schritt 1:
Wenn alle Koeffizienten der Matrix des Gleichungssystems gleich Null sind, hat das Gleichungssystems bereits Normalform mit Rang $k = 0$.
Wenn es einen nichtverschwindenden Koeffizienten gibt, können wir durch Zeilen– und Spaltenvertauschung erreichen, daß $\alpha_{1,1} \neq 0$. Nach $\mathrm{Z}_1^{\alpha_{1,1}^{-1}}$ ist

$$\alpha_{1,1} = 1.$$

(Man beachte, daß sich die Bedeutung der $\alpha_{i,j}$ geändert hat.) Dann wenden wir $\mathrm{Z}_{1,2}^{-\alpha_{2,1}}$,..., $\mathrm{Z}_{1,m}^{-\alpha_{m,1}}$ an und haben

$$\alpha_{2,1} = \ldots = \alpha_{m,1} = 0.$$

Wir notieren die so erreichte Gleichung durch

$$\left(\begin{array}{cccc|c} 1 & \alpha_{1,2} & \cdots & \alpha_{1,n} & \beta_1 \\ 0 & \alpha_{2,2} & \cdots & \alpha_{2,n} & \beta_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & \alpha_{m,2} & \cdots & \alpha_{m,n} & \beta_m \end{array}\right)$$

Schritt 2:
Wenn $\alpha_{i,j} = 0$ für alle $i, j \geq 2$ sind wir fertig. Das Gleichungssystems hat Normalform mit Rang $k = 1$.
Wenn $\alpha_{i,j} \neq 0$, für zwei Indizes ≥ 2, vertauschen wir die 2–te Zeile mit der i–ten Zeile und die 2–te Spalte mit der j–ten Spalte und erreichen $\alpha_{2,2} \neq 0$. Nach $\mathrm{Z}_2^{\alpha_{2,2}-1}$ ist

$$\alpha_{2,2} = 1.$$

Schließlich wenden wir $\mathrm{Z}_{2,1}^{-\alpha_{1,2}}, \mathrm{Z}_{2,3}^{-\alpha_{3,2}}, \ldots, \mathrm{Z}_{2,m}^{-\alpha_{m,2}}$ an und haben

$$\alpha_{1,2} = \alpha_{3,2} = \ldots = \alpha_{m,2} = 0,$$

In unserer symbolischen Notation

$$\left(\begin{array}{ccccc|c} 1 & 0 & \alpha_{1,3} & \cdots & \alpha_{1,n} & \beta_1 \\ 0 & 1 & \alpha_{2,3} & \cdots & \alpha_{2,n} & \beta_2 \\ 0 & 0 & \alpha_{3,3} & \cdots & \alpha_{3,n} & \beta_3 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \alpha_{m,3} & \cdots & \alpha_{m,n} & \beta_m \end{array}\right)$$

Auf diese Weise verwandeln wir der Reihe nach die Spalten der Matrix des Gleichungssystems in Nullspalten, in denen nur an der Diagonalstelle eine Eins steht. Dieses Verfahren bricht spätestens ab, wenn keine neuen Spalten oder Zeilen mehr zur Verfügung stehen, also nach höchstens $\min(m, n)$ Schritten:

$$\left(\begin{array}{ccccccc|c} 1 & 0 & \cdots & 0 & \alpha_{1,k+1} & \cdots & \alpha_{1,n} & \beta_1 \\ 0 & 1 & \cdots & 0 & \alpha_{2,k+1} & \cdots & \alpha_{2,n} & \beta_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \alpha_{k,k+1} & \cdots & \alpha_{k,n} & \beta_k \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \beta_{k+1} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \beta_m \end{array}\right)$$

□

Folgerung 1.1.4. *Sei A eine Matrix in Normalform mit m Zeilen, n Spalten. k sei der Rang einer Normalform von A. Ein Gleichungssystem mit Matrix A hat entweder keine Lösung oder ein $n-k$–parametriges Lösungssystem. Es ist $k = m$ genau dann, wenn alle Gleichungssysteme mit Matrix A lösbar sind, und $k = n$ genau dann, wenn alle Gleichungssystem mit Matrix A höchstens eine Lösung haben.*

Beweis: Das folgt aus Lemma 1.1.1 und der Umkehrbarkeit der Zeilenoperationen. □

Folgerung 1.1.5. *Ein homogenes Gleichungssystem mit weniger Gleichungen als Variablen hat eine nichttriviale Lösung.*

Beweis: Ein homogenes Gleichungssystem hat immer die triviale Lösung $(0,\ldots,0)$. Weil $k \leq m$, ist die Lösungsmenge mindestens $n-m>0$–parametrig. (Man kann also mindestens den Wert von ξ_n frei wählen.)

Wir geben noch einen alternativen Beweis, der unabhängig von Satz 1.1.3 ist. Wir zeigen die Behauptung durch Induktion über n, die Zahl der Variablen.
Induktionsanfang:
Wir beginnen bei $n=1$. Das Gleichungssystem enthält dann *keine* Gleichung. Jeder Wert für ξ_1 ist eine Lösung. Zum Beispiel $\xi_1 = 1$.

Induktionsschritt:
Nehmen wir an, daß $n \geq 2$ und die Behauptung für Gleichungssysteme mit $n-1$ Variablen gilt. G sei ein Gleichungssystems mit n Variablen und $m < n$ Gleichungen. Wir unterscheiden zwei Fälle:
Wenn alle Gleichungen trivial sind (d.h. alle Koeffizienten sind Null), sind wir fertig. Jedes n–Tupel $(\xi_1,\ldots,\xi_n)$ ist eine Lösung.
Sonst ordnen wir die Gleichungen und Variablen um, bis die erste Gleichung g_1 die Form

$$\alpha_1 x_1 + \ldots + \alpha_n x_n = 0$$

für ein $\alpha_1 \neq 0$ hat. Wir subtrahieren geeignete Vielfache von g_1 von den übrigen Gleichungen $g_2,\ldots,g_m$ bis die Variable x_1 in diesen Gleichungen nicht mehr vorkommt. Jetzt wenden wir die Induktionsvoraussetzung auf $g_2,\ldots,g_n$ an (ein Gleichungssystem mit $n-1$ Variablen und $m-1$ Gleichungen). Wir erhalten eine nicht–triviale Lösung $(\xi_2,\ldots,\xi_n)$, die wir durch

$$\xi_1 = -\frac{\sum_{j=2}^{n} \alpha_j \xi_j}{\alpha_1}$$

zu einer Lösung von G erweitern. □

Daß $(\xi_1,\ldots,\xi_n)$ ein Lösung von G ist, läßt sich als eine lineare Abhängigkeit

$$\xi_1 a_1 + \ldots + \xi_n a_n = b$$

zwischen den *Spalten*

$$a_1 = \begin{pmatrix} \alpha_{1,1} \\ \vdots \\ \alpha_{m,1} \end{pmatrix}, \ldots, a_n = \begin{pmatrix} \alpha_{1,n} \\ \vdots \\ \alpha_{m,n} \end{pmatrix}, b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}$$

ausdrücken. Wenn G in Normalform (von Rang k) ist, bestehen zwischen den $a_1,\ldots,a_k$ keine lineare Abhängigkeiten. Die übrigen Spalten a_i und – wenn das Gleichungssystem lösbar ist – auch b lassen sich eindeutig als Linearkombination der ersten k Spalten ausdrücken:

$$\alpha_{1,i}a_1 + \cdots + \alpha_{k,i}a_k = a_i$$
$$\beta_1 a_1 + \cdots + \beta_k a_k = b$$

Derartige lineare Abhängigkeiten zwischen Spalten eines Gleichungssystems werden offensichtlich von Zeilenoperationen nicht geändert. Daraus folgt

Folgerung 1.1.6. *Wenn man ein Gleichungssystem durch Zeilenoperationen*[3] *in Normalform bringt, erhält man immer die gleiche Matrix. Wenn das Gleichungssystem lösbar ist, ist auch die rechte Seite eindeutig bestimmt.*

Wenn man Spaltenvertauschung zuläßt, bleibt nur der Rang erhalten.

Satz 1.1.7. *Wenn man ein Gleichungssystem durch Zeilenoperationen und Spaltenvertauschungen in Normalform bringt, erhält man immer denselben Rang.*

Definition. *Eine Folge von Spalten gleicher Länge heißt linear unabhängig, wenn sich keine der Spalten als Linearkombination der anderen schreiben läßt.*

Bemerkung. $a_1, \cdots, a_n$ *sind genau dann linear unabhängig, wenn wenn für alle* $\xi_1, \ldots \xi_n \in \mathbb{R}$

$$0 = \xi_1 a_1 + \ldots + \xi_n a_n \quad \Longrightarrow \quad \xi_1 = \cdots = \xi_n = 0.$$

Beweis. Wenn $0 = \xi_1 a_1 + \ldots + \xi_n a_n$ und $\xi_1 \neq 0$, ist $a_1 = \sum_{i=2}^{n} -\frac{\xi_i}{\xi_1} a_i$. Wenn umgekehrt $a_1 = \sum_{i=2}^{n} \alpha_i a_i$, ist $0 = -a_1 + \alpha_2 a_2 + \ldots + \alpha_n a_n$. □

Eine einzelne Spalte a_1 ist linear unabhängig, wenn $a_1 \neq 0$. Die leere Folge $(n = 0)$ is linear unabhängig per Definition.

Beweis von 1.1.7. Wir überlegen uns, daß der Rang eines Gleichungssystems G in Normalform die maximale Zahl linear unabhängiger Spalten der Matrix von G ist. Die ersten k Spalten sind linear unabhängig. Seien $a_{i_1}, \ldots, a_{i_{k+1}}$ beliebige Spalten. Weil die Einträge dieser Spalten ab dem Zeilenindex $k+1$ gleich Null sind, hat das Gleichungssystem

$$x_1 a_{i_1} + \ldots + x_{k+1}, a_{i_{k+1}} = 0^4$$

nur k wirkliche Gleichungen. Nach 1.1.5 gibt es eine nicht–triviale Lösung, was zeigt, daß die $a_{i_1}, \ldots, a_{i_{k+1}}$ linear abhängig sind. □

[3] Also ohne Spalten zu vertauschen. Man kann die Folgerung auch so ausdrücken: Die Normalform hängt nur von den Spaltenvertauschungen ab.

[4] 0 bezeichnet hier die *Nullspalte* (siehe S.12)

1.2 Der $\mathbb{R}^n$

Daß *Produkt* $X_1 \times \ldots \times X_n$ einer Folge von Mengen ist die Menge

$$\{(x_1, \ldots, x_n) \mid x_i \in X_i (i = 1, \ldots, n)\}$$

aller n–Tupel, deren i–te Komponente aus X_i ist. X^n ist das n–fache direkte Produkt von X.

$$\mathbb{R}^n = \{(\xi_1, \ldots, \xi_n) \mid \ \xi_1, \ldots, \xi_n \in \mathbb{R}\}$$

ist der *n-dimensionale euklidische Raum*.

Spezialfälle:	$\mathbb{R}^0$	wird als der *Nullraum* $\mathbf{0} = \{0\}$ vereinbart.
	$\mathbb{R}^1$	sind die reellen Zahlen selbst.
	$\mathbb{R}^2$	ist die euklidische Ebene.

Je nach Zusammenhang heißen die Elemente des $\mathbb{R}^n$ Punkte oder Vektoren.

Vektoren können addiert und mit reellen Zahlen (Skalaren) multipliziert werden. Seien $a = (\xi_i)$ und $b = (\zeta_i)$ Vektoren und $\lambda \in \mathbb{R}$. Dann definiert man

$$\begin{aligned} a + b &= (\xi_i + \zeta_i) \\ \lambda a &= (\lambda \xi_i). \end{aligned}$$

Mit 0 bezeichnen wir den *Nullvektor* $(0, \ldots, 0)$.

Lemma 1.2.1. *Für die Addition gelten die folgenden Rechenregeln.*

a) $(x + y) + z = x + (y + z)$

b) $x + 0 = 0 + x = x$

c) Zu jedem x gibt es ein y mit $x + y = 0$.

d) $x + y = y + x$

und für die Multiplikation

a) $\lambda(x + y) = \lambda x + \lambda y$

b) $(\lambda + \mu)x = \lambda x + \mu x$

c) $\lambda(\mu x) = (\lambda \mu)x$

d) $1x = x$

Beweis. Diese Rechenregeln gelten für reelle Zahlen (den $\mathbb{R}^1$), also auch für $\mathbb{R}^n$. Zum Beispiel ist

$$\begin{aligned}(\lambda+\mu)(\xi_1,\dots,\xi_n) &= ((\lambda+\mu)\xi_1,\dots,(\lambda+\mu)\xi_n)\\ &= (\lambda\xi_1+\mu\xi_1,\dots,\lambda\xi_n+\mu\xi_n)\\ &= (\lambda\xi_1,\dots,\lambda\xi_n)+(\mu\xi_1,\dots,\mu\xi_n)\\ &= \lambda(\xi_1,\dots,\xi_n)+\mu(\xi_1,\dots,\xi_n)\end{aligned}$$

Das beweist b). □

1.3 Geraden und Ebenen

Definition. *Eine (affine) Gerade g in $\mathbb{R}^n$ ist eine Menge der Form*

$$g = \{a+\lambda v \mid \lambda \in \mathbb{R}\} = a + \mathbb{R}v,$$

wobei $v \neq 0$.

Der *Richtungsraum* $\mathbb{R}v$ ist durch g eindeutig bestimmt als die Menge der Differenzen $x-y$ der Elemente von g. Wenn w zum *Richtungsraum* $\mathbb{R}v$ gehört, ist $w+\mathbb{R}v = \mathbb{R}v$ und $g = (a+w)+\mathbb{R}v$ ist eine andere Darstellung von g. Für alle $b \in g$ ist also $g = b + \mathbb{R}v$.

Aus diesen Bemerkungen folgt

Lemma 1.3.1. *Zwei Geraden $a+\mathbb{R}v$ und $b+\mathbb{R}w$ sind genau dann gleich, wenn $\mathbb{R}v = \mathbb{R}w$ und $a-b \in \mathbb{R}v$.* □

Wenn v und w nicht null sind, ist $\mathbb{R}v = \mathbb{R}w$ genau dann, wenn v und w linear abhängig sind.

Lemma 1.3.2. *Durch zwei verschiedene Punkte des $\mathbb{R}^n$ geht genau eine Gerade.*

BEWEIS: Sei g eine Gerade durch a. Dann hat g die Gestalt $a+\mathbb{R}v$ und wenn g auch b enthält, ist $a-b \in \mathbb{R}v$. Daraus folgt $\mathbb{R}v = \mathbb{R}(a-b)$. Das beweist die Eindeutigkeit. □

Definition. *Zwei Geraden heißen* parallel*, wenn sie die gleichen Richtungsräume haben.*

Verschiedene parallele Geraden können sich nicht schneiden.

Definition. *Eine affine Ebene E ist eine Menge der Form $a+\mathbb{R}v+\mathbb{R}w$, für linear unabhängige v und w.*

Der Richtungsraum $\mathbb{R}v + \mathbb{R}w$ ist durch E eindeutig bestimmt als Menge aller Differenzen $a - b$ von Elementen von E.

Lemma 1.3.3. *Zwei nicht parallele Geraden, die in einer Ebene liegen, schneiden sich.*

Beweis: Sei $E = c + \mathbb{R}v_1 + \mathbb{R}v_2$ und $g_1 = a_1 + \mathbb{R}w_1$ und $g_2 = a_2 + \mathbb{R}w_2$ zwei Geraden in E. Wir suchen zwei Zahlen ξ_1 und ξ_2 mit

$$a_1 + \xi_1 w_1 = a_2 + \xi_2 w_2.$$

Wenn wir die Geraden in den Koordinaten von E beschreiben ($i = 1, 2$):

$$\begin{aligned} a_i &= c + \beta_{1i} v_1 + \beta_{2i} v_2 \\ w_i &= \alpha_{1i} v_1 + \alpha_{2i} v_2 \end{aligned}$$

werden wir auf das Gleichungssystem

$$\begin{aligned} \alpha_{11}\xi_1 - \alpha_{12}\xi_2 &= -\beta_{11} + \beta_{12} \\ \alpha_{21}\xi_1 - \alpha_{22}\xi_2 &= -\beta_{21} + \beta_{22} \end{aligned}$$

geführt. Weil g_1 und g_2 nicht parallel sind, sind w_1 und w_2 linear unabhängig und daher auch die Spalten der Matrix $\begin{pmatrix} \alpha_{11} & -\alpha_{12} \\ \alpha_{21} & -\alpha_{22} \end{pmatrix}$. Nach 1.1.7 und 1.1.4 hat Gleichungssystem eine Lösung. □

1.4 Das Skalarprodukt

Das Skalarprodukt zweier Vektoren $a = (\xi_i)$ und $b = (\zeta_i)$ ist

$$ab = \xi_1\zeta_1 + \ldots + \xi_n\zeta_n.$$

Lemma 1.4.1. *Das* Skalarprodukt *ist eine symmetrische, positiv definite Bilinearform. Das heißt*

- $(a + b)c = ac + bc$ *und* $a(b + c) = ab + ac$
- $(\lambda a)b = a(\lambda b) = \lambda(ab)$
- $ab = ba$
- $a^2 \geq 0$
- $a^2 = 0$ *gdw.* $a = 0$.

Die letzte Eigenschaft, die Positivdefinitheit, gilt, weil eine Quadratsumme $\xi_1^2 + \ldots + \xi_n^2$ niemals negativ ist, und nur gleich Null, wenn alle ξ_i Null sind.

Nach dem Satz von Pythagoras in der euklidischen Geometrie ist die Länge der Diagonale eines Quaders die Wurzel aus der Quadratsumme der Seiten. Wir definieren demgemäß:

Definition. *Die Norm oder Länge von a ist*

$$\|a\| = \sqrt{a^2}.$$

Zwei Vektoren a und b heißen *orthogonal*, wenn ihr Skalarprodukt ab gleich Null ist. Allgemeiner definiert man:

Definition. *Der Winkel α zwischen zwei nicht–trivialen Vektoren ist definiert durch $0 \leq \alpha \leq \pi$ und*

$$\cos(\alpha) = \frac{ab}{\|a\|\|b\|}. \tag{1.1}$$

Die Definition ist sinnvoll[5], weil aus dem nächsten Lemma folgt, daß

$$-1 \leq \frac{ab}{\|a\|\|b\|} \leq 1.$$

Lemma 1.4.2 (Cauchy-Schwarzsche Ungleichung)**.**

$$|ab| \leq \|a\|\|b\|$$

Die Gleichheit gilt genau dann, wenn a und b linear abhängig sind.

Beweis: Für beliebige λ ist

$$(a + \lambda b)^2 = a^2 + 2\lambda ab + \lambda^2 b^2 \geq 0. \tag{1.2}$$

Für $\lambda = -\dfrac{ab}{b^2}$ ergibt sich

$$a^2 - 2\frac{(ab)^2}{b^2} + \frac{(ab)^2}{b^2} \geq 0.$$

(Wir können uns auf den Fall $b \neq 0$ beschränken.) Daraus folgt $(ab)^2 \leq a^2b^2$, wie behauptet.

Wenn a und b linear unabhängig sind, ist die linke Seite von 1.2 immer ungleich Null und es folgt $(ab)^2 < a^2b^2$. □

[5] Daß die Definition mit der Anschauung der euklidischen Geometrie übereinstimmt, diskutieren wir auf Seite 16.

Lemma 1.4.3 (Dreiecksungleichung)**.**

$$\|a+b\| \leq \|a\| + \|b\|$$

Beweis. Aus der Cauchy–Schwarzschen Ungleichung folgt

$$\|a+b\|^2 = (a+b)^2 = a^2 + 2ab + b^2 \leq \|a\|^2 + 2\|a\|\|b\| + \|b\|^2 = \big(\|a\| + \|b\|\big)^2.$$

□

Folgerung 1.4.4. *Der $\mathbb{R}^n$ mit dem Abstand*

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

ist ein metrischer Raum. □

Ein metrischer Raum ist eine Menge X mit einer Funktion $d : X \times X \to \mathbb{R}$, für die

- $d(x,y) \geq 0$ für alle $x, y \in X$.
- $d(x,y) = 0$ genau dann, wenn $x = y$.
- $d(x,y) = d(y,x)$
- $d(x,z) \leq d(x,y) + d(y,z)$

1.5 Lineare Abbildungen

Eine Abbildung $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ heißt *linear*, wenn für reelle Zahlen $\alpha_1, \ldots, \alpha_n$

$$f(\xi_1, \ldots, \xi_n) = \alpha_1\xi_1 + \ldots + \alpha_n\xi_n.$$

f heißt auch *Linearform*[6]. Für die *kanonischen* Basisvektoren[7] $\mathrm{e}_i = (\xi_1, \ldots, \xi_n)$, definiert durch

$$\xi_j = \begin{cases} 0 & \text{, wenn } j \neq i \\ 1 & \text{, wenn } j = i, \end{cases}$$

gilt $f(\mathrm{e}_i) = \alpha_i$. Eine Abbildung $f = (f_1, \ldots, f_m)$ von $\mathbb{R}^n$ nach $\mathbb{R}^m$ heißt *linear*, wenn alle Komponenten $f_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ linear sind. $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, die Menge der linearen Abbildungen ist ein $\mathbb{R}$–Vektorraum unter wertweiser Addition und Multiplikation mit Skalaren.

Eine m–n–Matrix

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix}$$

[6] Oder *lineares Funktional*

[7] Die kanonischen Basisvektoren nennt man auch *Einheitsvektoren*.

ist eine mit Zahlenpaaren aus $\{1\dots m\}\times\{1\dots n\}$ indizierte Familie

$$(\alpha_{ij})_{\substack{i=1\dots m\\ j=1\dots n}}.$$[8]

Mit elementweiser Addition und Multiplikation mit Skalaren bildet die Menge $\mathrm{M}_{mn}(\mathbb{R})$ der m-n–Matrizen einen $\mathbb{R}$–Vektorraum. 1-n–Matrizen heißen *Zeilenvektoren*, m-1–Matrizen *Spaltenvektoren*.

Spezielle Matrizen: $\mathbf{0}$ ist die m–n–Matrix, deren Elemente Nullen sind. Die *Einheitsmatrix* $\mathbf{I}$ ist die m–m–Matrix (δ_{ij}), wobei $\delta_{ij} = \begin{cases} 0 & \text{wenn } i \neq j \\ 1 & \text{sonst} \end{cases}$. (Man nennt δ_{ij} "Kroneckers Delta".)

$$\mathbf{I} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$$

Eine n–n–Matrix heißt *quadratisch*. Wir schreiben $\mathrm{M}_n(\mathbb{R})$ für $\mathrm{M}_{nn}(\mathbb{R})$.

Eine m–n–Matrix A definiert eine lineare Abbildung $f_A : \mathbb{R}^n \to \mathbb{R}^m$ durch

$$f_A(\xi_1,\dots,\xi_n) = \big(\sum_{j=1}^n \alpha_{1j}\xi_j,\dots,\sum_{j=1}^n \alpha_{mj}\xi_j\big).$$

$f_{\mathbf{0}}$ ist die Nullabbildung, $f_{\mathbf{I}}$ die identische Abbildung.

Lemma 1.5.1. *Die durch $A \mapsto f_A$ definierte Bijektion* $\mathrm{M}_{mn}(\mathbb{R}) \to \mathrm{L}(\mathbb{R}^n,\mathbb{R}^m)$ *ist ein* Isomorphismus*. Das heißt, daß*

$$\begin{aligned} f_{A+B} &= f_A + f_B \\ f_{\beta A} &= \beta f_A. \end{aligned}$$

Beweis: Sei $A = (\alpha_{ij})$, $B = (\beta_{ij})$ und $x = (\xi_j)$. Dann ist

$$\begin{aligned} f_{A+B}(x) &= \big(\sum_{j=1}^n (\alpha_{ij}+\beta_{ij})\xi_j\big) \\ &= \big(\sum_{j=1}^n \alpha_{ij}\xi_j + \sum_{j=1}^n \beta_{ij}\xi_j\big) \\ &= \big(\sum_{j=1}^n \alpha_{ij}\xi_j\big) + \big(\sum_{j=1}^n \beta_{ij}\xi_j\big) \\ &= f_A(x) + f_B(x) \end{aligned}$$

[8] Kürzere, weniger präzise, Schreibweisen sind $(\alpha_{ij})_{i,j}$ und (α_{ij}).

Damit ist die erste Bedingung gezeigt. Mit einer ähnlichen Rechnung beweist man die zweite Bedingung:

$$\begin{aligned} f_{\beta A}(x) &= \big(\sum_{j=1}^{n}(\beta\alpha_{ij})\xi_j\big) \\ &= \big(\beta\sum_{j=1}^{n}\alpha_{ij}\xi_j\big) \\ &= \beta\big(\sum_{j=1}^{n}\alpha_{ij}\xi_j\big) \\ &= \beta f_A(x) \end{aligned}$$

□

Definition. *Das Produkt $C = (\gamma_{hj})$ einer l–m–Matrix $A = (\alpha_{hi})$ und einer m–n–Matrix $B = (\beta_{ij})$ ist eine l–n–Matrix definiert durch $\gamma_{hj} = \sum_{i=1}^{m}\alpha_{hi}\beta_{ij}$.*

Satz 1.5.2.

$$f_A \circ f_B = f_{AB}$$

Beweis: Sei $A = (\alpha_{hi})$ und $B = (\beta_{ij})$ und $x = (\xi_j)$. Dann ist

$$\begin{aligned} (f_A \circ f_B)(x) &= f_A(f_B(x)) \\ &= f_A\big(\big(\sum_j \beta_{ij}\xi_j\big)_i\big) \\ &= \big(\sum_i \alpha_{hi}\sum_j \beta_{ij}\xi_j\big)_h \\ &= \big(\sum_j\big(\sum_i \alpha_{hi}\beta_{ij}\big)\xi_j\big)_h \\ &= f_{AB}(x) \end{aligned}$$

□

Folgerung 1.5.3. *Das Matrizenprodukt ist assoziativ und bilinear. Es gilt $\mathbf{0}A = A\mathbf{0} = \mathbf{0}$ und $\mathbf{I}A = A\mathbf{I} = A$.*

Beweis: Die Bilinearität, das heißt

$$\begin{aligned} A(B + B') &= AB + AB' \\ (A + A')B &= AB + A'B \\ A(\gamma B) &= (\gamma A)B = \gamma(AB) \end{aligned}$$

(siehe Seite 38) folgt sofort aus der Definition. Zum Beispiel ist

$$
\begin{aligned}
(\alpha_{hi})\big((\beta_{ij}) + (\beta'_{ij})\big) &= (\alpha_{hi})(\beta_{ij} + \beta'_{ij}) \\
&= \big(\sum_i \alpha_{hi}(\beta_{ij} + \beta'_{ij})\big) \\
&= \big(\sum_i \alpha_{hi}\beta_{ij} + \sum_i \alpha_{ij}\beta'_{ij}\big) \\
&= \big(\sum_i \alpha_{hi}\beta_{ij}\big) + \big(\sum_i \alpha_{ij}\beta'_{ij}\big) \\
&= (\alpha_{hi})(\beta_{ij}) + (\alpha_{hi})(\beta'_{ij})
\end{aligned}
$$

Man kann die Assoziativität

$$A(BC) = (AB)C$$

entweder durch eine kurze Rechnung nachweisen, oder die Assoziativität der Komposition von Abbildungen verwenden:

$$
\begin{aligned}
f_{A(BC)} &= f_A \circ \big(f_B \circ f_C\big) \\
&= \big(f_A \circ f_B\big) \circ f_C \\
&= f_{(AB)C}
\end{aligned}
$$

□

Interpretation der Zeilen und Spalten von A Wenn man die Elemente von $\mathbb{R}^n$ bzw. von $\mathbb{R}^m$ als Spaltenvektoren auffaßt, ist f_A die Rechtsmultiplikation mit der Matrix A. Die Gleichung $f_A(\xi_1, \dots, \xi_n) = (\eta_1, \dots, \eta_m)$ ist nämlich gleichbedeutend mit

$$A \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_m \end{pmatrix}.$$

Wenn man den Einheitsvektor e_j als Spalte auffaßt, ist

$$A\mathrm{e}_j = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{mj} \end{pmatrix},$$

die j-te Spalte von A. Die Spalten von A sind also die Bilder der Einheitsvektoren unter f_A.

Aus

$$(\alpha_{i1} \dots \alpha_{in}) \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \eta_i$$

ergibt sich, daß die i–te Zeile von A die i–te Komponente von f_A definiert. Sei e_i^* der i-te Einheitsvektor als Zeile geschrieben. Dann ist $\mathrm{e}_i^* A$ die i–te Zeile von A. Daraus ergibt sich leicht, daß $\mathrm{e}_i^* A \mathrm{e}_j = \alpha_{ij}$.

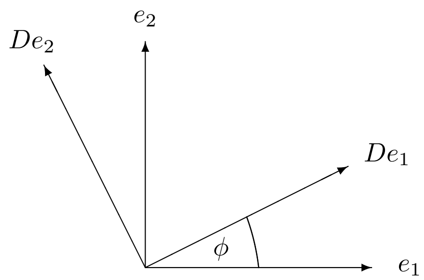
Lemma 1.5.4. *$a_1, \dots, a_l$ seien die Zeilen von A und $b_1, \dots, b_n$ die Spalten von B. Dann sind $a_1 B, \dots, a_l B$ die Zeilen und $Ab_1, \dots, Ab_n$ die Spalten von AB. Außerdem ist*

$$AB = (a_h b_j)_{\substack{h=1\dots l \\ j=1\dots n}}$$

Beweis: Die j–te Spalte von AB ist $AB\mathrm{e}_j = Ab_j$. Die h–te Zeile von AB ist $\mathrm{e}_h^* AB = a_h B$. Wenn $AB = (\gamma_{hj})$, ist $\gamma_{hj} = \mathrm{e}_h^* AB\mathrm{e}_j = a_h b_j$. □

Drehungen des $\mathbb{R}^2$

Wenn wir in der Schulgeometrie den $\mathbb{R}^2$ um den Winkel ϕ drehen, gehen die Basisvektoren $\begin{pmatrix}1\\0\end{pmatrix}$ und $\begin{pmatrix}0\\1\end{pmatrix}$ über in die Vektoren $\begin{pmatrix}\cos(\phi)\\\sin(\phi)\end{pmatrix}$ und $\begin{pmatrix}-\sin(\phi)\\\cos(\phi)\end{pmatrix}$.



Eine Linearkombination

$$\begin{pmatrix}\xi_1\\\xi_2\end{pmatrix} = \xi_1 \begin{pmatrix}1\\0\end{pmatrix} + \xi_2 \begin{pmatrix}0\\1\end{pmatrix}$$

wird dann abgebildet auf

$$\xi_1 \begin{pmatrix}\cos(\phi)\\\sin(\phi)\end{pmatrix} + \xi_2 \begin{pmatrix}-\sin(\phi)\\\cos(\phi)\end{pmatrix}.$$

Es folgt, daß unsere Drehung eine lineare Abbildung f_D ist, mit der Matrix

$$D = \begin{pmatrix}\cos(\phi) & -\sin(\phi)\\\sin(\phi) & \cos(\phi)\end{pmatrix}.$$

Wir wollen uns überlegen, daß die Formel (1.1) auf Seite 10

$$\cos(\alpha) = \frac{ab}{\|a\|\|b\|}$$

im schulgeometrischen Sinn den Winkel zwischen a und b richtig angibt. Das ist leicht zu sehen, wenn a auf der x–Achse liegt. Weil man jeden Vektor durch eine geeignete Drehung auf die x–Achse legen kann, genügt es zu zeigen, daß (1.1) invariant unter Drehungen ist. Das folgt aus der Tatsache, daß das Skalarprodukt invariant unter Drehungen ist. In der folgenden Rechnung verwenden wir (a, b) für das Skalarprodukt von a und b.

Sei $a = \alpha_1 \mathrm{e}_1 + \alpha_2 \mathrm{e}_2$ und $b = \beta_1 \mathrm{e}_1 + \beta_2 \mathrm{e}_2$. Dann ist

$$(Da, Db) = \alpha_1\beta_1(D\mathrm{e}_1, D\mathrm{e}_1) + \alpha_2\beta_2(D\mathrm{e}_2, D\mathrm{e}_2) + (\alpha_1\beta_2 + \alpha_2\beta_1)(D\mathrm{e}_1, D\mathrm{e}_2)$$

Die rechte Seite ist gleich

$$\alpha_1\beta_1 + \alpha_2\beta_2 = (a, b),$$

weil die Spalten $D\mathrm{e}_1 = \begin{pmatrix} \cos(\phi) \\ \sin(\phi) \end{pmatrix}$ und $D\mathrm{e}_2 = \begin{pmatrix} -\sin(\phi) \\ \cos(\phi) \end{pmatrix}$ die Länge 1 haben und aufeinander senkrecht stehen.

Diese Überlegung überträgt sich auch auf den $\mathbb{R}^3$: Man kann jede Drehung als ein Produkt von Drehungen um die Koordinatenachsen darstellen. Diese Drehungen verhalten sich wie Drehungen des $\mathbb{R}^2$ und lassen das Skalarprodukt invariant.

Kapitel 2

Vektorräume

2.1 Gruppen

Definition. *Eine Halbgruppe ist eine Menge S, auf der eine zweistellige assoziative Operation $\circ$ erklärt ist. Eine Halbgruppe ist also ein Paar*

$$(S, \circ\,),$$

wobei die Operation $\circ : S \times S \to S$ dem Assoziativgesetz

$$(x \circ y) \circ z = x \circ (y \circ z)$$

genügt.

Die Assoziativität hat zur Folge, daß es bei längeren Produkten nicht auf die Klammerung ankommt. Zum Beispiel ist

$$(a \circ b) \circ (b \circ c) = ((a \circ b) \circ b) \circ c.$$

Man läßt daher im allgemeinen die Klammern weg und schreibt einfach

$$a \circ b \circ b \circ c.$$

Definition. *Eine Halbgruppe S heißt abelsch, oder kommutativ, wenn*

$$s \circ t = t \circ s$$

für alle $s, t \in S$.

Die Operation einer Halbgruppe nennt man im allgemeinen Multiplikation und schreibt $s \cdot t$. Im abelschen Fall häufig auch Addition, wo man dann $s + t$ statt $s \cdot t$ schreibt.

Definition. *Ein Element e einer Halbgruppe S heißt linksneutral, wenn*

$$e \circ s = s$$

für alle $s \in S$. Wenn

$$s \circ f = s$$

für alle s, heißt f rechtsneutral.

Lemma 2.1.1. *Wenn eine Halbgruppe S ein linksneutrales Element e und ein rechtsneutrales f hat, stimmen e und f überein. Man nennt $e = f$ (das)* neutrale Element *von S.*

Beweis:

$$e = e \circ f = f$$

□

In multiplikativ geschrieben Halbgruppen nennt man das neutrale Element 1 (falls vorhanden) *Einselement*. In additiv geschriebenen (abelschen) Halbgruppen heißt das neutrale Element 0 die Null.

Sei nun S eine Halbgruppe mit Einselement 1. (Man nennt solche Halbgruppen auch *Monoide*.) Wenn

$$s \circ t = 1,$$

nennen wir s *Linksinverses* von t und t *Rechtsinverses* von s.

Lemma 2.1.2. *Sei S eine Halbgruppe mit Einselement 1. Wenn s ein Linksinverses s' und ein Rechtsinverses s''hat, stimmen s' und s'' überein. Man nennt $s' = s''$ das* Inverse *von s.*

Beweis:

$$s' = s' \circ 1 = s' \circ s \circ s'' = 1 \circ s'' = s''$$

□

In multiplikativen Halbgruppen schreibt man s^{-1} für das Inverse von s, in additiv geschriebenen (abelschen) Halbgruppen schreibt man $-s$.

Bemerkung.

$$(a \circ b)^{-1} = b^{-1} \circ a^{-1}$$

Definition. *Eine Gruppe G ist eine Halbgruppe mit neutralem Element, in der jedes Element ein Inverses hat.*

Satz 2.1.3. *Eine Halbgruppe G ist genau dann eine Gruppe, wenn es ein linksneutrales Element e gibt, für das jedes Element $g \in G$ ein Linksinverses hat, das heißt, eine Lösung der Gleichung $x \circ g = e$.*

Beweis: Sei g ein beliebiges Element, g' ein Linksinverses von g und g'' ein Linksinverses von g'. Aus

$$g \circ g' = e \circ g \circ g' = g'' \circ g' \circ g \circ g' = g'' \circ e \circ g' = g'' \circ g' = e$$

liest man ab, daß g' auch Rechtsinverses von g ist. Die Gleichung

$$g \circ e = g \circ g' \circ g = e \circ g = g$$

zeigt, daß e auch rechtsneutral ist. □

Lemma 2.1.4. *Eine Halbgruppe $(G, \circ)$ ist genau dann eine Gruppe, wenn für alle $a, b \in G$ die Gleichungen $x \circ a = b$ und $a \circ y = b$ lösbar sind.*

Beweis. Wenn $(G, \circ)$ eine Gruppe ist, ist $x = b \circ a^{-1}$ eine Lösung von $x \circ a = b$ und $y = a^{-1} \circ b$ eine Lösung von $a \circ y = b$.

Für die Umkehrung genügt es, ein linksneutrales Elemente e zu finden. Die Existenz von Linksinversen folgt dann aus der Lösbarkeit der Gleichung $x \circ a = e$. Dazu fixieren wir eine beliebiges Element a_0 und wählen e so, daß $e \circ a_0 = a_0$. Für gegebenes b sei y eine Lösung von $a_0 \circ y = b$. Dann ist $e \circ b = e \circ a_0 \circ y = a_0 \circ y = b$. □

Definition. *Sei n eine positive natürliche Zahl.*

- $a^0 = e$
- $a^n = \underbrace{a \circ \circ a \circ a}_{n \text{ Faktoren}}$
- $a^{-n} = (a^{-1})^n$

Lemma 2.1.5. *Sei G eine Gruppe. Dann gilt für alle $a \in G$ und $x, y \in \mathbb{Z}$*

a) *Wenn $a \circ b = b \circ a$, ist $(a \circ b)^x = a^x \circ b^x$.*

b) $a^{x+y} = a^x \circ a^y$

c) $a^{xy} = (a^x)^y$

d) $a^1 = a$

Beweis. Man kann sich das Lemma leicht durch Betrachtung einiger Beispiele klar machen. □

Beispiele abelscher Gruppen:

- $(\mathbb{Z}, +)$

- $\mathbb{Z}_n$, die additive Gruppe der Reste

  $$\bar{z} = z + n\mathbb{Z}$$

  ganzer Zahlen modulo n. Man addiert zwei Reste, indem man ihre Repräsentanten addiert.

  $$\bar{y} + \bar{z} = \overline{y + z}.$$

  Die Restklassen modulo n sind die Klassen der Äquivalenzrelation[1]

  $$y \equiv z \pmod{n}.$$

  Man beachte, daß $\mathbb{Z}_n = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$.

- $(\mathbb{R}, +)$

- $(\mathbb{R}^\bullet, \circ)$, die multiplikative Gruppe der von Null verschiedenen reellen Zahlen.

In additiv geschrieben abelschen Gruppen schreibt man za statt a^z. Während man in multiplikativ geschriebenen Gruppen die Gruppenoperation häufig nicht notiert:

$$a \cdot b \cdot b \cdot c = ab^2c,$$

kann man in der additiven Schreibweise $+$ nicht weglassen:

$$a + 2b + c.$$

Das nächste Lemma folgt sofort aus 2.1.5.

Lemma 2.1.6. *Sei $(G, +)$ eine abelsche Gruppe. Dann ist G ein $\mathbb{Z}$–Modul. Das heißt, daß für alle $a, b \in G$ und $x, y \in \mathbb{Z}$*

a) $x(a + b) = xa + xb$

b) $(x + y)a = xa + ya$

c) $(xy)a = x(ya)$

d) $1a = a$. □

Eine Abbildung $f : X \to Y$ heißt *injektiv*, wenn $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$ für alle $x_1, x_2 \in X$. f heißt *surjektiv* , wenn Y das *Bild*

$$f[X] = \{f(x) \mid x \in X\}$$

von f ist. Eine injektive und surjektive Abbildung heißt *bijektiv* oder *Bijektion*. Die *Umkehrabbildung* f^{-1} einer Bijektion f ist definiert durch

$$f^{-1}(y) = x \Leftrightarrow f(x) = y.$$

[1] Zum Begriff einer Äquivalenzrelation vergleiche Seite 36.

Eine Bijektion $f : X \to X$ heißt *Permutation* von X.

Zur Notation: $f \circ g$ ist die durch $(f \circ g)(x) = f(g(x))$ definierte *Verknüpfung* oder *Komposition* von f und g. id_X bezeichnet die *identische* Abbildung von X nach X: Es ist $\mathrm{id}_X(x) = x$ für alle $x \in X$.

Bemerkung. *Sei $f : X \to Y$ eine Abbildung zwischen zwei nicht-leeren Mengen.y*

1. *f ist genau dann injektiv, wenn es ein $f' : Y \to X$ gibt mit $f' \circ f = \mathrm{id}_X$.*
2. *f ist genau dann surjektiv, wenn es ein $f'' : Y \to X$ gibt mit $f \circ f'' = \mathrm{id}_Y$.*
3. *f ist genau dann bijektiv, wenn es ein $f^{-1} : Y \to X$ gibt mit $f^{-1} \circ f = \mathrm{id}_X$ und $f \circ f^{-1} = \mathrm{id}_Y$.*

Die Inverse Abbildung f^{-1} ist eindeutig bestimmt.

Definition. *Mit* $\mathrm{Sym}(X)$ *bezeichnet man die Gruppe der Permutationen von X, die Symmetrische Gruppe.* $\mathrm{S}_n = \mathrm{Sym}(1, \ldots, n)$ *ist die Gruppe der Permutationen von $\{1, \ldots, n\}$.*

S_1 und S_2 sind kommutativ. S_3 und alle weiteren S_n sind nicht kommutativ.

Man schreibt $\begin{pmatrix} 1 & 2 & \cdots & n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix}$ für das Element von S_n, das i die Zahl a_i zuordnet.

Definition. *Sei $(G, \circ)$ eine Gruppe. Eine Untergruppe U von G ist eine Teilmenge von G, die e enthält und unter $\circ$ und $^{-1}$ abgeschlossen ist.*

U ist offensichtlich wieder eine Gruppe mit der eingeschränkten Operation[2]

$$\circ_U = \circ \restriction (U \times U).$$

Bemerkung. *Eine unter $\circ$ abgeschlossene Teilmenge U von G ist mit der eingeschränkten Operation genau dann eine Gruppe, wenn U Untergruppe von G ist.*

Beweis. Wenn $(U, \circ_U)$ eine Gruppe ist, betrachten wir ihr Einselement e_U. Weil $e_U \circ e_U = e_U$, muß e_U gleich e sein. Inverse im Sinn von U sind also auch Inverse im Sinn von G, woraus folgt, daß U unter $^{-1}$ abgeschlossen ist. □

G selbst und $\mathrm{E} = \{e\}$ sind immer Untergruppen. Für jedes Element a ist

$$a^{\mathbb{Z}} = \{a^z \mid z \in \mathbb{Z}\}$$

eine Untergruppe von G. $a^{\mathbb{Z}}$ ist die kleinste Untergruppe, die a enthält.

Definition. *G und H seien Gruppen*[3]*. Eine Abbildung $f : G \to H$ ist ein*

[2] Die *Einschränkung* $f \restriction A$ einer Funktion $f : X \to Y$ auf eine Teilmenge A von X ist die auf A definierte Funktion, die auf A mit f übereinstimmt. Wenn man eine Funktion als eine Menge von Paaren auffaßt, ist also $f \restriction A = f \cap A \times Y$.

[3] Der Begriff ist sinnvoll für beliebige Mengen mit einer zweistelligen Operation.

Homomorphismus, wenn

$$f(ab) = f(a)f(b)$$

für alle $a, b \in G$.

BEISPIELE:

- Die identische Abbildung $\mathrm{id}_G : G \to G$.
- Die triviale Abbildung $g \mapsto e$ von G nach H.
- Für ein festes $a \in H$ die Abbildung $z \mapsto a^z$ von $\mathbb{Z}$ nach H.
- Die Restklassenabbildung $z \mapsto \bar{z}$ von $\mathbb{Z}$ nach $\mathbb{Z}_n$.
- Die Signaturabbildung

$$\mathrm{sign}(x) = \begin{cases} 1 & \text{wenn } x > 0 \\ -1 & \text{wenn } x < 0 \end{cases}$$

von $\mathbb{R}^\bullet$ in die Untergruppe $\{1, -1\}$.

Lemma 2.1.7. *Das Bild eines Homomorphismus $f : G \to H$ ist eine Untergruppe von H.*

Beweis. Das Produkt $z = f(x)f(y)$ von zwei Elementen von $f[G]$ ist wieder in G, weil $z = f(xy)$. $f(e)$ ist linksneutral in $f[G]$, weil $f(e)f(x) = f(ex) = f(x)$. $f(x^{-1})$ ist Linksinverses von $f(x)$, weil $f(x^{-1})f(x) = f(x^{-1}x) = f(e)$. $f[G]$ ist also eine Gruppe. Daraus folgt wie in der Bemerkung auf Seite 21, daß $f[G]$ Untergruppe ist. Insbesondere ist $f(e) = e$ und $f(a^{-1}) = f(a)^{-1}$ □

Ein bijektiver Homomorphismus f zwischen Gruppen heißt *Isomorphismus* . Zwei Gruppen G und H, zwischen denen es einen Isomorphismus gibt, heißen *isomorph.* Man notiert das als

$$f : G \xrightarrow{\cong} H$$

und

$$G \cong H.$$

BEISPIELE:

- Die Untergruppe $\{1, -1\}$ von $\mathbb{R}^\bullet$ ist isomorph zu $\mathbb{Z}_2$.
- Die Exponentialabbildung $\exp : (\mathbb{R}, +) \to (\mathbb{R}_{>0}, \cdot)$ ist ein Isomorphismus[4].

[4] $\mathbb{R}_{>0} = \{r \in \mathbb{R} \mid r > 0\}$

Lemma 2.1.8. *Isomorphie ist eine Äquivalenzrelation* [5] *.*

Beweis. Reflexivität: $\mathrm{id}_G : G \xrightarrow{\cong} G$

Transitivität: Wenn $f : F \xrightarrow{\cong} G$ und $g : G \xrightarrow{\cong} H$, dann ist $g \circ f : F \xrightarrow{\cong} H$.

Symmetrie: Wenn $f : G \xrightarrow{\cong} H$, dann ist $f^{-1} : H \xrightarrow{\cong} G$. □

Satz 2.1.9 (Cayley)**.** *Jede Gruppe G ist isomorph zu einer Untergruppe von* $\mathrm{Sym}(G)$*.*

Beweis. Sei $(G, \cdot)$ eine Gruppe. Wir ordnen jedem Element g von G die Linksmultiplikation $\lambda_g : G \to G$ zu, die definiert ist durch

$$\lambda_g(x) = g \cdot x.$$

Es gilt $\lambda_e = \mathrm{id}_G$ und

$$\lambda_{g \cdot h} = \lambda_g \circ \lambda_h,$$

weil $\lambda_{g \cdot h}(x) = (g \cdot h) \cdot x = g \cdot (h \cdot x) = \lambda_g(\lambda_h(x)) = (\lambda_g \circ \lambda_h)(x)$. Insbesondere gilt $\lambda_g \circ \lambda_{g^{-1}} = \lambda_{g^{-1}} \circ \lambda_g = \mathrm{id}_G$, woraus folgt, daß die λ_g Permutationen von G sind.

Also definiert $\lambda(g) = \lambda_g$ einen Homomorphismus

$$\lambda : G \to \mathrm{Sym}(G).$$

Das Bild $S = \lambda[G]$ von λ ist daher eine Untergruppe von $\mathrm{Sym}(G)$. Es bleibt noch zu zeigen, daß λ injektiv ist. Dann ist

$$\lambda : G \xrightarrow{\cong} S$$

der gesuchte Isomorphismus. λ ist injektiv, weil

$$\lambda(g) = \lambda(h) \quad \Rightarrow \quad g = g \cdot e = \lambda(g)(e) = \lambda(h)(e) = h \cdot e = h.$$

□

2.2 $\mathbb{R}$–Vektorräume

Definition. *Ein $\mathbb{R}$–Vektorraum V ist ein $\mathbb{R}$–Modul. Das heißt: eine abelsche Gruppe V mit einer Multiplikation $\mathbb{R} \times V \to V$, sodaß für alle $\alpha, \beta \in \mathbb{R}$ und $u, v \in V$ die folgenden Rechenregeln gelten:*

a) $\alpha(u + v) = \alpha u + \alpha v$

[5] Vgl. S. 36.

b) $(\alpha+\beta)u = \alpha u + \beta u$

c) $(\alpha\beta)u = \alpha(\beta u)$

d) $1u = u$

In einem Vektorraum gilt immer

- $0u = 0$
- $(-1)u = -u$,

weil $0u + 0u = (0+0)u = 0u$ und

$$(-1)u + u = (-1)u + 1u = (-1+1)u = 0u = 0.$$

Definition.

1. *Ein Unterraum U von V ist eine Untergruppe, die unter Multiplikation mit Skalaren abgeschlossen ist.*
2. *Eine lineare Abbildung $f : V \to U$ ist ein Homomorphismus, der mit der Multiplikation mit Skalaren kommutiert: $f(\alpha v) = \alpha f(v)$ für alle $v \in V$ und alle $\alpha \in \mathbb{R}$. Eine bijektive lineare Abbildung heißt Isomorphismus.*

Eine Teilmenge U von V ist genau dann ein Unterraum, wenn

- $0 \in U$
- $u, v \in U \;\Rightarrow\; u + v \in U$
- $\alpha \in \mathbb{R}, u \in U \;\Rightarrow\; \alpha u \in U$

Jeder Vektorraum V enthält den Nullraum $\{0\}$. Wir bezeichnen diese Nullräume immer mit dem gleichen Symbol $\mathbf{0}$.

Lemma 2.2.1. *Die linearen Abbildungen $f : \mathbb{R}^n \to \mathbb{R}^m$ sind genau die linearen Abbildungen im Sinn von Abschnitt 1.5.*

Beweis. Sei A eine m–n–Matrix. Die durch A definiert Abbildung $f_A : \mathbb{R}^n \to \mathbb{R}^m$ ist linear, weil (nach 1.5.3) für n–reihige Spaltenvektoren s und t

$$A(s+t) = As + At$$

und

$$A(\alpha s) = \alpha(As).$$

Wenn umgekehrt f linear ist, definieren wir die Matrix A als die Matrix mit den Spalten

$$f(\mathrm{e}_1), \ldots, f(\mathrm{e}_n).$$

Dann ist $f = f_A$, weil

$$\begin{aligned} f(\xi_1 \mathrm{e}_1 + \ldots + \xi_n \mathrm{e}_n) &= \xi_1 f(\mathrm{e}_1) + \ldots + \xi_n f(\mathrm{e}_n) \\ &= \xi_1 f_A(\mathrm{e}_1) + \ldots + \xi_n f_A(\mathrm{e}_n) \\ &= f_A(\xi_1 \mathrm{e}_1 + \ldots + \xi_n \mathrm{e}_n). \end{aligned}$$

□

Die Menge $\mathrm{L}(V, U)$ der linearen Abbildungen ist ein $\mathbb{R}$–Vektorraum unter wertweiser Addition und Multiplikation mit Skalaren. Aus dem letzten Lemma ergibt sich

$$\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m) \cong \mathrm{M}_{mn}(\mathbb{R}).$$

Das folgende rechnet man leicht nach (siehe auch Folgerung 1.5.3):

Lemma 2.2.2. *Die Kompositionsabbildung*

$$\circ : \mathrm{L}(U, V) \times \mathrm{L}(V, W) \to \mathrm{L}(U, W),$$

definiert durch

$$(f, g) \mapsto g \circ f,$$

ist bilinear. (Bilinearität wird definiert auf Seite 38.) □

2.3 Endlichdimensionale Vektorräume

Definition. *Eine Folge $b_1, \ldots, b_n$ heißt Basis von V, wenn sich jedes Element von V eindeutig in der Form*

$$v = \sum_{j=1}^{n} \xi_j b_j$$

schreiben läßt.

Eine Basis heißt auch *Koordinatensystem*, $\xi_1, \ldots, \xi_n$ sind die *Koordinaten* von v.

Die kanonischen Basisvektoren $\mathrm{e}_1, \ldots, \mathrm{e}_n$ bilden eine Basis des $\mathbb{R}^n$, weil

$$\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \xi_1 \mathrm{e}_1 + \ldots + \xi_n \mathrm{e}_n.$$

Die leere Folge (die Folge der Länge 0) ist Basis des Nullraums.

Satz 2.3.1. *Sei $v_1, \dots, v_n$ eine Basis von V, und $\mathrm{e}_1, \dots, \mathrm{e}_n$ die kanonische Basis von $\mathbb{R}^n$. Dann ist die lineare Abbildung $f : \mathbb{R}^n \to V$, die eindeutig durch $f(\mathrm{e}_j) = v_j$ bestimmt ist, ein Isomorphismus.*

Beweis. Wenn $f : \mathbb{R}^n \to V$ linear ist, ist

$$f(\xi_1 \mathrm{e}_1 + \ldots + \xi_n \mathrm{e}_n) = \xi_1 f(\mathrm{e}_1) + \ldots + \xi_n f(\mathrm{e}_n).$$

f ist also durch die Bilder der Basisvektoren eindeutig bestimmt.

Wenn wir uns umgekehrt Vektoren v_i vorgeben, definiert

$$f(\xi_1 \mathrm{e}_1 + \ldots + \xi_n \mathrm{e}_n) = \xi_1 v_1 + \ldots + \xi_n v_n$$

eine lineare Abbildung, die e_i auf v_i abbildet. Das ist leicht nachzurechnen: Sei $x = \xi_1 \mathrm{e}_1 + \ldots + \xi_n \mathrm{e}_n$ und $x' = \xi'_1 \mathrm{e}_1 + \ldots + \xi'_n \mathrm{e}_n$. Dann ist

$$\begin{aligned} f\big(x + x'\big) &= f\big((\xi_1 + \xi'_1)\mathrm{e}_1 + \ldots + (\xi_n + \xi'_n)\mathrm{e}_n\big) \\ &= (\xi_1 + \xi'_1) v_1 + \ldots + (\xi_n + \xi'_n) v_n \\ &= (\xi_1 v_1 + \ldots + \xi_n v_n) + (\xi'_1 v_1 + \ldots + \xi'_n v_n) \\ &= f(x) + f(x') \end{aligned}$$

und

$$\begin{aligned} f(\alpha x) &= f\big((\alpha\xi_1)\mathrm{e}_1 + \ldots + (\alpha\xi_n)\mathrm{e}_n\big) \\ &= (\alpha\xi_1) v_1 + \ldots + (\alpha\xi_n) v_n \\ &= \alpha(\xi_1 v_1 + \ldots + \xi_n v_n) \\ &= \alpha f(x). \end{aligned}$$

Wenn man jeden Vektor v aus V in der Form $v = \xi_1 v_1 + \ldots + \xi_n v_n$ schreiben kann, ist f surjektiv. Wenn v die ξ_i eindeutig bestimmt, ist f injektiv. $\square$

Die Wahl einer Basis von V ist gleichbedeutend mit der Wahl eines Isomorphismus zwischen $\mathbb{R}^n$ und V.

Sei $f : V \to U$ eine lineare Abbildung. $v_1, \dots, v_n$ und $u_1, \dots, u_m$ seien Basen von V und U. Wenn die Isomorphismen $\epsilon : \mathbb{R}^n \to V$ und $\delta : \mathbb{R}^m \to U$ die kanonischen Basen von $\mathbb{R}^n$ und $\mathbb{R}^m$ auf diese Basen abbilden, ist die lineare Abbildung $\delta^{-1} \circ f \circ \epsilon : \mathbb{R}^n \to \mathbb{R}^m$ durch eine m–n–Matrix $A = (\alpha_{ij})$ gegeben. Es gilt dann

$$f(\sum_{j=1}^{n} \xi_j v_j) = \sum_{i=1}^{m} (\sum_{j=1}^{n} \alpha_{ij} \xi_j) u_i.$$

Insbesondere ist

$$f(v_j) = \sum_{i=1}^{m} \alpha_{ij} u_i.$$

Die j–te Spalte der m–n–Matrix $A = (\alpha_{ij})$ besteht also aus den Koeffizienten des Bildes von v_j bezogen auf die Basis $u_1, \ldots, u_m$.

Wenn man umgekehrt A dadurch definiert, hat man in symbolischer Schreibweise

$$f(v_j) = (u_1 \ldots u_m) \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{mj} \end{pmatrix},$$

zusammengefaßt:

$$f(v_1 \ldots v_n) = (u_1 \ldots u_m)A. \tag{2.1}$$

Ein beliebiger Vektor

$$v = (v_1 \ldots v_n) \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

wird dann abgebildet auf

$$f(v) = f(v_1 \ldots v_n) \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = (u_1 \ldots u_m) A \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}.$$

$f \circ \epsilon = \delta \circ f_A$ bedeutet, daß das folgende Diagramm *kommutativ* ist:

$$\begin{tikzcd} V \arrow[r, "f"] & U \\ \mathbb{R}^n \arrow[u, "\epsilon"] \arrow[r, "f_A"'] & \mathbb{R}^m \arrow[u, "\delta"'] \end{tikzcd}$$

Wir sagen, daß f bezüglich der Basen (v_j) und (u_i) zur Matrix A gehört.

Definition. *Eine Folge $a_1, \ldots, a_n$ heißt Erzeugendensystem von V, wenn $V = \mathbb{R}a_1 + \cdots + \mathbb{R}a_n$. Das heißt, daß sich jedes Element von V in der Form*

$$\xi_1 a_1 + \ldots + \xi_n a_n \tag{2.2}$$

darstellen läßt.

Man sieht leicht, daß für jede Folge $a_1, \ldots, a_n$ die Menge

$$\langle a_1, \ldots, a_n \rangle = \{\xi_1 a_1 + \cdots + \xi_n a_n \mid a_1, \ldots, a_n \in A,\ \xi_1, \ldots, \xi_n \in \mathbb{R}\}$$

ein Unterraum von V ist. Man nennt diesen Unterraum das *Erzeugnis* von $a_1, \ldots, a_n$ oder den von $a_1, \ldots, a_n$ erzeugten (oder aufgespannten) Unterraum.

Definition. *Eine Folge $u_1, \ldots, u_n$ heißt linear unabhängig, wenn für alle $\xi_1, \ldots \xi_n \in \mathbb{R}$*

$$0 = \xi_1 u_1 + \ldots + \xi_n u_n \quad \Longrightarrow \quad \xi_1 = \cdots = \xi_n = 0.$$

Wenn die u_j paarweise verschieden sind, heißt die Menge $\{u_1, \ldots, u_n\}$ linear unabhängig, wenn $u_1, \ldots, u_n$ linear unabhängig ist.

Eine Folge, in der 0 oder zwei gleiche Vektoren vorkommen, ist linear abhängig.

Lemma 2.3.2. *$u_1, \ldots, u_n$ ist genau linear unabhängig, wenn die Darstellung der Elemente von $\langle u_1, \ldots, u_n \rangle$ als $\xi_1 u_1 + \ldots + \xi_n u_n$ eindeutig ist.*

Beweis. Wenn $u_1, \ldots, u_n$ linear abhängig sind, hat man zwei verschieden Darstellungen der Null, nämlich $0 = \xi_1 u_1 + \ldots + \xi_n u_n$ und $0 = 0u_1 + \ldots + 0u_n$.

Nehmen wir umgekehrt an, daß $u_1, \ldots, u_n$ linear unabhängig sind und daß

$$\xi_1 u_1 + \ldots + \xi_n u_n = \xi_1' u_1 + \ldots + \xi_n' u_n.$$

Dann folgt

$$(\xi_1 - \xi_1') u_1 + \ldots + (\xi_n - \xi_n') u_n = 0.$$

Und daraus

$$\xi_1 - \xi_1' = \cdots = \xi_n - \xi_n' = 0.$$

Also ist $\xi_1 = \xi_1', \ldots, \xi_n = \xi_n'$. $\square$

Folgerung. *Eine Basis ist ein linear unabhängiges Erzeugendensystem.*

Wir nennen eine endliche Menge E ein Erzeugendensystem, wenn jedes Element von V eine Linearkombination von Elementen von E ist. Eine Menge U heißt linear unabhängig, wenn $0 = \sum_{u \in U} \xi_u u$, impliziert, daß alle $\xi_u = 0$. Eine Basis ist ein linear unabhängiges Erzeugendensystem. Wenn $B = \{b_1, \ldots, b_n\}$ für paarweise verschiedene b_i, dann ist B genau dann eine Basis, wenn die Folge $b_1, \ldots, b_n$ eine Basis ist.

Lemma 2.3.3. *Sei V ein $\mathbb{R}$–Vektorraum mit Erzeugendensystem E und U eine linear unabhängige Teilmenge von E. Dann hat V eine Basis B, die zwischen U und E liegt:*

$$U \subset B \subset E.$$

Beweis. Wir erweitern U zu einer maximalen linear unabhängigen Teilmenge $B \subset E$. B ist ein Erzeugendensystem. Dazu genügt es zu zeigen, daß alle $e \in E$ im Erzeugnis von B liegen. Sei $e \in E \setminus B$. Weil $B \cup \{e\}$ linear abhängig sein muß, gibt es $b_1, \ldots, b_n \in B$ und $\xi_1, \ldots, \xi_n, \zeta \in \mathbb{R}$, nicht alle Null, mit

$$\xi_1 b_1 + \cdots + \xi_n b_n + \zeta e = 0.$$

Weil die $b_1, \ldots, b_n$ linear unabhängig sind[6], kann ζ nicht Null sein. Dann liegt

$$e = \frac{-\xi_1}{\zeta} b_1 + \ldots + \frac{-\xi_n}{\zeta} b_n$$

im Erzeugnis von B. □

Definition. *V heißt endlich erzeugt, wenn V ein (endliches) Erzeugendensystem hat.*

Folgerung.

1. *Jedes Erzeugendensystem enthält eine Basis.*

2. *Wenn V endlich erzeugt ist, läßt sich jede unabhängige Menge zu einer Basis erweitern.* □

Es folgt, daß jeder endlich erzeugte Vektorraum zu einem $\mathbb{R}^n$ isomorph ist.

Satz 2.3.4. *Sei V ein $\mathbb{R}$–Vektorraum. Eine linear unabhängige Teilmenge hat höchstens so viele Elemente wie ein Erzeugendensystem von V.*

Beweis. Sei $e_1, \ldots, e_m$ ein Erzeugendensystem, $u_1, \ldots, u_n$ linear unabhängigbeliebige Vektoren. Für eine geeignete Matrix $(\alpha_{ij}) = A$ von Koeffizienten ist

$$u_j = \sum_{i=1}^{m} \alpha_{ij} e_i \quad (j = 1, \ldots, n).$$

In symbolischer Schreibweise

$$(u_1 \ldots u_n) = (e_1 \ldots e_m) A.$$

Wenn $m < n$, hat nach 1.1.5 das Gleichungssystem

$$\sum_{j=1}^{n} \alpha_{ij} \xi_j = 0 \quad (i = 1, \ldots, m)$$

eine nicht–triviale Lösung. Symbolisch

$$A \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = 0.$$

[6]wenn wir die b_i paarweise verschieden gewählt haben

Wenn wir diese Gleichungen mit den e_i multiplizieren erhalten wir eine lineare Abhängigkeit zwischen den u_j:

$$\sum_{j=1}^{n} u_j \xi_j = \sum_{j=1}^{n} \sum_{i=1}^{m} \alpha_{ij} \xi_j e_i = \sum_{i=1}^{m} (\sum_{j=1} \alpha_{ij} \xi_j) e_i = 0,$$

symbolisch:

$$(u_1 \dots u_n) \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = (e_1 \dots e_m) A \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = (e_1 \dots e_m) \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = 0.$$

Wenn $m < n$, sind die u_j also nicht linear unabhängig. □

Folgerung 2.3.5. *Alle Basen eines endlich erzeugten Vektorraums V haben die gleiche Mächtigkeit: die Dimension*

$$\dim(V)$$

von V. Zwei endlich erzeugte Vektorräume sind genau dann isomorph, wenn sie die gleiche Dimension haben. □

Folgerung 2.3.6.

1. *Ein Unterraum U eines endlich erzeugten Vektorraums V ist endlich erzeugt. Es gilt* $\dim(U) \leq \dim(V)$.

2. *Sei V endlich erzeugt und $f : V \to U$ linear. Dann ist auch das Bild $f[V]$ endlich erzeugt und hat höchstens die Dimension von V.*

Beweis.

1): Sei V von m Elementen erzeugt. Weil linear unabhängige Mengen höchstens m Elemente haben, enthält U eine maximale linear unabhängige Menge B. B ist eine Basis von U.

2): Wenn E ein Erzeugendensystem von V ist, ist $f[E]$ ein Erzeugendensystem von $f[V]$. □

2.4 Unendlichdimensionale Vektorräume

Definition. *Für eine Menge I sei $\mathbb{R}^{(I)}$ die Menge alle Funktionen $f : I \to \mathbb{R}$, die an fast allen Stellen (d.h. an allen bis auf endlich vielen) den Wert* 0 *hat.* $\mathbb{R}^{(I)}$ *ist ein Vektorraum mit wertweiser Addition und Multiplikation mit Skalaren.*

$\mathbb{R}^{(I)}$ ist ein Unterraum von $\mathbb{R}^I$, dem Raum *aller* Funktionen von I nach $\mathbb{R}$.

Wir wollen zeigen, daß jeder Vektorraum zu einem $\mathbb{R}^{(I)}$ isomorph ist. Wir verallgemeinern dazu die Begriffe aus Abschnitt 2.2 auf unendliche Mengen.

Definition. *Sei V ein Vektorraum und $(v_i)_{i\in I}$ eine Familie von Vektoren.*

- *$(v_i)_{i\in I}$ ist eine Basis von V, wenn sich jeder Vektor $v \in V$ eindeutig in der Form*
$$v = \sum_{i\in I} \xi_i v_i$$
schreiben läßt. Dabei dürfen nur endlich viele der ξ_i ungleich Null sein, damit die eventuell unendliche Summe sinnvoll ist.

- *$(v_i)_{i\in I}$ ist linear unabhängig, wenn*
$$0 = \sum_{i\in I} \xi_i v_i \quad \Longrightarrow \quad \xi_i = 0 \text{ für alle } i \in I.$$

- *$(v_i)_{i\in I}$ ist ein Erzeugendensystem, wenn sich jedes Element von V in der Form $\sum_{i\in I} \xi_i v_i$ schreiben läßt.*

Eine Menge B heißt Basis (linear unabhängig, Erzeugendensystem), wenn $(b)_{b\in B}$ eine Basis (linear unabhängig, Erzeugendensystem) ist.

Wie in der Bemerkung S.28 sieht man, daß Basen nichts anderes als linear unabhängige Erzeugendensysteme sind.

Satz 2.4.1. *Jeder Vektorraum ist zu einem $\mathbb{R}^{(I)}$ isomorph.*

Beweis. Sei V ein Vektorraum. Wir betrachten die Menge P aller linear unabhängigen Teilmengen von V und ordnen sie durch Inklusion. Weil die Vereinigung einer Kette von linear unabhängigen Mengen wieder linear unabhängig ist, können wir Zorns Lemma (2.4.2) auf P anwenden und erhalten eine maximale linear unabhängige Menge B. Wie im Beweis von 2.3.3 sieht man, daß B eine Basis von V ist. Die Zuordnung

$$(\xi_b)_{b\in B} \longmapsto \sum_{b\in B} \xi_b b$$

definiert offensichtlich ein Isomorphismus zwischen $\mathbb{R}^{(B)}$ und V. $\square$

Definition. *Eine partielle Ordnung $\leq$ auf P ist eine zweistellige Relation, die reflexiv, transitiv und antisymmetrisch ist. Dabei bedeutet*

reflexiv	$x \leq x$
transitiv	$x \leq y \;\&\; y \leq z \Rightarrow x \leq z$
antisymmetrisch	$x \leq y \;\&\; y \leq x \Rightarrow x = y$

für alle $x, y, z \in P$.

In einer totalen (oder linearen) Ordnung sind alle Elemente x, y vergleichbar, das heißt $x \leq y$ oder $y \leq x$.

Eine Teilmenge A von P ist eine *Kette*, wenn A durch $\leq$ linear geordnet ist.

$p \in P$ ist eine *obere Schranke* von A, wenn $a \leq p$ für alle $a \in A$.

$p \in P$ ist ein *maximales* Element von P, wenn $p \leq p' \Rightarrow p = p'$ für alle $p' \in P$.

Satz 2.4.2 (Zornsches Lemma)**.** *Sei $(P, \leq)$ eine partielle Ordnung. Wenn jede Kette in P eine obere Schranke hat, hat P ein maximales Element.*

Beweis. Das Zornsche Lemma wird in der Mengenlehre bewiesen. Es ist zum *Auswahlaxiom* äquivalent. □

2.5 Der Verband der Unterräume

Sei $(P, \leq)$ eine partielle Ordnung. p heißt *größtes* Element von P, wenn $p' \leq p$ für alle $p' \in P$. P hat höchstens ein größtes Element. Kleinstes Element von P wird analog definiert.

Sei A eine Teilmenge von P. Eine (die) kleinste obere Schranke von A heißt *Supremum* $\sup(A)$ von A. Das *Infimum* $\inf(A)$ ist größte untere Schranke.

Definition. *Eine partielle Ordnung heißt Verband, wenn je zwei Elemente ein Supremum und ein Infimum haben.*

Satz 2.5.1. *Sei V ein Vektorraum. Dann bilden die durch Inklusion geordneten Unterräume von V einen Verband. Dabei ist $\sup(U_1, U_2) = U_1 + U_2$ und $\inf(U_1, U_2) = U_1 \cap U_2$.*

Beweis. Offensichtlich ist $U_1 \cap U_2$ der größte gemeinsame Unterraum von U_1 und U_2 und

$$U_1 + U_2 = \{u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2\}$$

ist $\langle U_1 \cup U_2 \rangle$, der kleinste Unterraum, der U_1 und U_2 enthält. □

Der Nullraum $\mathbf{0}$ ist kleinstes Element, V größtes Element dieses Verbandes. Der Verband der Unterräume ist zwar im Allgemeinen nicht distributiv[7], es gilt aber eine Abschwächung des Distributivgesetzes, das *Modularitätsgesetz*:

$$U_2 \subset W \quad \Rightarrow \quad (U_1 + U_2) \cap W = (U_1 \cap W) + U_2.$$

[7] Das Distributivgesetz $(U_1 + U_2) \cap W = (U_1 \cap W) + (U_2 \cap W)$ gilt nicht, wenn $\dim(V) \geq 2$.

Die rechte Seite ist klarerweise in der linken enthalten. Sei umgekehrt $w \in (U_1 + U_2) \cap W$. Wenn wir $w = u_1 + u_2$ für $u_i \in U_i$ schreiben, folgt daß u_1 in W liegt, weil w und u_2 zu W gehören. Also ist $u_1 + u_2 \in (U_1 \cap W) + U_2$.

Lemma 2.5.2. *Jeder Unterraum U von V hat ein* Komplement. *Ein Komplement von U ist ein Unterraum U' mit*

$$U \cap U' = \mathbf{0} \quad \textit{und} \quad U + U' = V.$$

Beweis. Wir nennen U und U' *unabhängig*, wenn ihr Durchschnitt Null ist, das heißt, wenn für alle $u \in U$ und $u' \in U'$

$$u + u' = 0 \quad \Rightarrow \quad u = u' = 0.$$

Mit Zorns Lemma verschaffen wir einen maximalen Unterraum U', der unabhängig von U ist. Wir zeigen $U + U' = V$. Sei $v \in V$ ein beliebiger Vektor. Wir können annehmen, daß $v \notin U'$. Dann muß $U' + \mathbb{R}v$ abhängig von U sein. Es gibt also $u \in U, u' \in U', \alpha \in \mathbb{R}$ mit $u = u' + \alpha v \neq 0$. Weil U und U' unabhängig sind, muß $\alpha \neq 0$ sein. Es folgt

$$v = \frac{1}{\alpha}(u - u') \in U + U'.$$

□

U' läßt sich auch folgendermaßen konstruieren: Man wählt eine Basis B von U und setzt sie (mit Zorns Lemma) zu einer Basis $B \mathbin{\dot\cup} B'$ von V fort. Für U' nehmen wir den von B' aufgespannten Unterraum. Weil $B \mathbin{\dot\cup} B'$ linear unabhängig ist, sind U und U' unabhängig. Weil $B \mathbin{\dot\cup} B'$ ein Erzeugendensystem von V ist, ist $V = U + U'$. Der Beweis des nächsten Lemmas zeigt, daß man alle Komplemente von U auf diese Weise erhält.

Lemma 2.5.3. *Sei V endlichdimensional, U ein Unterraum und U' eine Komplement von U in V. Dann ist*

$$\dim(V) = \dim(U) + \dim(U').$$

Beweis. Wähle eine Basis B von U und eine Basis B' von U'. Dann ist $B \mathbin{\dot\cup} B'$ eine Basis von V. □

Das direkte Produkt $U_1 \times U_2$ zweier Vektorräume wird mit elementweisen Operationen ein Vektorraum, die *direkte Summe* $U_1 \bigoplus U_2$.

Lemma 2.5.4. *Wenn U_1 und U_2 unabhängig sind, ist die lineare Abbildung von $U_1 \bigoplus U_2 \longrightarrow U_1 + U_2$, definiert durch $(u_1, u_2) \mapsto u_1 + u_2$, ein Isomorphismus.*

Beweis. Die Abbildung ist offensichtlich linear und surjektiv. Sie ist injektiv, weil, U_1 und U_2 unabhängig sind: Aus $u_1 + u_2 = u_1' + u_2'$ folgt $u_1 - u_1' = u_2' - u_2$ und daraus $u_1 - u_1' = u_2' - u_2 = 0$ und $u_1 = u_1',\ u_2 = u_2'$. □

Folgerung. *Wenn U' Komplement von U ist, ist $V \cong U \bigoplus U'$.*

Zwei Komplemente U' und U'' von U sind kanonisch isomorph. Man definiert durch
$$f(u') = u'' \Leftrightarrow u' - u'' \in U$$
einen Isomorphismus $f : U' \to U''$. f ist auf ganz U' wohldefiniert, weil es zu jedem $u' \in U$ eine eindeutige Darstellung $u' = u + u''$ gibt. Vertauscht man die Rollen von U' und U'', erkennt man, daß f injektiv und surjektiv ist. Die Linearität von f ist klar.

Sei U ein Unterraum von V. Mengen der Form $v + U$ heißen *Nebenklassen* von U.

Lemma 2.5.5. *Nebenklassen von U sind gleich oder disjunkt. Es gilt $v_1 + U = v_2 + U \Leftrightarrow v_1 - v_2 \in U$.*

Beweis. Wenn $v_1 + U$ und $v_2 + U$ nicht disjunkt sind, gibt es $u_i \in U$ mit $v_1 + u_1 = v_2 + u_2$. Es folgt $v_1 - v_2 = u_2 - u_1 \in U$. Wenn umgekehrt $v_1 - v_2 \in U$, folgt $v_1 + U = v_2 + (v_1 - v_2) + U = v_2 + U$. □

Definition. *$V/U = \{v + U \mid v \in V\}$ heißt der Quotient von V nach U. Die Abbildung*
$$\pi : v \mapsto v + U$$
heißt kanonische Projektion.

Satz 2.5.6. *V/U trägt eine Vektorraumstruktur, die eindeutig dadurch bestimmt ist, daß die kanonische Projektion*
$$\pi : V \to V/U$$
linear ist.

Beweis. Sei V/U mit einer Vektorraumstruktur versehen und π linear. Dann lassen sich in V/U die Summe zweier Elemente a und b und das Produkt mit einem $\alpha \in \mathbb{R}$ so ausrechnen: Wir wählen Repräsentanten x, y der Nebenklassen, sodaß $\pi(x) = a$ und $\pi(y) = b$. Dann ist
$$a + b = \pi(x + y) \tag{2.3}$$
und
$$\alpha \cdot a = \pi(\alpha \cdot x) \tag{2.4}$$
Es folgt daß Addition und Multiplikation eindeutig bestimmt sind.

Um die Existenz einer Vektorraumstruktur einzusehen, definieren wir Addition und Multiplikation durch (2.3) und (2.4). Wir müssen aber nachrechnen, daß

$a+b$ und αa unabhängig von der Wahl von x und y definiert sind.
Sei also $\pi(x') = a$ und $\pi(y') = b$. Dann gehören $x - x'$ und $y - y'$ zu U und damit auch ihre Summe $(x+y) - (x'+y')$. Es folgt $\pi(x+y) = \pi(x'+y')$ und $a+b$ ist wohldefiniert.
Um die Wohldefiniertheit der Multiplikation einzusehen, nehmen wir an, daß $\pi(x') = a$. Dann folgt $x - x' \in U$ und damit auch $\alpha x - \alpha x' = \alpha(x - x') \in U$ und $\pi(\alpha x) = \pi(\alpha x')$.

Die Vektorraumaxiome übertragen sich dann sofort auf dem Quotienten. Betrachten wir als ein Beispiel die Assoziativität der Addition:

$$\begin{aligned}(\pi(x) + \pi(y)) + \pi(z) &= \pi(x+y) + \pi(z) = \pi((x+y)+z)\\ &= \pi(x+(y+z)) = \pi(x) + \pi(y+z)\\ &= \pi(y) + (\pi(y) + \pi(z)).\end{aligned}$$

□

Lemma 2.5.7. *Sei U' ein Komplement von U in V. Dann induziert π einen Isomorphismus $U' \to V/U$.*

Beweis. Sei π' die Einschränkung von π auf U'. Wenn $v = u + u'$ für $u \in U,\ u' \in U'$, ist $\pi'(u') = v + U$. Aus $U + U' = V$ folgt also die Surjektivität von π'. Und aus $U \cap U' = \mathbf{0}$ die Injektivität:

$$\pi'(u_1') = \pi'(u_2') \quad \Rightarrow \quad u_1' - u_2' \in U \quad \Rightarrow \quad u_1' = u_2'.$$

□

Folgerung 2.5.8. *Wenn V endlichdimensional ist, ist*

$$\dim(U) + \dim(V/U) = \dim(V).$$

Beweis. Die Behauptung folgt aus 2.5.3 und 2.5.7 □

Definition. *Die* Kodimension *von U in V ist die Dimension von V/U.*

$$\operatorname{codim}_V(U) = \dim(V/U)$$

Satz 2.5.9. *Für endlichdimensionale Vektorräume gilt*

$$\dim(U_1 + U_2) + \dim(U_1 \cap U_2) = \dim(U_1) + \dim(U_2)$$

Beweis. Sei H ein Komplement von $U_1 \cap U_2$ in U_2. Dann ist H ein Komplement von U_1 in $U_1 + U_2$: Klar ist, daß $U_1 + H = U_1 + (U_1 \cap U_2) + H = U_1 + U_2$. Andererseits ist $U_1 \cap H = U_1 \cap U_2 \cap H = \mathbf{0}$.

Es folgt $\operatorname{codim}_{U_1+U_2}(U_1) = \operatorname{codim}_{U_2}(U_1 \cap U_2)$. Nach 2.5.8 ist

$$\operatorname{codim}_{U_1+U_2}(U_1) = \dim(U_1 + U_2) - \dim(U_1)$$

und

$$\operatorname{codim}_{U_2}(U_1 \cap U_2) = \dim(U_2) - \dim(U_1 \cap U_2).$$

Daraus folgt die Behauptung durch Gleichsetzen der rechten Seiten. □

Wenn $\mathrm{U}_1 + U_2 = V$, läßt sich die Behauptung des letzten Satzes ausdrücken durch

$$\operatorname{codim}(U_1 \cap U_2) = \operatorname{codim}(U_1) + \operatorname{codim}(U_2).$$

Äquivalenzrelationen und Partitionen:

Sei X eine Menge. Eine *Äquivalenzrelation* E ist eine zweistellige reflexive, transitive und symmetrische (d.h. $xEy \Rightarrow yEx$) Relation auf X. Eine *Partition* von X ist eine Teilmenge $\mathfrak{B}$ der Potenzmenge $\mathfrak{P}(X) = \{B \mid B \subset X\}$ von X, deren Elemente nichtleer, paarweise disjunkt sind und deren Vereinigung X ist.

Äquivalenzrelationen und Partitionen entsprechen einander: Setzt man $x/E = \{y \in X \mid yEx\}$ für $x \in X$ und eine Äquivalenzrelation E, so ist

$$\mathfrak{B} = \{x/E \mid x \in X\}$$

die zugehörige Partition von X. Wenn umgekehrt $\mathfrak{B}$ eine Partition ist, so erhält man mit

$$xEy \quad \Leftrightarrow \quad \exists B \in \mathfrak{B} \quad x \in B \;\&\; y \in B$$

die zugehörige Äquivalenzrelation.

Eine Unterraum U von V bestimmt auf V die Äquivalenzrelation $x - y \in U$ mit der Partition $\{v + U \mid v \in V\}$.

2.6 Körper

Ein *Ring* R ist ein Tripel $(R, +, \cdot)$, wobei

- $(R, +)$ eine abelsche Gruppe,
- $(R, \cdot)$ eine Halbgruppe ist und
- die Distributivgesetze

$$x(y + z) = xy + xz$$

und

$$(x + y)z = xz + yz$$

gelten.

Ein *Einselement* 1, falls vorhanden, erfüllt $x1 = 1x = x$.
R ist *kommutativ*, wenn die Multiplikation kommutativ ist.

Wie auf Seite 24 sieht man, daß in einem Ring immer $0x = x0 = 0$.

Definition. *Ein Körper ist ein kommutativer Ring mit Einselement* 1*, in dem jedes von Null verschiedene Element ein Inverses bezüglich der Multiplikation besitzt. Außerdem fordert man*[8]*, daß* $0 \neq 1$.

[8] Dadurch schließt man den Nullring $0 = \{0\}$ aus

Beispiele:

- $\mathbb{Z}$ ist ein kommutativer Ring mit 1.
- $\mathbb{Q}$ und $\mathbb{R}$ sind Körper.
- $\mathbb{Z}_n$ ist, mit repräsentantenweiser Multiplikation
$$\bar{y} \cdot \bar{z} = \overline{yz},$$
ein kommutativer Ring mit 1.
- $\mathbb{Z}_p$ ist genau dann ein Körper, wenn p eine Primzahl ist[9]. Man bezeichnet diesen Körper auch mit $\mathbb{F}_p$.

Um den Körper der komplexen Zahlen einzuführen, brauchen wir den Begriff einer $\mathbb{R}$–Algebra.

Definition. *Eine $\mathbb{R}$–Algebra A ist ein $\mathbb{R}$–Vektorraum mit einer bilinearen Operation (der Multiplikation).*

Dabei heißt eine Abbildung $\circ : A \times A \to A$ bilinear, wenn sie in beiden Argumenten linear ist, das heißt, wenn

- $(x + x') \circ y = x \circ y + x' \circ y$
- $x \circ (y + y') = x \circ y + x \circ y'$
- $(\alpha x) \circ y = x \circ (\alpha y) = \alpha(x \circ y)$

Beispiel: $\mathrm{M}_n(\mathbb{R})$ ist eine $\mathbb{R}$–Algebra der Dimension n^2.

Lemma 2.6.1. *Sei $b_1, \dots, b_n$ eine Basis von A. Für beliebige Vektoren a_{ij} $(i, j = 1, \dots, n)$ gibt es genau eine Algebren–Multiplikation $\circ$ auf V mit $b_i \circ b_j = a_{ij}$.*

Beweis. Seien $x = \sum_i \xi_i b_i$ und $y = \sum_j \zeta_j b_j$ zwei beliebige Elemente von A. Dann ist
$$x \circ y = \sum_{i,j} \xi_i \zeta_j (b_i \circ b_j).$$
Das zeigt, daß $\circ$ durch die $b_i \circ b_j$ eindeutig bestimmt ist. Umgekehrt definiert
$$x \circ y = \sum_{i,j} \xi_i \zeta_j a_{ij}$$
eine bilineare Abbildung. $\square$

[9] Ein Körper ist immer *nullteilerfrei*, das heißt, daß die Null nicht Produkt von Elementen ungleich Null sein kann. Man sieht sofort, daß $\mathbb{Z}_m$ genau dann nullteilerfrei ist, wenn m prim ist. In einem nullteilerfreien Ring R ist die Multiplikation λ_a mit einem $a \neq 0$ injektiv, weil $ax = ax' \Rightarrow a(x - x') = 0 \Rightarrow x - x' = 0$. Wenn R endlich ist, ist also λ_a surjektiv und a hat ein Rechtsinverses. Das beweist, daß die $\mathbb{Z}_p$, für Primzahlen p, Körper sind.

BEISPIEL: $\mathbb{R}^3$ wird mit dem *Kreuzprodukt* eine Algebra, wenn man die Produkte $\mathrm{e}_i \times \mathrm{e}_j$ der kanonischen Basisvektoren definiert durch

$$\begin{array}{rclcrclcrcl}
\mathrm{e}_1 \times \mathrm{e}_2 &=& \mathrm{e}_3 &,& \mathrm{e}_2 \times \mathrm{e}_3 &=& \mathrm{e}_1 &,& \mathrm{e}_3 \times \mathrm{e}_1 &=& \mathrm{e}_2 \\
\mathrm{e}_2 \times \mathrm{e}_1 &=& -\mathrm{e}_3 &,& \mathrm{e}_3 \times \mathrm{e}_2 &=& -\mathrm{e}_1 &,& \mathrm{e}_1 \times \mathrm{e}_3 &=& -\mathrm{e}_2 \\
\mathrm{e}_1 \times \mathrm{e}_1 &=& 0 &,& \mathrm{e}_2 \times \mathrm{e}_2 &=& 0 &,& \mathrm{e}_3 \times \mathrm{e}_3 &=& 0.
\end{array}$$

Das Kreuzprodukt ist nicht assoziativ. Es handelt sich vielmehr um eine *Liealgebra* (vgl. Seite 137).

Definition. *Die $\mathbb{R}$–Algebra $\mathbb{C}$ der komplexen Zahlen entsteht aus $\mathbb{R}^2$ durch folgende Multiplikation: (Die kanonische Basis des $\mathbb{R}^2$ sei hier mit $\mathbf{1}, \mathrm{i}$ bezeichnet.)*

$$\mathbf{1}\mathbf{1} = \mathbf{1}, \quad \mathbf{1}\mathrm{i} = \mathrm{i}\mathbf{1} = \mathrm{i}, \quad \mathrm{i}\mathrm{i} = -\mathbf{1}$$

Lemma 2.6.2. $\mathbb{C}$ *ist ein Körper.*

Beweis. Drei Dinge sind zu zeigen:

- $\mathbf{1}$ ist Einselement: $\mathbf{1}$ ist linksneutral, weil $\mathbf{1}(\alpha\mathbf{1} + \beta\,\mathrm{i}) = \alpha(\mathbf{1}\mathbf{1}) + \beta\,(\mathbf{1}\mathrm{i}) = \alpha\mathbf{1} + \beta\,\mathrm{i}$. Eine ähnliche Rechnung zeigt, daß $\mathbf{1}$ auch rechtsneutral ist.
- Die Kommutativität der Multiplikation: Ergibt sich leicht aus der Vertauschbarkeit der Basiselemente $\mathbf{1}\mathrm{i} = \mathrm{i}\mathbf{1}$.
- Die Assoziativität der Multiplikation: Man überprüft leicht, daß die Multiplikation auf den beiden Basiselementen assoziativ ist: Weil $\mathbf{1}$ Einselement ist, braucht man nur

  $$(\mathrm{i} \circ \mathrm{i}) \circ \mathrm{i} = (-\mathbf{1}) \circ \mathrm{i} = -\mathrm{i} = \mathrm{i} \circ (-\mathbf{1}) = \mathrm{i} \circ (\mathrm{i} \circ \mathrm{i})$$

  zu betrachten. Wenn aber die Multiplikation einer Algebra A auf einer Basis assoziativ ist, ist sie auf ganz A assoziativ. Die dreistellige multilineare[10] Funktion

  $$(x \circ y) \circ z - x \circ (y \circ z)$$

  verschwindet nämlich auf ganz A, wenn sie auf einer Basis verschwindet. (Das ist eine Verallgemeinerung von 2.6.1.)
- Existenz von Inversen: Sei $z = \alpha\mathbf{1} + \beta\,\mathrm{i}$. Man nennt

  $$\overline{z} = \alpha\mathbf{1} - \beta\,\mathrm{i}$$

  die konjugiert komplexe Zahl. Wenn z ungleich Null ist, ist das Produkt

  $$z\,\overline{z} = \alpha^2 + \beta^2$$

  ebenfalls ungleich Null und $\dfrac{\overline{z}}{\alpha^2 + \beta^2}$ ist Inverses von z.

[10] $\mu : V_1 \times V_2 \times \ldots \times V_m \longrightarrow W$ heißt multilinear, wenn für alle $v_1 \ldots, v_m \in V$ und für alle $i = 1, \ldots, m$ die Abbildung $\nu : V_i \to W$ definiert durch $\nu(x) = \mu(v_1, \ldots, v_{i-1}, x, v_{i+1}, \ldots, v_m)$ linear ist.

$\square$

Wir fassen $\mathbb{R}$ vermöge der Einbettung $\alpha \mapsto \alpha\,\mathbf{1}$ als Unterraum von $\mathbb{C}$ auf. Die kanonische Basis ist dann $1, \mathrm{i}$.

Definition. *Sei R ein Ring mit Einselement. Ein R–Modul ist eine abelsche Gruppe M mit einer Multiplikation $R \times M \to M$, die den folgenden Regeln genügt:*

a) $r(x+y)=rx+ry$

b) $(r+s)x=rx+sx$

c) $(rs)x=r(sx)$

d) $1x=x$

Ein K–Modul über einem Körper K heißt K–Vektorraum.

Alle Sätze des Kapitels gelten für beliebige K–Vektorräume.

Kapitel 3

Lineare Abbildungen

Wir fixieren einen Körper K. Alle in diesem Kapitel vorkommenden Vektorräume sind K–Vektorräume.

3.1 Der Noethersche Isomorphiesatz

Sei $f : V \to W$ eine lineare Abbildung.

Lemma 3.1.1. *Das Bild*

$$\mathrm{Im}(f) = f[V]$$

von f ist ein Unterraum von W. Der Kern

$$\mathrm{Ker}(f) = \{v \in V \mid f(v) = 0\}$$

ein Unterraum von V.

Beweis. Bild und Kern enthalten den Nullvektor, weil aus $f(0) = f(0 + 0) = f(0) + f(0)$ folgt, daß

$$f(0) = 0.$$

Die Summe $f(v_1) + f(v_2)$ zweier Elemente von $\mathrm{Im}(f)$ ist gleich $f(v_1 + v_2)$, also wieder im Bild von f. Wenn $f(v_1) = f(v_2) = 0$, ist auch $f(v_1 + v_2) = f(v_1) + f(v_2) = 0 + 0 = 0$.
Mit $f(v_1)$ ist auch $\alpha f(v_1) = f(\alpha v_1)$ im Bild. Wenn v_1 im Kern ist, ist $f(\alpha v_1) = \alpha f(v_1) = 0$. Also ist auch αv_1 im Kern. $\square$

f ist genau dann injektiv, wenn $\operatorname{Ker}(f) = \mathbf{0}$. Denn

$$f(v_1) = f(v_2) \Leftrightarrow v_1 - v_2 \in \operatorname{Ker}(f) \tag{3.1}$$

Beispiele:
Sei U ein Unterraum von V.

- Für die Inklusionsabbildung $\iota : U \to V$ ist $\operatorname{Im}(\iota) = U$.
- Für die kanonische Projektion $\pi : V \to V/U$ ist $\operatorname{Ker}(\pi) = U$.
- Sei U' ein Komplement von U und $\pi : V \to U'$ die *Projektion auf U' entlang U*,
  $$\pi(u + u') = u'.$$
  Dann ist $\operatorname{Im}(\pi) = U'$ und $\operatorname{Ker}(\pi) = U$.

Notation
Sei $f : X \to Y$ eine Abbildung, a ein Element von Y und A eine Teilmenge von Y. Dann bezeichnen wir mit $f^{-1}(a)$ und $f^{-1}(A)$ jeweils die Menge $\{x \mid f(x) = a\}$ bzw. $\{x \mid f(x) \in A\}$ der Urbilder von a bzw. A. (Wenn f eine Bijektion ist, hat $f^{-1}(a)$ zwei Bedeutungen!)

Offenbar ist $\operatorname{Ker}(f) = f^{-1}(0)$. Wenn $f(a) = b$, ist $a + \operatorname{Ker}(f) = f^{-1}(b)$.

Satz 3.1.2 (Noetherscher Isomorphiesatz)**.** *f induziert einen Isomorphismus*

$$\overline{f} : V/\operatorname{Ker}(f) \to \operatorname{Im}(f)$$

.

Beweis. Sei π die kanonische Projektion von V auf $V/\operatorname{Ker}(f)$. Dann ist

$$f(v_1) = f(v_2) \;\Leftrightarrow\; v_1 - v_2 \in \operatorname{Ker}(f) \;\Leftrightarrow\; \pi(v_1) = \pi(v_2).$$

Aus "$\Leftarrow$" folgt sofort, daß man f durch π faktorisieren kann. Das heißt, daß $f = \overline{f} \circ \pi$ für eine Abbildung $\overline{f} : V/\operatorname{Ker}(f) \to W$.

$$\begin{tikzcd}
V \arrow[r, "f"] \arrow[d, "\pi"'] & W \\
V/\operatorname{Ker}(f) \arrow[ur, "\overline{f}"']
\end{tikzcd}$$

Weil π surjektiv ist, ist $\overline{f}$ eindeutig bestimmt und hat $\operatorname{Im}(f)$ als Bild. Aus "$\Rightarrow$" folgt, daß $\overline{f}$ injektiv ist. Es bleibt die Linearität von $\overline{f}$ zu zeigen. Seien dazu c_1

und c_2 beliebige Elemente von $V/\operatorname{Ker}(f)$ und $\alpha \in K$. Für geeignete $v_1, v_2 \in V$ ist $\pi(v_1) = c_1$ und $\pi(v_2) = c_2$. Wir haben dann

$$\begin{aligned}\overline{f}(c_1 + c_2) &= \overline{f}(\pi(v_1) + \pi(v_2)) = \overline{f}(\pi(v_1 + v_2)) \\ &= f(v_1 + v_2) = f(v_1) + f(v_2) \\ &= \overline{f}(c_1) + \overline{f}(c_2)\end{aligned}$$

und

$$\begin{aligned}\overline{f}(\alpha c_1) &= \overline{f}(\alpha\pi(v_1)) = \overline{f}(\pi(\alpha v_1)) \\ &= f(\alpha v_1) = \alpha f(v_1) \\ &= \alpha\overline{f}(c_1).\end{aligned}$$

□

Die Abbildung $\overline{f}$ ist gegeben durch $\overline{f}(v + \operatorname{Ker}(f)) = f(v)$ die Umkehrabbildung durch $\overline{f}^{-1}(w) = f^{-1}(w)$ für $w \in \operatorname{Im}(f)$.

Folgerung 3.1.3.

$$\dim(V) = \dim(\operatorname{Ker}(f)) + \dim(\operatorname{Im}(f))$$

Beweis. Nach 2.5.8 ist

$$\dim(V) = \dim(\operatorname{Ker}(f)) + \dim\bigl(V/\operatorname{Ker}(f)\bigr).$$

Aus dem Satz folgt andererseits

$$\dim\bigl(V/\operatorname{Ker}(f)\bigr) = \dim(\operatorname{Im}(f)).$$

□

Folgerung 3.1.4. *Sei $f : V \to W$ eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen.*

1. *Für alle $b \in W$ ist $f^{-1}(b)$ leer oder eine Nebenklasse von* $\operatorname{Ker}(f)$.
2. *Wenn* $\dim(V) > \dim(W)$, *ist* $\operatorname{Ker}(f) \neq \mathbf{0}$. *Wenn* $\dim(V) < \dim(W)$, *ist* $\operatorname{Im}(f) \neq W$.
3. *Wenn* $\dim(V) = \dim(W)$, *so sind äquivalent:*
   (a) $\operatorname{Ker}(f) = \mathbf{0}$
   (b) *f ist injektiv.*
   (c) *f ist surjektiv.*

Beweis.

1: folgt sofort aus der Äquivalenz (3.1) auf Seite 42.

2: Wenn $\dim(V) > \dim(W)$, kann $f : V \to \mathrm{Im}(f)$ kein Isomorphismus sein, weil dann auch $\dim V > \dim(\mathrm{Im}(f))$. f ist also nicht injektiv, d.h. $\mathrm{Ker}(f) \neq 0$.

Wenn $\dim(V) < \dim(W)$ ist auch $\dim(\mathrm{Im}(f)) < \dim(W)$, also $\mathrm{Im}(f) \neq W$

3: f ist genau dann injektiv, wenn der Kern von f Null ist, wegen der Äquivalenz (3.1). Aus 3.1.3 folgt, daß der Kern genau dann Null ist, wenn V und $\mathrm{Im}(f)$ die gleiche Dimension haben. Wenn V die gleiche Dimension wie W hat, bedeutet das $\mathrm{Im}(f) = W$. □

Wendet man diese Folgerung auf $f_A : K^n \to K^m$ an, übersetzt sie sich in:

Folgerung 3.1.5. *Sei A eine m–n–Matrix und H die Menge der Lösungen des homogenen Gleichungssystems $Ax = 0$. Dann gilt*

1. *Für alle Spaltenvektoren b der Länge m ist die Lösungsmenge des Gleichungssystems $Ax = b$ leer oder eine Nebenklasse von H.*
2. *Wenn $m < n$, ist $H \neq \mathbf{0}$. Wenn $n < m$ haben nicht alle Gleichungen $Ax = b$ eine Lösung.*
3. *Wenn $m = n$, so sind äquivalent:*
   - *(a) $H = \mathbf{0}$*
   - *(b) Für alle b hat $Ax = b$ höchstens eine Lösung.*
   - *(c) $Ax = b$ ist lösbar für alle b.* □

Der erste Teil von 2. wurde schon in 1.1.5 bewiesen. Wir geben noch einen Beweis von 3. , der von 3.1.2 unabhängig ist: Daß die Eindeutigkeit der Lösungen von $Ax = b$ zu $H = \mathbf{0}$ äquivalent ist, ist klar, weil sich je zwei Lösungen durch ein Element von H unterscheiden.
$Ax = b$ ist für alle b genau dann lösbar, wenn die Spalten ein Erzeugendensystem des K^n bilden. H ist genau dann Null, wenn die Spalten linear unabhängig sind. Eine n–elementige Menge ist aber genau ein Erzeugendensystem von K^n, wenn sie linear unabhängig ist. (Das folgt aus 2.3.3 und 2.3.6.)

Definition. *Der Rang von f ist die Dimension des Bildes von f. Der Rang einer Matrix ist die Maximalzahl linear unabhängiger Spalten.*

Der hier definierte Rang stimmt mit dem in Abschnitt 1.1 definierten Rang überein. Man vergleiche den Beweis von 1.1.7.

Lemma 3.1.6. *Sei A eine m–n–Matrix und $f_A : K^n \to K^m$ die dadurch definierte lineare Abbildung. Dann ist der Rang von A gleich dem Rang von f_A.*

Beweis. Wenn $a_1, \ldots, a_n$ die Spalten von A sind, ist

$$f_A(\xi_1, \ldots, \xi_n) = \xi_1 a_1 + \cdots + \xi_n a_n.$$

Daran sieht man, daß die Spalten von A ein Erzeugendensystem von $\mathrm{Im}(f)$ bilden. Jede maximale linear unabhängige Teilmenge von $\{a_1, \ldots, a_n\}$ ist eine Basis von $\mathrm{Im}(f)$. □

Bemerkung. *Wenn $v_1, \ldots, v_n$ ein Erzeugendensystem von V ist, ist $f(v_1), \ldots, f(v_n)$ ein Erzeugendensystem von $\mathrm{Im}(f)$. Wenn $f(v_1), \ldots, f(v_n)$ linear unabhängig sind, sind auch $v_1, \ldots, v_n$ linear unabhängig.*

Beweis. Die Elemente $v \in V$ sind Linearkombinationen $\sum_{i=1}^n \alpha_i v_i$. Also sind die Bilder $f(v)$ Linearkombinationen $\sum_{i=1}^n \alpha_i f(v_i)$. Schließlich führt eine lineare Abhängigkeit $\sum_{i=1}^n \alpha_i v_i = 0$ zu einer linearen Abhängigkeit $\sum_{i=1}^n \alpha_i f(v_i)$. □

3.2 Die lineare Gruppe

Ein *Endomorphismus* von V ist eine lineare Abbildung von V nach V. Die Endomorphismen von V bilden eine K–Algebra[1], $\mathrm{End}(V)$.

Ein bijektiver Endomorphismus heißt *Automorphismus*. Ein Automorphismus ist also ein Isomorphismus von V mit sich selbst. Die Automorphismen von V bilden mit der Komposition als Verknüpfung eine Gruppe, die *lineare Gruppe* $\mathrm{Gl}(V)$ mit der identischen Abbildung id_V als neutralem Element[2].

Das folgende Lemma ist eine Umformulierung von 3.1.4.3.

Lemma 3.2.1. *Sei V endlichdimensional und f ein Endomorphismus von V. Dann sind äquivalent:*

1. *f ist ein Automorphismus*
2. *f ist injektiv*
3. *f ist surjektiv* □

Sei $b_1, \ldots, b_n$ eine Basis von V. Dann ist f genau dann injektiv, wenn die Bilder der b_i unter f linear unabhängig sind, und genau dann surjektiv, wenn die Bilder V erzeugen.

[1]Das ist ein K–Vektorraum mit einer bilinearen Operation. $\mathbb{R}$-Algebren wurden schon auf Seite 38 definiert.

[2]$\mathrm{Gl}(V)$ ist die Einheitengruppe von $\mathrm{End}(V)$. Dabei heißt ein Element eines Ringes *Einheit*, wenn es ein Links- und ein Rechtsinverses hat.

Wenn $V = K^n$, ist $\mathrm{End}(V)$ auf natürliche Weise isomorph zur K–Algebra $\mathrm{M}_n(K)$. Eine n–n–Matrix A heißt *regulär* , wenn $f_A : K^n \to K^n$ ein Automorphismus ist, sonst *singulär*. Die Menge $\mathrm{Gl}_n(K)$ der regulären n-n–Matrizen mit der Matrizenmultiplikation als Operation ist eine Gruppe mit der Einheitsmatrix als Einselement. Der Isomorphismus zwischen $\mathrm{M}_n(K)$ und $\mathrm{End}(K^n)$ überführt $\mathrm{Gl}(K^n)$ in $\mathrm{Gl}_n(K)$.

Lemma 3.2.2. *Für eine n–n–Matrix A sind äquivalent:*

1. *A ist regulär*
2. *A hat Rang n.*
3. *A hat ein Linksinverses: $BA = \mathbf{I}$*
4. *A hat ein Rechtsinverses: $AC = \mathbf{I}$*

Beweis. Wenn A regulär ist, ist f_A ein Automorphismus. Weil f_A surjektiv ist, hat f_A und damit auch A den Rang n. Wenn f_B die Umkehrabbildung von f_A ist, ist $BA = AB = \mathbf{I}$. (Natürlich ist B das einzige Links- bzw. Rechtsinverse.)

Wenn A ein Links- oder Rechtsinverses hat, hat f_A ein Links- oder Rechtsinverses und ist damit injektiv oder surjektiv. Aus 3.2.1 folgt, daß f_A ein Automorphismus und A regulär ist. □

Wenn A regulär ist, bezeichnen wir mit A^{-1} das Inverse von A.

BEISPIELE:

- Es ist $\mathrm{Gl}_1(K) = K^\bullet$, die multiplikative Gruppe der von Null verschiedenen Elemente von K.
- $\mathrm{Gl}_n(\mathbb{F}_p)$ hat
$$(p^n - 1)(p^n - p) \dots (p^n - p^{n-1})$$
viele Elemente. Eine n–n–Matrix mit Spalten $a_1, \dots, a_n$ ist nämlich genau dann regulär, wenn die $a_1, \dots, a_n$ linear unabhängig sind. Das heißt, daß
$$a_1 \notin \mathbf{0},\ a_2 \notin \langle a_1 \rangle,\ \dots\ , a_n \notin \langle a_1, \dots, a_{n-1} \rangle.$$
Für a_1 bleiben also $p^n - 1$ viele Möglichkeiten, für a_2 $p^n - p$ viele Möglichkeiten,..., und schließlich $p^n - p^{n-1}$ viele Möglichkeiten für a_n.

Die *Elementarmatrizen*

$$\mathrm{E}_{ij}^{\lambda} \quad (1 \leq i \neq j \leq n,\ \lambda \in K)$$

sind n-n–Matrizen, in deren Diagonalen Einsen stehen und deren andere Einträge Null sind. Nur an der Stelle (i, j) steht λ.

Die *Elementarmatrizen*

$$\mathrm{E}_j^\lambda \quad (1 \leq i \leq n,\ \lambda \in K^\cdot)$$

haben Einsen in der Diagonalen und Nullen sonst. Nur an der Stelle (j,j) steht λ.

Für die Einheitsvektoren $\mathrm{e}_1, \ldots, \mathrm{e}_n$ gilt

$$\mathrm{E}_{ij}^\lambda\, \mathrm{e}_k = \begin{cases} \mathrm{e}_j + \lambda \mathrm{e}_i & \text{wenn } k = j \\ \mathrm{e}_k & \text{sonst} \end{cases}$$

und

$$\mathrm{E}_j^\lambda\, \mathrm{e}_k = \begin{cases} \lambda \mathrm{e}_j & \text{wenn } k = j \\ \mathrm{e}_k & \text{sonst} \end{cases}$$

Damit sieht man sofort:

Folgerung. $\mathrm{E}_{ij}^\lambda A$ *entsteht aus* A *durch Anwenden der Zeilenoperation* Z_{ji}^λ*,* $\mathrm{E}_i^\lambda A$ *durch Anwenden der Zeilenoperation* Z_i^λ*.* □

Elementarmatrizen sind regulär. Es ist $(\mathrm{E}_{ij}^\lambda)^{-1} = \mathrm{E}_{ij}^{-\lambda}$ und $(\mathrm{E}_i^\lambda)^{-1} = \mathrm{E}_i^{\lambda^{-1}}$.

Satz 3.2.3. $\mathrm{Gl}_n(K)$ *wird von Elementarmatrizen erzeugt. Das heißt, daß sich jede reguläre Matrix als Produkt von Elementarmatrizen (und ihren Inversen) schreiben läßt.*

Beweis. Wir haben in 1.1.3 gesehen, daß sich jede Matrix A durch Zeilenoperationen in Normalform N bringen läßt. Es gibt also ein Produkt S von Elementarmatrizen, sodaß $SA = N$. Weil N wie A den Rang n hat, muß $N = \mathbf{I}$ sein. Daraus folgt, daß $A = S^{-1}$ Produkt von Elementarmatrizen ist.

In 1.1.3 mußte man allerdings Spalten vertauschen. Wenn A quadratisch ist und Rang $k = n$ hat, sind Spaltenvertauschungen unnötig. Man sieht das leicht am Beweis von 1.1.3. [3]. Hier ist aber nochmal das Argument:

Nehmen wir an, wir hätten A durch Zeilenoperation umgeformt in eine Matrix A', deren i erste Spalten die Einheitsvektoren $\mathrm{e}_1, \ldots, \mathrm{e}_i$ sind[4]. Sei a die $i+1$–te Spalte von A'. Weil A' regulär ist, ist a nicht linear abhängig von den vorangehenden Spalten. Also muß a unterhalb der i–ten Zeile ein von Null verschiedenes Element enthalten. Weil wir Zeilen vertauschen dürfen, können wir annehmen, daß der $i+1$–te Koeffizient von Null verschieden ist. Durch Anwenden von geeigneten Operationen Z_i^λ und Z_{ij}^λ können wir a in e_{i+1} verwandeln. Es resultiert eine Matrix A'', deren $i+1$ erste Spalten die $i+1$ ersten Einheitsvektoren sind. □

[3] Man muß nämlich die Spalten nur vertauschen, um zu erreichen, daß die ersten k Spalten linear unabhängig sind.

[4] Man beginnt mit $i = 0$

Ein Gleichungssystem $Ax = b$ läßt sich durch Multiplikation mit einer regulären Matrix B passender Größe äquivalent umformen in

$$x = (BA)x = Bb.$$

3.2.3 zeigt, daß man diese Umformungen immer durch Zeilenoperation erzielen kann.

3.3 Basiswechsel

Definition. *Seien $\mathfrak{B} = (b_1, \ldots, b_n)$ und $\mathfrak{B}' = (b'_1, \ldots, b'_n)$ zwei Basen von V. Der* Wechsel *von $\mathfrak{B}$ nach $\mathfrak{B}'$ wird durch die Matrix $B = (\beta_{ij})_{\substack{i=1\ldots n\\ j=1\ldots n}}$ beschrieben, wenn*

$$b'_j = \sum_{i=1}^{n} \beta_{ij} b_i. \tag{3.2}$$

In symbolischer Notation bedeutet das

$$(b'_1 \ldots b'_n) = (b_1 \ldots b_n)B.$$

Wie wir auf Seite 27 gesehen haben, heißt das, daß bezüglich der Basis $\mathfrak{B}$ die Matrix B die lineare Abbildung g beschreibt, die b_i auf b'_i abbildet. Anders ausgedrückt, wenn $\epsilon : K^n \to V$ die kanonische Basis auf $\mathfrak{B}$ abbildet, ist das Diagramm

$$\begin{tikzcd} V \arrow[r, "g"] & V \\ K^n \arrow[u, "\epsilon"] \arrow[r, "f_B"'] & K^n \arrow[u, "\epsilon"'] \end{tikzcd} \tag{3.3}$$

kommutativ. Weil g ein Automorphismus ist, ist B regulär.

Lemma 3.3.1. *Die Koordinaten eines Vektors*

$$\xi_1 b_1 + \cdots + \xi_n b_n = \xi'_1 b'_1 + \cdots + \xi'_n b'_n$$

bezüglich $\mathfrak{B}$ und $\mathfrak{B}'$ gehen vermöge

$$\xi_i = \sum_{j=1}^{n} \beta_{ij} \xi'_j$$

auseinander hervor.

Beweis. Die Behauptung folgt aus

$$\sum_i \xi_i b_i = \sum_j \xi'_j b'_j = \sum_j \xi'_j \Big(\sum_i \beta_{ij} b_i\Big) = \sum_i \Big(\sum_j \beta_{ij} \xi'_j\Big) b_i$$

durch Koeffizientenvergleich. □

In symbolischer Schreibweise sieht der Beweis so aus:

$$(b_1 \dots b_n) \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = (b'_1 \dots b'_n) \begin{pmatrix} \xi'_1 \\ \vdots \\ \xi'_n \end{pmatrix} = (b_1 \dots b_n) B \begin{pmatrix} \xi'_1 \\ \vdots \\ \xi'_n \end{pmatrix}$$

impliziert

$$\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = B \begin{pmatrix} \xi'_1 \\ \vdots \\ \xi'_n \end{pmatrix}.$$

Die Abbildung $\epsilon' = g \circ \epsilon$ bildet die kanonische Basis auf $\mathfrak{B}'$ ab. Aus (3.3) folgt die Kommutativität von

$$\begin{tikzcd} V \arrow[r, "\mathrm{id}_V"] & V \\ K^n \arrow[u, "\epsilon'"] \arrow[r, "f_B"'] & K^n \arrow[u, "\epsilon"'] \end{tikzcd}$$

B beschreibt also bezüglich der Basis $\mathfrak{B}'$ im Definitionsbereich und der Basis $\mathfrak{B}$ im Bildbereich die identische Abbildung von V nach V. Das ist ein Umformulierung von 3.3.1.

Satz 3.3.2. *U,V und W seien endlichdimensionale Vektorräume mit ausgewählten Basen $\mathfrak{A}$, $\mathfrak{B}$ und $\mathfrak{C}$.*

$f : V \to U$ und $h : W \to V$ seien lineare Abbildungen, die Matrizen A und C entsprechen. Durch B sei ein Basiswechsel von $\mathfrak{B}$ nach $\mathfrak{B}'$ gegeben. Dann gilt:

1. *Bezüglich der Basen $\mathfrak{B}'$ und $\mathfrak{A}$ gehört zu f die Matrix*

$$AB.$$

2. *Bezüglich der Basen $\mathfrak{C}$ und $\mathfrak{B}'$ gehört zu h die Matrix*

$$B^{-1}C.$$

Beweis. Sei $v \in V$ und x bzw. x' die Koordinaten von v bezüglich $\mathfrak{B}$ bzw. $\mathfrak{B}'$. Die Koordinaten von $f(v)$ bezüglich $\mathfrak{A}$ sind dann $Ax = A(Bx') = (AB)x'$. Das beweist die erste Behauptung.

Die zweite Behauptung läßt sich ebenso beweisen. Ich gebe aber zur Abwechslung einen anderen Beweis: Wenn wir die Basen als Zeilen von Vektoren auffassen, haben wir $h(\mathfrak{C}) = \mathfrak{B}C$ und $\mathfrak{B}' = \mathfrak{B}B$. Daraus folgt $\mathfrak{B} = \mathfrak{B}'B^{-1}$ und, wie behauptet, $h(\mathfrak{C}) = \mathfrak{B}C = \mathfrak{B}'(B^{-1}C)$. □

Folgerung 3.3.3. *Sei V ein endlichdimensionaler Vektorraum mit ausgezeichneter Basis $\mathfrak{B}$. $f : V \to V$ sei ein durch die Matrix A beschriebener Endomorphismus. Wenn wir mit Hilfe der Matrix B zur Basis $\mathfrak{B}'$ übergehen, wird f durch die Matrix $B^{-1}AB$ beschrieben.* □

Zwei quadratische Matrizen A und A' heißen *ähnlich* oder *konjugiert* , wenn $A' = B^{-1}AB$ für eine reguläre Matrix B.

Satz 3.3.4 (Normalform für lineare Abbildungen)**.** *U und V seien Vektorräume der Dimension m und n, $f : V \to U$ eine lineare Abbildung vom Rang k. Dann kann man Basen in U und V so wählen, daß f durch die m–n–Matrix*

$$\left(\begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array}\right) \tag{3.4}$$

dargestellt wird, wobei I die k–dimensionale Einheitsmatrix ist.

Beweis. Sei V' ein Komplementärraum von $\mathrm{Ker}(f)$ in V. Nach 2.5.7 und 3.1.2 induziert f eine Isomorphismus von V' mit $\mathrm{Im}(f)$. V' hat also Rang k. Wir wählen als Basis von V eine Basis $b_1, \dots, b_k$ von V' gefolgt von einer Basis $b_{k+1}, \dots, b_n$ von $\mathrm{Ker}(f)$. Auf der Bildseite setzen wir $f(b_1), \dots, f(b_k)$ zu einer Basis von U fort. □

Folgerung 3.3.5. *Eine m–n–Matrix A hat genau dann den Rang k, wenn es eine reguläre n–n–Matrix B und eine reguläre m–m–Matrix B' gibt, sodaß $B'AB$ die Gestalt* (3.4) *hat.*

Beweis. Satz 3.3.4 auf $f_A : K^n \to K^m$ angewendet liefert Basen in K^n und K^m, bezüglich der f_A die gewünschte Matrixdarstellung A' hat. Wenn die Matrizen B und C die beiden Basiswechsel (von den kanonischen Basen zu den neuen Basen) beschreiben, ist $A' = C^{-1}AB$. Setze $B' = C^{-1}$.

Umgekehrt hat eine Matrix der Form (3.4) Rang k. Damit haben auch f_A und A den Rang k. □

Folgerung 3.3.6. *Jede m–n–Matrix vom Rang k läßt sich durch Zeilen– und Spaltenoperationen in die Gestalt* (3.4) *bringen.*

Beweis. Sei $A' = B'AB$ für reguläre B' und B. Weil B' ein Produkt von Elementarmatrizen ist, läßt sich A' aus AB durch Zeilenoperationen gewinnen.

Es ist leicht zu sehen, daß Rechtsmultiplikation mit Elementarmatrizen die Ausführung einer Spaltenoperation bedeutet: $A\mathrm{E}_{ij}^{\lambda}$ addiert das λ–fache der j–ten Spalte von A zur i–ten Spalte, $A\mathrm{E}_{i}^{\lambda}$ multipliziert die i–te Spalte mit λ.

Daraus folgt, daß AB aus A durch Spaltenoperationen hervorgeht. □

Satz 3.3.7. *U und V seien Vektorräume der Dimension m und n, $f : V \to U$ eine lineare Abbildung vom Rang k. In V sei eine Basis $\mathfrak{B}$ fixiert. Dann kann man $\mathfrak{B}$ so zu einer Basis $\mathfrak{B}'$ umordnen und eine Basis $\mathfrak{A}$ von U so wählen, daß f bezüglich der neuen Basen durch eine Matrix in* Normalform

$$\left(\begin{array}{c|c} I & * \\ \hline 0 & 0 \end{array}\right) \tag{3.5}$$

dargestellt wird, wobei I die k–dimensionale Einheitsmatrix und $*$ eine beliebige k–$(n-k)$–Matrix ist.

Beweis. Der Beweis ist eine Variante des Beweises von 3.3.4. Wir wählen die Umordnung $\mathfrak{B}'$ so, daß $f(b_1'), \ldots, f(b_k')$ eine Basis von $\mathrm{Im}(f)$ ist. $\mathfrak{A}$ ist irgendeine Fortsetzung von $f(b_1'), \ldots, f(b_k')$ zu einer Basis von U. □

Als Folgerung erhält man noch einmal Satz 1.1.3: Jede Matrix vom Rang k läßt sich durch Zeilenoperationen und Umordnen der Spalten in eine Matrix in Normalform vom Rang k bringen.

Jede Normalform läßt sich durch Addieren von geeigneten Vielfachen der ersten k Spalten zu den letzten $n-k$ Spalten in die Form (3.4) bringen. Damit hat man einen alternativen Beweis von 3.3.4.

Kapitel 4

Determinanten

4.1 Die Signatur einer Permutation

X sei eine endliche Menge. Wir wollen jeder Permutation $\pi \in \mathrm{Sym}(X)$ ein Vorzeichen zuordnen, ihre *Signatur* $\mathrm{sign}(\pi)$. Dazu wählen wir eine *Orientierung* der zweielementigen Teilmengen von X. Das heißt eine Abbildung $s : X^2 \setminus \Delta \to \{1, -1\}$, mit[1]

$$s(x, y) = -s(y, x)$$

für alle $x \neq y \in X$, und definieren

Definition.

$$\mathrm{sign}(\pi) = \prod \frac{s(\pi x, \pi y)}{s(x, y)}.$$

Dabei wird das Produkt gebildet über alle *ungeordneten* Paare[2] x, y. Auf die Wahl einer Reihenfolge von x und y kommt es nicht an, weil $\dfrac{s(\pi x, \pi y)}{s(x, y)} = \dfrac{s(\pi y, \pi x)}{s(y, x)}$.

Wenn t eine andere Orientierung ist, die sich von s bei n ungeordneten Paaren unterscheidet, ist

$$\prod \frac{t(x, y)}{s(x, y)} = \prod \frac{t(\pi x, \pi y)}{s(\pi x, \pi y)} = (-1)^n.$$

Daraus folgt, daß es bei der Definition der Signatur auf die Wahl der Orientierung nicht ankommt. Zu Beispiel liefert jede lineare Ordnung $<$ von X eine

[1] $\Delta = \{(x, x) \mid x \in X\}$ ist die *Diagonale* von X

[2] Das heißt über alle Zweiermengen $\{x, y\}$ mit $x \neq y$.

Orientierung, definiert durch

$$\operatorname{sign}(x,y) = \begin{cases} 1 & \text{wenn } x < y \\ -1 & \text{sonst} \end{cases}$$

Man nennt eine zweielementige Menge $\{x, y\}$ einen *Fehlstand* von π, wenn ihre Ordnung von π umgedreht wird, wenn also z.B. $x < y$ und $\pi(x) > \pi(y)$. Man erhält sofort die äquivalente Definition

Definition (Variante)**.**

$$\operatorname{sign}(\pi) = (-1)^{\textit{Zahl der Fehlstände}}$$

Man nennt eine Permutation *gerade*, wenn ihre Signatur 1 ist, und *ungerade*, wenn ihre Signatur -1 ist. Eine Permutation ist also genau dann gerade, wenn sie eine gerade Zahl von Fehlständen hat.

Satz 4.1.1. *Die Signaturabbildung*

$$\operatorname{sign} : \operatorname{Sym}(X) \to \{1, -1\}$$

ist ein Gruppenhomomorphismus.

Beweis.

$$\begin{aligned} \operatorname{sign}(\pi\sigma) &= \prod \frac{s(\pi\sigma x, \pi\sigma y)}{s(x,y)} \\ &= \prod \frac{s(\pi\sigma x, \pi\sigma y)}{s(\sigma x, \sigma y)} \prod \frac{s(\sigma x, \sigma y)}{s(x,y)} \\ &= \operatorname{sign}(\pi)\operatorname{sign}(\sigma) \end{aligned}$$

□

Eine Folge $x_1, \dots, x_k$ von paarweise verschiedenen Elementen von X definiert einen *Zyklus*

$$(x_1 \dots x_k).$$

Das ist die Permutation, die die Elemente x_i (für $i < k$) auf x_{i+1} abbildet, x_k auf x_1 und alle anderen Elemente von X auf sich selbst.

Bemerkung. *Ein Zyklus der Länge k hat die Signatur* $(-1)^{k-1}$

Beweis. Wir ordnen X so, daß $x_1, \dots, x_k$ die ersten k Elemente sind. Die Fehlstände von $(x_1 \dots x_k)$ sind dann die Paare $\{x_1, x_k\}, \{x_2, x_k\}, \dots, \{x_{k-1}, x_k\}$. □

Lemma 4.1.2. *Jede Permutation läßt sich (bis auf die Reihenfolge) eindeutig als Produkt von disjunkten Zyklen schreiben.*

Beweis. Sei $\pi \in \mathrm{Sym}(X)$, Wir zerlegen X mit der Äquivalenzrelation

$$x \sim y \Leftrightarrow \exists z \in \mathbb{Z} \;\; \pi^z x = y$$

in die *Bahnen von* π.

Sei x ein beliebiges Element von X. Die $x, \pi x, \pi^2 x, \ldots$ können nicht alle verschieden sein. Sei n die kleinste positive natürliche Zahl mit $\pi^n x \in \{x, \ldots, \pi^{n-1} x\}$. Dann ist $\pi^n x = x$ und

$$B = \{x, \pi x, \ldots, \pi^{n-1} x\}$$

ist eine Bahn von π. Auf B operiert π wie der Zyklus $\sigma_B = (x \; \pi x \ldots \pi^{n-1} x)$ und es ist

$$\pi = \prod_{B \text{ Bahn}} \sigma_B$$

. □

Ein Zyklus der Länge zwei heißt *Transposition*.

Folgerung 4.1.3. *Jede Permutation ist Produkt von Transpositionen.*

Beweis.

$$(x_1 \ldots x_k) = (x_k x_1)(x_{k-1} x_1) \ldots (x_2 x_1)$$

□

Satz 4.1.4. sign *ist der einzige nicht–triviale Homomorphismus*

$$\mathrm{Sym}(X) \to \{1, -1\}.$$

Beweis. Wir beginnen mit einer Vorüberlegung. Betrachte zwei Transpositionen (ab) und $(a'b')$. Wähle eine Permutation π mit $\pi(a') = a$ und $\pi(b') = b$. Dann ist $(a'b') = \pi^{-1}(ab)\pi$.[3] Wenn nun $S : \mathrm{Sym}(X) \to \{1, -1\}$ ein Homomorphismus ist, ist

$$S(a'b') = S(\pi)^{-1} S(ab) S(\pi) = S(ab).$$

Es gibt also zwei Fälle:

Fall 1: $S(\tau) = 1$ für alle Transpositionen τ.
Dann hat S auch auf allen Produkten von Transpositionen den Wert 1. Also ist $S(\sigma) = 1$ für alle σ.

Fall 2: $S(\tau) = -1$ für alle Transpositionen τ. Wenn σ Produkt von n Transpositionen ist, gilt daher

$$S(\sigma) = (-1)^n = \mathrm{sign}(\sigma).$$

□

[3] Wir sagen dazu: (ab) und $(a'b')$ sind *konjugiert*.

4.2 k–Formen

Sei V ein K–Vektorraum.

Definition. *Eine k–stellige multilineare Abbildung $\mu : V^k \to K$ (eine* Multilinearform*) heißt alternierend oder k–Form, wenn sie die zwei folgenden äquivalenten Bedingungen erfüllt:*

a) $\mu(a_1, \ldots, a_k) = 0$ für alle k–Tupel $a_1, \ldots, a_k$, in denen ein Vektor zweimal vorkommt.

b) Für alle $a_1, \ldots, a_k \in V$, alle $1 \leq i \neq j \leq n$ und alle $\lambda \in K$ ist

$$\mu(a_1, \ldots, a_i + \lambda a_j, \ldots, a_k) = \mu(a_1, \ldots, a_i, \ldots, a_k)$$

Die Äquivalenz der beiden Bedingungen folgt aus

$$\mu(\ldots, a_i + \lambda a_j, \ldots, a_j, \ldots) = \mu(\ldots, a_i, \ldots, a_j, \ldots) + \lambda\mu(\ldots, a_j, \ldots, a_j, \ldots).$$

Lemma 4.2.1. *Eine Multilinearform μ ist genau dann alternierend, wenn $\mu(a_1, \ldots, a_k) = 0$ für alle linear abhängigen Tupel $a_1, \ldots, a_k$.*

Beweis. Sei μ alternierend. Aus Bedingung b) der Definition folgt

$$0 = \mu(a_1, \ldots, a_{k-1}, 0) = \mu(a_1, \ldots, a_{k-1}, \lambda_1 a_1 + \cdots + \lambda_{k-1} a_{k-1}).$$

Wenn $a_1, \ldots, a_k$ linear abhängig sind, weil zum Beispiel $\lambda_1 a_1 + \cdots + \lambda_{k-1} a_{k-1} = a_k$, ist daher $0 = \mu(a_1, \ldots, a_k)$.

Die Umkehrung ist klar. □

Das folgende Lemma ist leicht zu beweisen:

Lemma 4.2.2.

1. *Jede Linearkombination von k–Formen ist wieder eine k–Form.*
2. *Sei $\phi : V \to U$ eine lineare Abbildung und μ eine k–Form auf U. Dann ist μ^ϕ, definiert durch*

$$\mu^\phi(a_1, \ldots, a_k) = \mu(\phi(a_1), \ldots, \phi(a_k))$$

eine k–Form auf V. □

Man bezeichnet mit $\mathrm{A}^k V$ den Raum aller k–Formen auf V.

Lemma 4.2.3. *Sei μ eine k–Form auf V. Dann ist für alle $a_1, \ldots, a_k \in V$, und alle $1 \leq i < j \leq k$*

$$\mu(a_1, \ldots, a_i, \ldots, a_j \ldots, a_k) = -\mu(a_1, \ldots, a_j, \ldots, a_i \ldots, a_k)$$

Beweis. Wir wenden den Trick an, mit der wir im Beweis von Satz 1.1.3 gezeigt haben, daß man mit Zeilenoperationen Zeilen vertauschen kann:

$$\begin{aligned}\mu(\ldots, a_i, \ldots, a_j, \ldots) &= \mu(\ldots, a_i, \ldots, a_i + a_j, \ldots)\\ &= \mu(\ldots, -a_j, \ldots, a_i + a_j, \ldots)\\ &= \mu(\ldots, -a_j, \ldots, a_i, \ldots)\\ &= -\mu(\ldots, a_j, \ldots, a_i, \ldots)\end{aligned}$$

□

Wenn K nicht die Charakteristik 2 hat[4], ist jede Multilinearform mit der im Lemma gegebenen Eigenschaft alternierend. Man schließt nämlich so:

$$\mu(\ldots, a_i, \ldots, a_i, \ldots) = -\mu(\ldots, a_i, \ldots, a_i, \ldots) \Rightarrow \mu(\ldots, a_i, \ldots, a_i, \ldots) = 0$$

Für die nächste Folgerung erweitern wir den Begriff der Signatur auf beliebige Abbildungen $\tau : \{1, \ldots, k\} \to \{1, \ldots, k\}$: Die Signatur $\operatorname{sign}(\tau)$ von τ ist 0, wenn τ keine Permutation ist. Satz 4.1.1 gilt immer noch in der Form

$$\operatorname{sign}(\pi \circ \sigma) = \operatorname{sign}(\pi) \operatorname{sign}(\sigma)$$

für beliebige π und σ.

Folgerung 4.2.4. *Sei μ eine k–Form auf V Dann ist für jede Abbildung $\tau : \{1, \ldots, k\} \to \{1, \ldots, k\}$ und alle $a_1, \ldots, a_k \in V$*

$$\mu(a_{\tau(1)}, \ldots, a_{\tau(k)}) = \operatorname{sign}(\tau)\, \mu(a_1, \ldots, a_k).$$

Beweis. Wenn τ keine Permutation ist, gibt es $i < j$ mit $a_{\tau(i)} = a_{\tau(j)}$. Also ist $\mu(a_{\tau(1)}, \ldots, a_{\tau(k)}) = 0$. Wenn τ eine Permutation ist und Produkt von n Transpositionen, ist

$$\mu(a_{\tau(1)}, \ldots, a_{\tau(k)}) = (-1)^n \mu(a_1, \ldots, a_k) = \operatorname{sign}(\tau)\, \mu(a_1, \ldots, a_k).$$

□

Folgerung 4.2.5. *Sei $a_1, \ldots, a_n$ eine Basis von V. Dann ist μ bestimmt durch die Werte $\mu(a_{i_1}, \ldots, a_{i_k})$ für alle $1 \leq i_1 < \ldots < i_k \leq n$.*

[4]Sei K ein Körper und p die kleinste positive Zahl, sodaß in K

$$\underbrace{1 + 1 + \cdots + 1}_{p} = 0.$$

p muß eine Primzahl sein. (Siehe Fußnote 9, Seite 38). p heißt die Charakteristik von K. Wenn p nicht existiert hat K die Charakteristik Null.

Beweis. Wie in 2.6.1 sieht man, daß eine Multilinearform bestimmt ist durch ihre Werte auf der Basis $a_1, \ldots, a_n$, daß heißt durch die

$$\alpha_j = \mu(a_{j_1}, \ldots, a_{j_k}).$$

Wenn die $j_1, \ldots, j_k$ alle verschieden sind, ordnen wir sie aufsteigend als $i_1, \ldots, i_k$. Dann ist $\alpha_j = \mathrm{sign}(\pi)\mu(a_{i_1}, \ldots, a_{i_k})$, wenn π die Permutation bezeichnet, die $i_1, \ldots, i_k$ auf $j_1, \ldots, j_k$ abbildet.

Wenn die $j_1, \ldots, j_k$ nicht paarweise verschieden sind, ist $\alpha_j = 0$. □

Folgerung 4.2.6. *Wenn $n < k$, ist 0 die einzige k–Form auf V.*

4.3 Determinanten

Satz 4.3.1. *Sei $b_1, \ldots, b_n$ eine Basis von V und $\beta \in K$. Dann gibt es genau eine n–Form μ mit $\mu(b_1, \ldots, b_n) = \beta$.*

Beweis. Die Eindeutigkeit ergibt sich aus 4.2.5. Für die Existenz definieren wir μ als die Multilinearform, die auf den Basiselementen die richtigen Werte hat, für die also

$$\mu(b_{\tau(1)}, \ldots, b_{\tau(n)}) = \mathrm{sign}(\tau)\beta$$

für alle $\tau : \{1, \ldots, n\} \to \{1, \ldots, n\}$ (siehe 2.6.1).

Betrachte für ein festgehaltenes $\pi \in \mathrm{S}_n$ die Multilinearform

$$\nu(x_1, \ldots, x_n) = \mu(x_{\pi(1)}, \ldots, x_\pi(n)).$$

Auf den Basiselementen nimmt ν die gleichen Werte an wie $\mathrm{sign}(\pi)\mu$, weil

$$\begin{aligned}
\nu(b_{\tau(1)}, \ldots, b_{\tau(n)}) &= \mu(b_{\tau(\pi(1))}, \ldots, b_{\tau(\pi(n))}) \\
&= \mathrm{sign}(\tau\pi)\beta \\
&= \mathrm{sign}(\pi)\,\mathrm{sign}(\tau)\beta \\
&= \mathrm{sign}(\pi)\mu(b_{\tau(1)}, \ldots, b_{\tau(n)})
\end{aligned}$$

Es folgt, daß

$$\mu(x_{\pi(1)}, \ldots, x_{\pi(n)}) = \mathrm{sign}(\pi)\mu(x_1, \ldots, x_n).$$

Wenn K nicht die Charakteristik 2 hat, wissen wir jetzt, daß μ alternierend ist und sind fertig. Im allgemeinen Fall schließt man so: Wir müssen zeigen, daß μ Null wird, wenn zwei Argumente gleich sind. Um die Notation zu vereinfachen, nehmen wir an, daß es sich um die ersten beiden Argumente handelt. Es ist also zu zeigen, daß für alle a die $(n-2)$–Form $\mu(a, a, x_3, \ldots)$ auf den Basiselementen verschwindet. Das heißt, daß für alle $i_3, \ldots, i_n \in \{1, \ldots, n\}$

$$\mu(a, a, b_{i_3}, \ldots, b_{i_n}) = \mu(a, a, \ldots) = 0.$$

Für $a = \xi_1 b_1 + \cdots + \xi_n b_n$ rechnet man:

$$\begin{aligned}\mu(a, a, \ldots) &= \sum_i \xi_i^2 \mu(b_i, b_i, \ldots) \\ &\quad + \sum_{i<j} \xi_i \xi_j (\mu(b_i, b_j, \ldots) + \mu(b_j, b_i, \ldots)) \\ &= 0.\end{aligned}$$

Denn $\mu(b_i, b_i, \ldots) = 0$ und $\mu(b_i, b_j, \ldots) = -\mu(b_j, b_i, \ldots)$. □

Folgerung 4.3.2. *Sei V n–dimensional und μ eine nicht–triviale n–Form auf V. Dann ist jede andere n–Form ein Vielfaches von μ.*

Beweis. Sei $b_1, \ldots, b_n$ eine Basis. Weil μ nicht Null ist, ist $\beta = \mu(b_1, \ldots, b_n)$ von Null verschieden. Sei μ' eine andere n–Form und $\beta' = \mu'(b_1, \ldots, b_n)$. Dann ist $\mu' = \frac{\beta'}{\beta}\mu$. □

Sei $\mathrm{e}_1, \ldots, \mathrm{e}_n$ die kanonische Basis des K^n und μ_0 die n–Form, für die $\mu_0(\mathrm{e}_1, \ldots, \mathrm{e}_n) = 1$. Wir nennen μ_0 die *Standard n–Form* des K^n.

Die *Determinante* einer n–n–Matrix A ist

$$\det(A) = \mu_0(a_1, \ldots, a_n),$$

wobei $a_1, \ldots, a_n$ die Spalten von A sind. Anders ausgedrückt:

Definition. *Die Determinante* $\det(A)$ *ist eine alternierende Form in den Spalten von A, festgelegt durch* $\det(\mathbf{I}) = 1$.

Man schreibt auch $|A|$ für die Determinante von A.

Lemma 4.3.3. *Wenn $A = (\alpha_{ij})_{i,j=1\ldots n}$, ist*

$$\det(A) = \sum_{\pi \in \mathrm{S}_n} \operatorname{sign}(\pi) \prod_{j=1}^{n} \alpha_{\pi(j),j}. \tag{4.1}$$

Beweis. Die Formel, nur eine explizite Version von 4.2.5 für $k = n$, wird einfach nachgerechnet: (Dabei läuft τ über alle Abbildungen $\{1, \ldots, n\} \to \{1, \ldots, n\}$

und π über alle Permutationen.)

$$\begin{aligned}
\det(A) &= \mu_0\big(\sum_i \alpha_{i1}\mathrm{e}_i, \ldots, \sum_i \alpha_{in}\mathrm{e}_i\big) \\
&= \sum_\tau \mu_0(\alpha_{\tau(1)1}\mathrm{e}_{\tau(1)}, \ldots, \alpha_{\tau(n)n}\mathrm{e}_{\tau(n)}) \\
&= \sum_\tau \big(\prod_{j=1}^n \alpha_{\tau(j),j}\big)\mu_0(\mathrm{e}_{\tau(1)}, \ldots, \mathrm{e}_{\tau(n)}) \\
&= \sum_\tau \operatorname{sign}(\tau) \prod_{j=1}^n \alpha_{\tau(j),j} \\
&= \sum_{\pi \in \mathrm{S}_n} \operatorname{sign}(\pi) \prod_{j=1}^n \alpha_{\pi(j),j}
\end{aligned}$$

□

Wir rechnen einige Beispiele aus. Für $n = 1, 2, 3$ haben wir:

$$\begin{aligned}
\det(a) &= a \\
\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} &= a_1b_2 - a_2b_1 \\
\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} &= a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_2b_1c_3 - a_1b_3c_2 - a_3b_2c_1
\end{aligned}$$

Die Determinante einer *oberen Dreiecksmatrix* ist das Produkt ihrer Diagonalelemente,

$$\begin{vmatrix} \lambda_1 & & & \\ 0 & \lambda_2 & * & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \ldots & \lambda_n \end{vmatrix} = \lambda_1\lambda_2\ldots\lambda_n$$

wie man mit Hilfe von (4.1) leicht sieht. Insbesondere gilt für die Elementarmatrizen:

$$\begin{aligned}
\det(E_i^\lambda) &= \lambda \\
\det(E_{ij}^\lambda) &= 1
\end{aligned}$$

Eine Variante davon ist die folgenden Formel. Sei A eine m–m–Matrix, B eine m–n–Matrix und C eine n–n–Matrix. Dann ist

$$\begin{vmatrix} A & B \\ \mathbf{0} & C \end{vmatrix} = |A| \cdot |C|. \tag{4.2}$$

Das ist mit Gleichung (4.1) auf Seite 58 leicht nachzurechnen. Man sieht, daß man sich auf $\pi \in \mathrm{S}_{m+n}$ beschränken kann, die $\{1,\dots,m\}$ (und daher auch $\{m+1,\dots,m+n\}$) invariant lassen. Weiterhin ist klar, daß die Signatur einer solchen Permutation das Produkt der Signaturen ihrer Einschränkungen auf $\{1,\dots,m\}$ und $\{m+1,\dots,m+n\}$ ist.

Satz 4.3.4.

1. $\det(AB) = \det(A)\det(B)$
2. *Die Determinante von A ist genau dann 0, wenn A singulär ist.*

Beweis.
1: Wenn $b_1,\dots,b_n$ die Spalten von B sind, ist (für festes A)

$$\mu(b_1,\dots,b_n) = \det(Ab_1 \cdots Ab_n) = \det(AB)$$

eine n–Form auf K^n (siehe 4.2.2). Weil $\mu(\mathrm{e}_1,\dots,\mathrm{e}_n) = \det(A)$, ist μ das $\det(A)$–fache der Standardform:

$$\mu(b_1,\dots,b_n) = \det(A)\cdot\mu_0(b_1,\dots,b_n) = \det(A)\det(B).$$

2: Wenn A regulär ist, ist $\det(A)\cdot\det(A^{-1}) = 1$. $\det(A)$ kann also nicht Null sein. Wenn A singulär ist, sind die Spalten von A linear abhängig und $\det(A)$ ist Null nach 4.2.1. $\square$

Die Determinante ist ein Homomorphismus

$$\det : \mathrm{Gl}_n(K) \to K^\bullet.$$

Der Kern dieser Abbildung ist die *spezielle* lineare Gruppe

$$\mathrm{Sl}_n(K) = \{A \in \mathrm{M}_n(K) \mid \det(A) = 1\}.$$

Bemerkung. $\mathrm{Sl}_n(K)$ *wird erzeugt von den Elementarmatrizen* $\mathrm{E}_{ij}^{\lambda}$.

Beweis. Der Beweis von 3.2.3 zeigt, daß man jede reguläre Matrix mit Hilfe von Zeilenoperationen $\mathrm{Z}_{ii'}^{\lambda}$ in eine Diagonalmatrix (mit gleicher Determinante) umformen kann. Es genügt also zu zeigen, daß sich jede Diagonalmatrix D aus $\mathrm{Sl}_n(K)$ mit Hilfe der $\mathrm{Z}_{ii'}^{\lambda}$ in die Einheitsmatrix umformen läßt. Wir können die Diagonalelemente von D sukzessive durch 1 ersetzen, wenn uns die Umformung

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 \\ 0 & ab \end{pmatrix}$$

gelingt. Aber das geht so:

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \xrightarrow{\mathrm{Z}_{12}^{1/a}} \begin{pmatrix} a & 0 \\ 1 & b \end{pmatrix} \xrightarrow{\mathrm{Z}_{21}^{1-a}} \begin{pmatrix} 1 & b-ab \\ 1 & b \end{pmatrix} \xrightarrow{\mathrm{Z}_{12}^{-1}} \begin{pmatrix} 1 & b-ab \\ 0 & ab \end{pmatrix} \xrightarrow{\mathrm{Z}_{21}^{1-1/a}} \begin{pmatrix} 1 & 0 \\ 0 & ab \end{pmatrix}$$

$\square$

Definition. *Die Transponierte einer m–n–Matrix $A = (\alpha_{ij})_{\substack{i=1\dots m\\ j=1\dots n}}$ ist die n–m–Matrix*

$$A^\top = (\alpha_{ij})_{\substack{j=1\dots n\\ i=1\dots m}}$$

Die Spalten von $A^\top$ sind also die Zeilen von A. Man rechnet leicht nach, daß

$$(AB)^\top = B^\top A^\top.$$

Die beiden Produkte AB und $B^\top A^\top$ werden gleich ausgerechnet. Die Resultate werden nur anders indiziert. Wenn $a_1, \dots, a_l$ die Zeilen von A sind und $b_1, \dots, b_n$ die Spalten von B (allen von der Länge m), ist in symbolischer Schreibweise

$$AB = \begin{pmatrix} a_1 \\ \vdots \\ a_l \end{pmatrix} (b_1 \cdots b_n)$$

und daher

$$(AB)^\top = \begin{pmatrix} b_1^\top \\ \vdots \\ b_n^\top \end{pmatrix} (a_1^\top \cdots a_l^\top) = B^\top A^\top.$$

Satz 4.3.5.

1. $\det(A^\top) = \det(A)$
2. $\det$ *ist eine alternierende Multilinearform der Zeilen von A.*

Beweis. Weil $\prod_{i=1}^n \alpha_{i,\pi(i)} = \prod_{j=1}^n \alpha_{\pi^{-1}(j),j}$, ist

$$\begin{aligned}
\det(A^\top) &= \sum_{\pi \in \mathrm{S}_n} \mathrm{sign}(\pi) \prod_{i=1}^n \alpha_{i,\pi(i)} \\
&= \sum_{\pi \in \mathrm{S}_n} \mathrm{sign}(\pi) \prod_{j=1}^n \alpha_{\pi^{-1}(j),j} \\
&= \sum_{\pi \in \mathrm{S}_n} \mathrm{sign}(\pi) \prod_{j=1}^n \alpha_{\pi(j),j} \\
&= \det(A)
\end{aligned}$$

Es folgt, daß $\det(A)$ eine alternierende Multilinearform der Spalten von $A^\top$, und damit der Zeilen von A, ist. $\square$

4.4 Der Laplacesche Entwicklungssatz

Satz 4.4.1 (Die Cramersche Regel)**.** *Sei $Ax = b$ ein quadratisches Gleichungssystem, $a_1, \ldots, a_n$ die Spalten von A. Wenn A regulär ist, errechnet sich der Lösungsvektor $x_1, \ldots, x_n$ als*

$$x_j = \frac{\det(a_1 \cdots a_{j-1} b a_{j+1} \cdots a_n)}{\det(A)}.$$

Beweis. Die Gleichung $Ax = b$ bedeutet $x_1 a_1 + \cdots + x_n a_n = b$. Dann ist

$$\begin{aligned}
\det(a_1 \cdots a_{j-1} b a_{j+1} \cdots a_n) = \\
&= \det(a_1 \cdots \sum_i x_i a_i \cdots a_n) \\
&= \sum_i x_i \det(a_1 \cdots a_i \cdots a_n) \\
&= x_j \det(a_1 \cdots a_j \cdots a_n) \\
&= x_j \det(A).
\end{aligned}$$

□

<u>Notation</u>
Für eine n–n–Matrix A und zwei Indizes i und j bezeichnen wir mit A_{ij} die Matrix, die aus A entsteht, wenn man die i–te Zeile und die j–te Spalte streicht.

Satz 4.4.2 (Entwicklungssatz von Laplace)**.**
Sei $A = (\alpha_{ij})_{\substack{i=1\ldots n \\ j=1\ldots n}}$ eine n–n–Matrix und j_0 ein Spaltenindex. Dann ist

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j_0} \alpha_{ij_0} \det(A_{ij_0}).$$

Die Determinante wird hier nach der j_0–ten Spalte *entwickelt.*

Beweis. Wir schreiben

$$\det(A) = \left| a_1 \cdots a_n \right|,$$

wobei die a_j die Spalten von A sind. Weil $a_{j_0} = \sum_i \alpha_{ij_0} \mathrm{e}_i$, ist

$$\left| a_1 \cdots a_{j_0} \cdots a_n \right| = \sum_i \alpha_{ij_0} \left| a_1 \cdots \mathrm{e}_i \cdots a_n \right|.$$

Wenn wir in der Matrix $\big(a_1 \cdots \mathrm{e}_i \cdots a_n\big)$ die j_0–te Spalte nach vorn schieben und die i–te Zeile nach oben, erhalten wir eine Matrix der Form

$$\begin{pmatrix} 1 & * \\ \mathbf{0} & A_{ij_0} \end{pmatrix}.$$

Die verwendeten Permutationen $(1 \cdots j_0)$ und $(1 \cdots i)$ haben die Signaturen $(-1)^{j_0-1}$ und $(-1)^{i-1}$. Also ist (wegen 4.3.5)

$$\big|a_1 \cdots \mathrm{e}_i \cdots a_n\big| = (-1)^{(i-1)+(j_0-1)} \begin{vmatrix} 1 & * \\ \mathbf{0} & A_{ij_0} \end{vmatrix} = (-1)^{i+j_0} |A_{ij_0}|.$$

Daraus ergibt sich die Behauptung. □

Definiere die *Adjunkte* $\mathrm{adj}(A) = (\gamma_{ij})$ einer n–n–Matrix A durch

$$\gamma_{ij} = (-1)^{i+j} \det(A_{ji}).$$

Sei c_j die j–te Zeile von $\mathrm{adj}(A)$. Aus dem Laplaceschen Entwicklungssatz folgt dann für alle Spaltenvektoren b

$$c_j b = \big|a_1 \cdots a_{j-1} b a_{j+1} \cdots a_n\big|.$$

Also ist

$$c_j a_i = \begin{cases} \det(A), & (j = i) \\ 0, & (j \neq i) \end{cases}$$

Das ergibt

Satz 4.4.3. $\mathrm{adj}(A)\ A = \det(A)\ \mathbf{I}$ □

Wenn A regulär ist, ist also

$$A^{-1} = \frac{\mathrm{adj}(A)}{\det(A)}.$$

Die Cramersche Regel folgt hieraus. Denn, wenn $Ax = b$, ist $x = A^{-1}b$. Das heißt

$$x_j = \frac{c_j b}{\det(A)} = \frac{\det(a_1 \cdots b \cdots a_n)}{\det(A)}.$$

Weil das Linksinverse von A auch Rechtsinverses ist, ist für reguläre A auch $A\ \mathrm{adj}(A) = \det(A)\ \mathbf{I}$. Diese Gleichung gilt beliebige A:

Bemerkung. $A\ \mathrm{adj}(A) = \det(A)\ \mathbf{I}$

Beweis. Man überlegt zuerst, daß $\mathrm{adj}(A)^\top = \mathrm{adj}(A^\top)$. Dann folgt

$$A\ \mathrm{adj}(A) = \big(\mathrm{adj}(A)^\top A^\top\big)^\top = \big(\mathrm{adj}(A^\top)\ A^\top\big)^\top = \big(\det(A^\top)\ \mathbf{I}\big)^\top = \det(A)\ \mathbf{I}$$

□

4.5 Geometrische Interpretation

Orientierung

Wie in Abschnitt 4.3 bezeichnen wir mit μ_0 die Standard n–Form des $\mathbb{R}^n$.

Definition. *Eine Basis $\mathfrak{B}$ des $\mathbb{R}^n$ heißt positiv orientiert, wenn $\mu_0(\mathfrak{B}) > 0$.*

Wenn wir eine Basis $\mathfrak{B}_0$ *stetig* in eine Basis $\mathfrak{B}_1$ überführen, ändert sich die Orientierung nicht. Wenn nämlich für eine stetige Funktion

$$B : [0,1] \to \{\text{Basen}\}$$

$B(0) = \mathfrak{B}_0$ und $B(1) = \mathfrak{B}_1$, ist $\mu_0(B(t))$ ist eine stetige Funktion von t, die nirgends Null wird. Also haben $\mu_0(\mathfrak{B}_0)$ und $\mu_0(\mathfrak{B}_1)$ dasselbe Vorzeichen.

Wenn V ein beliebiger n–dimensionaler $\mathbb{R}$–Vektorraum ist, definiert jede nicht–triviale n–Form μ eine Orientierung von V durch die Maßgabe

$$\mathfrak{B} \text{ positiv orientiert} \iff \mu(\mathfrak{B}) > 0.$$

Eine andere nicht–triviale n–Form, $\mu' = \lambda\mu$, definiert dieselbe Orientierung, wenn $\lambda > 0$ und die entgegengesetzte Orientierung, wenn $\lambda < 0$. V hat also genau zwei Orientierungen.

Volumen

Für Vektoren $a_1, \ldots, a_n$ des $\mathbb{R}^n$ sei $\mathrm{vol}(a_1, \ldots, a_n)$ das in der Analysis (oder für $n = 1, 2, 3$ in der Schulgeometrie) definierte Volumen des Parallelepipeds

$$\mathrm{PE}(a_1, \ldots, a_n) = \{\lambda_1 a_1 + \cdots + \lambda_n a_n \mid 0 \leq \lambda_1 \leq 1, \ldots, 0 \leq \lambda_n \leq 1\}.$$

Satz 4.5.1. $\mathrm{vol}(a_1, \ldots, a_n) = |\,\mu_0(a_1, \ldots, a_n)|$

Beweis. Sei

$$\mathrm{Vol}(a_1, \ldots, a_n) = \begin{cases} \mathrm{vol}(a_1, \ldots, a_n) & \text{wenn } \mu_0(a_1, \ldots, a_n) \geq 0 \\ -\mathrm{vol}(a_1, \ldots, a_n) & \text{wenn } \mu_0(a_1, \ldots, a_n) < 0 \end{cases}$$

das *orientierte* Volumen von $\mathrm{PE}(a_1, \ldots, a_n)$. Wenn man der geometrischen Anschauung vertraut, sieht man leicht, daß Vol eine alternierend Multilinearform ist: Wenn $a_1, \ldots, a_n$ linear abhängig sind, ist $\mathrm{PE}(a_1, \ldots, a_n)$ zu einer “Fläche” entartet und hat Volumen 0. Um zu sehen, daß $\mathrm{Vol}(a_1, \ldots, a_n)$ linear in (zum Beispiel) a_1 ist, schreiben wir

$$\mathrm{Vol}(a_1, \ldots, a_n) = d(a_1)\,\mathrm{Vol}(a_2, \ldots, a_n),$$

wobei $d(a_1)$ der mit dem richtigen Vorzeichen versehene Abstand von a_1 zur Fläche $\langle a_2, \dots, a_n \rangle$ ist. $d(x)$ ist klarerweise linear. Weil $\mathrm{Vol}(\mathrm{e}_1 \dots, \mathrm{e}_n) = 1$, ist $\mathrm{Vol} = \mu_0$. $\square$

Endomorphismen

Die *Determinante* $\det(\phi)$ *eines Endomorphismus* ϕ von V definiert man als die Determinante einer Matrixdarstellung A von ϕ. Wenn man die Basis wechselt, hat die neue Matrix die Form $B^{-1}AB$ und man hat

$$\det(B^{-1}AB) = \det(B)^{-1} \det(A) \det(B) = \det(A).$$

Die Definition von $\det(\phi)$ hängt also nicht von der Wahl der Basis ab. Die Invarianz der Definition wird klarer, wenn wir eine andere Definition verwenden. Sei μ eine Volumenform auf V. Dann ist die n–Form μ^ϕ ein Vielfaches von μ. Der Faktor ist von der Wahl von μ unabhängig und heißt die *Determinante* von ϕ:

Definition. *Sei V ein n–dimensionaler K–Vektorraum. Eine Volumenform ist eine nicht–triviale n–Form auf V.*

Definition. $\det(\phi)$ *ist definiert durch*[5]

$$\mu^\phi = \det(\phi)\mu.$$

Dabei ist μ eine beliebige Volumenform auf V.

Es gilt also

$$\mu(\phi(a_1), \dots, \phi(a_n)) = \det(\phi)\mu(a_1, \dots, a_n)$$

für alle n–Formen μ und alle $a_1, \dots, a_n \in V$. Wenn $V = \mathbb{R}^n$, ist das Volumen des Bildes $\phi(P)$ eines Parallelepipeds das $\det(\phi)$–fache des Volumens von P.

Das nächste Lemma sagt, daß beide Definition von $\det(\phi)$ übereinstimmen.

Lemma 4.5.2. *Im K^n ist* $\det(f_A) = \det(A)$

Beweis. Die Bilder der kanonischen Basisvektoren unter f_A sind die Spalten von A. Also ist für die Standardvolumenform μ_0

$$\det(A) = \mu_0(f_A(\mathrm{e}_1), \dots, f_A(\mathrm{e}_n)) = \det(f_A)\mu_0(\mathrm{e}_1, \dots, e_n) = \det(f_A).$$

$\square$

Wir erhalten einen neuen Beweis von Satz 4.3.4:

[5] Wenn V 0–dimensional ist, ist zwar $\phi = 0$, man setzt aber $\det(\phi) = 1$.

Satz 4.5.3. $\det(\phi \circ \psi) = \det(\phi)\det(\psi)$

Beweis. Weil

$$\begin{aligned}\mu^{\phi\circ\psi}(a_1,\ldots,a_n) &= \mu\big(\phi(\psi(a_1)),\ldots,\phi(\psi(a_n))\big)\\ &= \mu^{\phi}(\psi(a_1),\ldots,\psi(a_n))\\ &= (\mu^{\phi})^{\psi}(a_1,\ldots,a_n),\end{aligned}$$

ist

$$\mu^{\phi\circ\psi} = (\mu^{\phi})^{\psi} = \det(\psi)\mu^{\phi} = \det(\psi)\det(\phi)\mu.$$

$\square$

Kapitel 5

Endomorphismen

V sei in diesem Kapitel ein n–dimensionaler K-Vektorraum und ϕ ein Endomorphismus von V.

Wenn man eine Basis festlegt, wird ϕ durch eine n–n–Matrix A beschrieben. Eine *Invariante* eines Endomorphismus ϕ berechnet sich aus A, hängt aber nur von ϕ ab und nicht von der Wahl der Basis.

5.1 Diagonalisierbare Endomorphismen

Matrizen der Form

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \tag{5.1}$$

heißen *Diagonalmatrizen*.

Definition. *ϕ heißt diagonalisierbar, wenn ϕ bei geeigneter Basiswahl durch eine Diagonalmatrix dargestellt wird.*

Wenn ϕ bezüglich der Basis $b_1, \dots, b_n$ durch (5.1) dargestellt wird, ist

$$\phi(b_i) = \lambda_i b_i$$

für $i = 1, \dots, n$. Die b_i sind also Eigenvektoren von ϕ im Sinn der folgenden Definition.

Definition. *Ein von Null verschiedener Vektor v heißt Eigenvektor, wenn $\phi(v)$ Vielfaches von v ist, wenn also*

$$\phi(v) = \lambda v$$

für ein $\lambda \in K$. Der Faktor λ heißt Eigenwert von ϕ.

0 ist genau dann Eigenwert, wenn ϕ singulär ist. Wenn λ Eigenwert von ϕ ist, ist $\lambda + \mu$ Eigenwert von $\phi + \mu\mathrm{id}$ und $\lambda\mu$ Eigenwert von $\mu\phi$.

Das nächste Lemma ist klar.

Lemma 5.1.1. *ϕ ist genau dann diagonalisierbar, wenn V eine Basis aus Eigenvektoren hat.* □

Eigenvektoren und Eigenwerte einer quadratischen Matrix A sind die Eigenvektoren und Eigenwerte von f_A.

Beispiel

Die Eigenwerte einer oberen Dreiecksmatrix

$$A = \begin{pmatrix} \lambda_1 & * & \dots & * \\ 0 & \lambda_2 & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

sind $\lambda_1, \dots, \lambda_n$.

λ ist nämlich genau dann Eigenwert von A, wenn

$$\lambda\mathbf{I} - A$$

singulär ist. Eine obere Dreiecksmatrix ist genau dann singulär, wenn eine Null in der Diagonalen steht. Die Diagonalelemente von $\lambda\mathbf{I} - A$ sind aber $\lambda - \lambda_1, \dots, \lambda - \lambda_n$.

Definition.

$$V_\lambda = \{v \in V \mid \phi(v) = \lambda v\}$$

heißt der Eigenraum von λ.

V_λ ist ein Unterraum, der von ϕ in sich abgebildet wird, ein sogenannter *ϕ–invarianter* Unterraum. V_0 ist der Kern von ϕ. Wenn λ kein Eigenwert ist, ist $V_\lambda = 0$.

Lemma 5.1.2. *Wenn $\lambda_1, \dots, \lambda_k$ paarweise verschieden sind, sind die V_i linear unabhängig. Das heißt, daß eine Summe $v_1 + \cdots + v_k$ mit $v_i \in V_{\lambda_i}$ nur Null sein kann, wenn alle v_i Null sind.*

Beweis. Durch Induktion über k. Der Fall $k = 1$ ist klar. Aus dem Bestehen der Gleichung $v_1 + \cdots + v_k = 0$, erhalten wir durch Anwendung von ϕ einerseits und durch Multiplikation mit λ_k andererseits die beiden Gleichungen

$$\begin{array}{ccccccccc} \lambda_1 v_1 & + & \cdots & + & \lambda_{k-1} v_{k-1} & + & \lambda_k v_k & = & 0 \\ \lambda_k v_1 & + & \cdots & + & \lambda_k v_{k-1} & + & \lambda_k v_k & = & 0 \end{array}$$

Wir subtrahieren und erhalten

$$(\lambda_1 - \lambda_k) v_1 \quad + \quad \cdots \quad + \quad (\lambda_{k-1} - \lambda_k) v_{k-1} \quad = \quad 0$$

Induktion ergibt $(\lambda_i - \lambda_k) v_i = 0$ und daher $v_i = 0$ für $i = 1, \dots, k-1$. Es folgt auch $v_k = 0$. □

Seien $\lambda_1, \dots, \lambda_e$ die verschiedenen Eigenwerte von ϕ. Das letzte Lemma besagt, daß die Summe E der Eigenräume $V_{\lambda_1}, \dots, V_{\lambda_e}$ ist *direkt* ist: Das heißt, daß sich jedes Element von E eindeutig als Summe $v_1 + \cdots + v_e$ $(v_i \in V_{\lambda_i})$ schreiben läßt. Es folgt

$$\dim E = \sum_{i=1}^{e} \dim V_{\lambda_i} \tag{5.2}$$

Satz 5.1.3. *ϕ ist genau dann diagonalisierbar, wenn V Summe von Eigenräumen ist.*

Beweis. Wähle in jedem V_{λ_i} eine Basis $b^i_1, \dots, b^i_{n_i}$. Das letzte Lemma bedeutet, daß die Familie

$$\mathfrak{B} = (b^1_1, \dots, b^e_{n_e})$$

linear unabhängig ist. Die Voraussetzung hat also zur Folge, daß $\mathfrak{B}$ eine Basis von V ist, eine Basis aus Eigenvektoren. □

Folgerung 5.1.4. *ϕ hat höchstens n Eigenwerte. Wenn ϕ n Eigenwerte hat, ist ϕ diagonalisierbar.*

Beweis. Nach (5.2) ist

$$e \leq \dim E \leq n.$$

Es folgt $e \leq n$. Und $\dim E = n$, wenn $e = n$. □

Der Eigenraum V_λ einer Diagonalmatrix (5.1) (Seite 67) wird aufgespannt von den kanonischen Basisvektoren e_i, für die $\lambda_i = \lambda$. Die Dimension von V_λ ist also die Vielfachheit des Vorkommens von λ in der Diagonalen der Matrix. Die Dimensionen der Eigenräume bilden ein System von Invarianten, das die Diagonalmatrix bis auf eine Permutation der Diagonalen (d.h. eine Permutation der Basis) eindeutig festlegt.

5.2 Das charakteristische Polynom

Sei K ein Körper.

Definition. *Der Polynomring $K[x]$ ist eine assoziative, kommutative K–Algebra mit Einselement* 1 *und einem ausgezeichneten Element x, in dem sich jedes Element (*"Polynom"*) eindeutig in der Form*

$$p = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_0 1$$

schreiben läßt.

Die angegebene Bedingung bedeutet, daß die Folge $1, x, x^2, \ldots$ eine Basis von $K[x]$ bildet. Wenn man $\alpha \in K$ mit $\alpha 1$ identifiziert, wird K zu einer Unteralgebra von $K[x]$.

Man kann $K[x]$ folgendermaßen konstruieren. Sei A ein unendlichdimensionaler K–Vektorraum mit Basis $m_0.m_1, \ldots$ (zum Beispiel $K^{(\mathbb{N})}$, vgl. S.30). Mit Lemma 2.6.1 macht man A zu einer K–Algebra, indem man das Produkt von Basiselementen definiert durch

$$m_i m_j = m_{i+j}.$$

Wie im Beweis von 2.6.2, sieht man, daß A assoziativ und kommutativ ist mit m_0 als Einselement, weil die Multiplikation auf den Basiselementen (offensichtlich) diese Eigenschaften hat. Man identifiziert nun $\alpha \in K$ mit αm_0 und setzt $x = m_1$. Dann ist $x^i = m_i$ und A hat die Eigenschaften des Polynomrings. Man sieht leicht, daß $K[x]$ durch die angegebenen Eigenschaften bis auf Isomorphie eindeutig bestimmt ist.

Der *Grad* des Polynoms $p = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_0$ ist n, wenn $a_n \neq 0$. Das Polynom 0 hat den Grad $-\infty$. Man überlegt sich leicht, daß

$$\begin{aligned} \deg(pq) &= \deg(p) + \deg(q) \qquad \text{und} \\ \deg(p+q) &\leq \max\bigl(\deg(p), \deg(q)\bigr). \end{aligned}$$

p heißt *normiert*, wenn $a_n = 1$, und *konstant*, wenn $\deg(p) \leq 0$.

Sei A eine assoziative K–Algebra mit Einselement 1 und a ein Element von A. Dann ist die lineare Abbildung

$$p \longmapsto p(a) = a_n a^n + a_{n-1} a^{n-1} + \ldots + a_0 1$$

ein *Homomorphismus* von $K[x]$ nach A, der sogenannte *Einsetzungshomomorphismus*. Das heißt, daß

$$(pq)(a) = p(a)q(a)$$

für alle $p, q \in K[x]$. Das folgt daraus, daß die bilineare Abbildung $(p,q) \mapsto (pq)(a) - p(a)q(a)$ auf der Basis $1, x, x^2, \ldots$ und daher auf ganz $K[x]$ verschwindet.

Wenn A der Körper K selbst ist, ist der Einsetzungshomomorphismus das wohlbekannte Einsetzen eines Körperelements a in Polynome. Man nennt a *Nullstelle* von p, wenn $p(a) = 0$. Wir haben aber jetzt auch die Möglichkeit Matrizen aus $A = \mathrm{M}_n(K)$ und Endomorphismen aus $A = \mathrm{End}(V)$ in Polynome einzusetzen.

$K[x]$ ist ein *Integritätsbereich*. Das ist ein kommutativer Ring mit $1 \neq 0$, der *nullteilerfrei* ist: $rs = 0 \Rightarrow r = 0 \,\text{oder}\, s = 0$. Ein Integritätsbereich R läßt sich immer in einen Körper Q, seinen *Quotientenkörper* einbetten. Man definiert auf der Menge aller Paare (r, s), $(r, s \in R, s \neq 0)$ eine Äquivalenzrelation durch $(r, s) \sim (r', s') \Leftrightarrow rs' = sr'$ und setzt

$$Q = \{\frac{r}{s} \mid r, s \in R, s \neq 0\},$$

wobei $\frac{r}{s}$ die Äquivalenzklasse von (r, s) bezeichnet. Mit den Operationen

$$\frac{r}{s} + \frac{r'}{s'} = \frac{rs' + sr'}{ss'} \qquad \text{und}$$
$$\frac{r}{s}\frac{r'}{s'} = \frac{rr'}{ss'}$$

wird Q zu einem Körper. Wenn man r mit $\frac{r}{1}$ identifiziert, erhält man R als Unterring. Der Quotientenkörper von $K[x]$ ist der *rationale Funktionenkörper* $K(x)$.

Aus der Gleichung (4.1) auf Seite 58 folgt sofort:

Lemma 5.2.1. *Sei $A(x)$ eine Matrix aus Polynomen, dann ist* $\det(A(x))$ *ein Polynom.* □

Definition. *Das charakteristische Polynom* P_A *einer quadratischen Matrix A ist*

$$\mathrm{P}_A = \det(x\mathbf{I} - A).$$

Das charakteristische Polynom einer 2–2–Matrix $A = (a_{ij})$ ist zum Beispiel

$$\mathrm{P}_A = (x - a_{11})(x - a_{22}) - a_{21}a_{12} = x^2 - (a_{11} + a_{22})x + (a_{11}a_{22} - a_{21}a_{12}).$$

Auf Seite 59 haben wir die Determinante einer oberen Dreiecksmatrix ausgerechnet. Daraus folgt, daß das charakteristische Polynom von

$$\begin{pmatrix} \lambda_1 & * & \dots & * \\ 0 & \lambda_2 & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

das Produkt

$$\mathrm{P}_A = (x - \lambda_1) \cdots (x - \lambda_n)$$

ist. Wir sagen, daß P_A in *Linearfaktoren zerfällt.*

Wenn A und B ähnlich sind, sind auch $x\mathbf{I} - A$ und $x\mathbf{I} - B$ ähnlich. Ähnliche Matrizen haben also das gleiche charakteristische Polynom (vgl. S.65). Man kann daher das charakteristische Polynom $\mathrm{P}_\phi(x)$ eines Endomorphismus durch das charakteristische Polynom einer zugehörigen Matrix definieren. $\mathrm{P}_\phi(x)$ ist eine Invariante von ϕ.

Lemma 5.2.2. *Sei ϕ Endomorphismus eines n–dimensionalen Vektorraums.*

1. $\mathrm{P}_\phi(x)$ *ist ein normiertes Polynom vom Grad n.*
2. *Die Eigenwerte von ϕ sind die Nullstellen von* $\mathrm{P}_\phi(x)$*.*

Beweis. Für $A = (\alpha_{ij})$ folgt aus (4.1) (Seite 58)

$$\mathrm{P}_A(x) = \prod_{j=1}^{n}(x - a_{jj}) + \text{Polynome vom Grad} \leq n-2 \tag{5.3}$$

Daraus ergibt sich die erste Behauptung. λ ist genau dann Eigenwert von A, wenn $\lambda\mathbf{I}-A$ singulär ist, das heißt, wenn $\det(\lambda\mathbf{I}-A) = 0$. Die zweite Behauptung folgt jetzt sofort aus

$$\mathrm{P}_A(\lambda) = \det(\lambda\mathbf{I} - A). \tag{5.4}$$

□

Wenn man in die letzte Gleichung 0 einsetzt ergibt sich $p_A(0) = \det(-A) = (-1)^n \det(A)$. Wir finden also im charakteristischen Polynom

$$\mathrm{P}_\phi = x^n + a_{n-1}x^{n-1} + \ldots + a_0$$

die Determinante von ϕ wieder als

$$\det(\phi) = (-1)^n a_0.$$

Aus (5.3) entnimmt man, daß $a_{n-1} = -\sum_{j=1}^n \alpha_{jj}$. Man nennt die Summe der Diagonalelemente die *Spur* $\mathrm{Tr}(A)$ von A. Die Spur ist wie das ganze charakteristische Polynom eine Invariante von ϕ. Wir notieren

$$\mathrm{Tr}(\phi) = -a_{n-1}.$$

Daß ähnliche Matrizen dieselbe Spur haben, sieht man auch so: Die Gleichung

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$$

ist leicht nachzurechnen. Für reguläre B folgt dann

$$\mathrm{Tr}\big(B^{-1}AB\big) = \mathrm{Tr}\big(BB^{-1}A\big) = \mathrm{Tr}(A).$$

Satz 5.2.3. *ϕ läßt sich genau dann bei geeigneter Basiswahl durch eine obere Dreiecksmatrix darstellen, wenn P_ϕ in Linearfaktoren zerfällt.*

Daß das charakteristische Polynom einer oberen Dreiecksmatrix in Linearfaktoren zerfällt, wissen wir schon. Für die Umkehrung brauchen wir einen Hilfssatz:

Lemma 5.2.4. *Sei U ein ϕ–invarianter Unterraum von V. Durch*

$$\overline{\phi}(v+U) = \phi(v) + U$$

wird ein Endomorphismus von V/U definiert. Für die charakteristischen Polynome gilt

$$\mathrm{P}_\phi = \mathrm{P}_{\phi\restriction U}\ \mathrm{P}_{\overline{\phi}}.$$

Beweis. Zuerst müssen wir zeigen, daß $\overline{\phi}$ wohldefiniert ist: Sei $v+U = v'+U$. Dann ist $v-v' \in U$ und folglich $\phi(v)-\phi(v') = \phi(v-v') \in U$. Also ist $\phi(v)+U = \phi(v')+U$. Die Linearität von $\overline{\phi}$ folgt einfach aus der Linearität von ϕ.

Wir wählen eine Basis $b_1, \dots, b_m$ von U und setzen sie zu einer Basis $b_1, \dots, b_n$ von V fort. ϕ wird jetzt von einer Matrix der Form

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

dargestellt. Die m–m–Matrix A ist dabei die Matrix von $\phi \restriction U$ und C die Matrix von $\overline{\phi}$ bezogen auf die Basis $b_{m+1}+U, \dots, b_n+U$ von V/U. Die Behauptung folgt jetzt aus Gleichung (4.2) auf Seite 59. □

Wir zeigen Satz 5.2.3 durch Induktion über $n = \dim V$. Sei

$$\mathrm{P}_\phi = (x-\lambda_1)\cdots(x-\lambda_n).$$

Für $n = 0$ (und $n = 1$) ist nichts zu zeigen. Nehmen wir also an, daß $n > 0$. Dann ist λ_1 Eigenwert von ϕ mit Eigenvektor b_1. Wir wenden das Lemma auf $U = Kb_1$ an und erhalten

$$\prod_{i=1}^{n}(x-\lambda_i) = (x-\lambda_1)\ \mathrm{P}_{\overline{\phi}}.$$

Weil $K[x]$ ein Integritätsbereich ist[1], können wir schließen, daß

$$\mathrm{P}_{\overline{\phi}} = \prod_{i=2}^{n}(x-\lambda_i).$$

Nach Induktionsannahme finden wir eine Basis $b_2+U, \dots, b_n+U$ von V/U, für die $\overline{\phi}$ die obere Dreiecksmatrix C hat. Bezogen auf die Basis $b_1, \dots, b_n$ hat dann ϕ die Matrix

$$\begin{pmatrix} \lambda_1 & B \\ \mathbf{0} & C \end{pmatrix}.$$

[1] Aus $ab = ac$ folgt $a(b-c) = 0$ und $b-c = 0$, wenn $a \neq 0$.

$\square$

Der Beweis konstruiert eine obere Dreiecksmatrix, in deren Diagonale $\lambda_1, \ldots, \lambda_n$ stehen.

Satz 5.2.5 (Hamilton)**.** *Jeder Endomorphismus ist Nullstelle seines charakteristischen Polynoms:*

$$\mathrm{P}_\phi(\phi) = 0.$$

Gemeint ist das Folgende: Sei V ein endlichdimensionaler Vektorraum und ϕ ein Endomorphismus. Dann bildet der Einsetzungshomomorphismus $K[x] \to \mathrm{End}(V)$, der x auf ϕ abbildet, das Polynom P_ϕ auf 0 ab. Man rechnet zum Beispiel leicht nach, daß für jede 2–2–Matrix $A = (\alpha_{ij})$

$$(A - \alpha_{11}\mathbf{I})(A - \alpha_{22}\mathbf{I}) - \alpha_{21}\alpha_{12}\mathbf{I} = \mathbf{0}.$$

Beweis. [2] Wir zeigen zuerst einen Hilfssatz:

Sei V ein Modul über einem kommutativen Ring R, $b_1, \ldots, b_n$ eine Folge von Elementen von V und A eine n–n–Matrix mit Koeffizienten in R. Wenn

$$A \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

werden alle b_i von der Determinante von A annulliert.

Beweis: Satz 4.4.3 drückt eine Identität aus, die auch in allen kommutativen Ringen gültig sein muß. Es ist also

$$\det(A)\,\mathbf{I} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \mathrm{adj}(A)\,A \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \mathrm{adj}(A) \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Damit ist der Hilfssatz bewiesen.

Sei nun

$$R = \bigl\{ p(\phi) \bigm| p \in K[x] \bigr\}$$

die von ϕ erzeugte Unteralgebra von $\mathrm{End}(V)$. V ist natürlich ein R–Modul. Sei $b_1, \ldots, b_n$ eine Basis von V und A die dazugehörige Matrix für ϕ. Das bedeutet

$$\phi\,\mathbf{I} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = A^\top \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

[2] Ich danke C. Schötz für den Hinweis auf einen Fehler in einer früheren Fassung des Beweises.

oder

$$(\phi\,\mathbf{I} - A^\top)\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Der Hilfssatz impliziert

$$\det(\phi\,\mathbf{I} - A)\,b_i = \det((\phi\,\mathbf{I} - A)^\top)\,b_i = 0$$

für $i = 1, \ldots, n$. Also ist $p_A(\phi) = \det(\phi\,\mathbf{I} - A)$ der Nullendomorphismus. □

Wenn P_ϕ in Linearfaktoren zerfällt[3], folgt der Satz von Hamilton wie im Beweis von 5.2.3 durch Induktion: Sei U von einem Eigenvektor von λ_1 aufgespannt und $\overline{\phi}$ der auf V/U induzierte Endomorphismus. Nach Induktionsvoraussetzung ist $\overline{\phi}$ Nullstelle seines charakteristischen Polynoms. Das heißt, daß das Bild von

$$\prod_{i=2}^{n}(\phi - \lambda_i)$$

in U liegt. Weil U im Kern von $\phi - \lambda_1$ liegt, ist $(\phi - \lambda_1)\prod_{i=2}^{n}(\phi - \lambda_i)$ der Nullendomorphismus.

[3] Weil sich jeder Körper in einen algebraisch abgeschlossenen Körper einbetten läßt, ist diese Voraussetzung harmlos. Vergleiche 5.3.7.

5.3 Zerlegung in Haupträume

Sei V ein Vektorraum und $\lambda \in K$. In diesem Abschnitt werden wir für den Endomorphismus $\lambda \cdot \mathrm{id}_V$ einfach λ schreiben.

Definition. *Der Hauptraum V'_λ von λ ist die Menge aller Vektoren, die von einer Potenz von $\phi - \lambda$ auf Null abgebildet werden.*

V'_λ ist als Vereinigung der aufsteigenden Folge

$$\mathbf{0} = \mathrm{Ker}(1) \subset \mathrm{Ker}(\phi - \lambda) \subset \mathrm{Ker}(\phi - \lambda)^2 \subset \dots$$

ein Unterraum von V. V'_λ ist invariant unter ϕ, weil aus $(\phi - \lambda)^m x = 0$ folgt, daß $(\phi - \lambda)^m \phi x = \phi(\phi - \lambda)^m x = \phi 0 = 0$.

Wenn $x \in V'_\lambda \setminus \mathbf{0}$ und m minimal mit $(\phi - \lambda)^m x = 0$, ist $(\phi - \lambda)^{m-1} x$ ein Element von $V_\lambda \setminus \mathbf{0}$. Also ist $V'_\lambda = \mathbf{0}$, wenn $V_\lambda = \mathbf{0}$.

Beispiel

Sei $\lambda_1, \dots, \lambda_n$ eine Folge von Körperelementen. Wir nehmen an, daß die ersten k Zahlen gleich λ_0 sind und $\lambda_{k+1}, \dots, \lambda_n$ verschieden von λ_0. Dann wird der zu λ_0 gehörende Hauptraum der Matrix

$$A = \begin{pmatrix} \lambda_1 & * & \dots & * \\ 0 & \lambda_2 & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

von $\mathrm{e}_1 \dots, \mathrm{e}_k$ aufgespannt. Das läßt sich folgendermaßen einsehen:

$A - \lambda_0$ bildet alle Vektoren aus $\langle \mathrm{e}_1, \dots, \mathrm{e}_i \rangle$ modulo $\langle \mathrm{e}_1, \dots, \mathrm{e}_{i-1} \rangle$ auf ihr $(\lambda_i - \lambda_0)$–faches ab. Daraus folgt, daß genau die Elemente von $\langle \mathrm{e}_1, \dots, \mathrm{e}_k \rangle$ von einer Potenz von $A - \lambda_0$ auf 0 abgebildet werden.

In diesem Beispiel ist die Dimension von V'_{λ_0} gleich k. k ist die *Vielfachheit*[4] der Nullstelle λ_0 von $\mathrm{P}_A = (x - \lambda_0)^k (x - \lambda_{k+1}) \dots (x - \lambda_n)$.

Lemma 5.3.1. *Die Dimension von V'_λ ist die Vielfachheit*[4] *von λ in P_ϕ. Auf V'_λ läßt sich ϕ durch eine obere Dreiecksmatrix*

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

darstellen, in deren Diagonale λ steht.

Beweis. Wenn k die Vielfachheit von λ in P_ϕ ist, ist $\mathrm{P}_\phi = (x-\lambda)^k q(x)$, mit $q(\lambda) \neq 0$. Man zeigt wie im Beweis von Satz 5.2.3, die Existenz eines ϕ-invarianten Unterraums U der Dimension k, auf dem ϕ durch eine Dreiecksmatrix der behaupteten Form dargestellt werden kann. U ist ein Unterraum von V'_λ. Wenn U ein echter Unterraum von V'_λ wäre, gäbe es ein $v \in V'_\lambda \setminus U$ mit $(\phi - \lambda)v \in U$. Dann wäre also λ ein Eigenwert des induzierten Endomorphismus $\overline{\phi} : V/U \to V/U$. Das ist nicht möglich, weil q das charakteristische Polynom von $\overline{\phi}$ ist. □

Das folgende Lemma ist eine Verschärfung von 5.1.2.

Lemma 5.3.2. *Die Haupträume von ϕ sind unabhängig.*

Beweis. Sei ein Körperelement λ festgehalten. Während jedes $x \in V'_\lambda$ von einer Potenz von $\phi - \lambda$ auf Null abgebildet wird, operieren für $\mu \neq \lambda$ die $\phi - \mu$ und damit auch alle Potenzen als Automorphismus von V'_λ. Sonst hätte nämlich $\phi - \mu$ einen nicht–trivialen Kern in V'_λ; mit anderen Worten: V_μ enthielte ein nicht–triviales Element von V'_λ. Das geht aber nicht, weil $\phi - \lambda$ auf V_μ als Automorphismus (nämlich als $\mu - \lambda$) operiert.

Seien nun $\lambda_1, \ldots, \lambda_e$ paarweise verschieden und

$$v_1 + \cdots + v_e = 0$$

mit $v_i \in V_{\lambda_i}$, wählen wir für jedes $i \neq 1$ eine Potenz q_i von $\phi - \lambda_i$, die v_i auf Null abbildet. Das Produkt $q = q_2 q_3 \ldots q_n$ annulliert alle $v_2, \ldots, v_n$ und operiert als Automorphismus auf V'_{λ_1}. Es folgt also $qv_1 = 0$ und daraus $v_1 = 0$. □

Aus den letzten beiden Lemmas folgt:

Folgerung 5.3.3. *Wenn P_ϕ in Linearfaktoren zerfällt, ist V die direkte Summe der Haupträume von ϕ.*

Wenn V direkte Summe von ϕ–invarianten Unterräumen $V_1, \ldots, V_e$ ist, kann man Basen $\mathfrak{B}_\mathrm{i}$ der V_i zu einer Basis $\mathfrak{B}$ von V zusammenfassen. Wenn die $\phi \restriction V_i$ bezüglich der $\mathfrak{B}_\mathrm{i}$ durch die Matrizen A_i dargestellt werden, wird ϕ bezüglich $\mathfrak{B}$ durch die *Blockmatrix*

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_e \end{pmatrix}$$

dargestellt.

Die nächste Folgerung wird im nächsten Abschnitt wesentlich verbessert zum Satz von der Jordanschen Normalform (5.4.5).

[4] Siehe Seite 79

Folgerung 5.3.4. *Wenn*

$$p_\phi(x) = \prod_{i=1}^{e} (x - \lambda_i)^{n_i},$$

(die λ_i paarweise verschieden), kann ϕ bezüglich einer geeigneten Basis durch eine Blockmatrix dargestellt werden, deren Blöcke n_i-n_i–Matrizen der Form

$$A_i = \begin{pmatrix} \lambda_i & * & \cdots & * \\ 0 & \lambda_i & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_i \end{pmatrix} \tag{5.5}$$

sind.

Beweis. Das charakteristische Polynom der Einschränkung ϕ_i von ϕ auf V'_λ ist $(x-\lambda)^{n_i}$. Bei geeigneter Wahl einer Basis wird (nach 5.2.3) ϕ_i von einer oberen Dreiecksmatrix dargestellt, in deren Diagonale λ_i steht. □

Exkurs: Nullstellen von Polynomen

Lemma 5.3.5. *λ ist genau dann Nullstelle von $p(x)$, wenn $x - \lambda$ ein Teiler von $p(x)$ ist.*

Beweis. Sei λ ein Nullstelle von p. Wir zeigen durch Induktion über den Grad von p, daß $x-\lambda$ ein Teiler von p ist. Wenn $p = 0$, ist die Behauptung klar. Wenn p nicht Null ist, ist der Grad n von $p = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0 1$ mindestens 1. Dann ist λ auch Nullstelle von

$$q = p - a_n(x - \lambda)^n.$$

Weil der Grad von q kleiner als n ist, wird q, und damit auch p, von $x - \lambda$ geteilt. □

Ein alternativer Beweis geht so:

$$x^n - \lambda^n = (x - \lambda)(x^{n-1} + x^{n-2}\lambda + \cdots + \lambda^{n-1})$$

wird von $x - \lambda$ geteilt, also auch $p(x) - p(\lambda) = p(x)$.

Es folgt sofort, daß ein Polynom vom Grad n höchstens n verschiedene Nullstellen haben kann. Daraus ergibt ein neuer Beweis der Tatsache, daß ϕ höchstens n Eigenwerte hat (5.1.4).

Folgerung 5.3.6. *Jedes $p(x) \neq 0$ läßt sich (bis auf die Reihenfolge) eindeutig zerlegen als*

$$(x - \lambda_1) \cdots (x - \lambda_n) p_0(x),$$

wobei p_0 keine Nullstellen hat.

Beweis. Nur die Eindeutigkeit ist noch zu zeigen: Sei

$$(x-\lambda'_1)\cdots(x-\lambda'_{n'})p'_0(x)$$

eine andere Zerlegung von p. Wir können annehmen, daß $n > 0$. Weil λ_1 eine Nullstelle von p ist, muß eins der λ'_i gleich λ_1 sein, sagen wir $\lambda'_1 = \lambda_1$. Dann ist

$$(x-\lambda_2)\cdots(x-\lambda_n)p_0(x) = (x-\lambda'_2)\cdots(x-\lambda'_{n'})p'_0(x),$$

und die Eindeutigkeit folgt per Induktion. □

Definition. *Die Vielfachheit der Nullstelle λ von p ist die Anzahl der Faktoren der Form $(x-\lambda)$ in der obigen Zerlegung von p.*

Man überlegt sich leicht, daß die Vielfachheit von λ das größte k ist, sodaß $(x-\lambda)^k$ das Polynom p teilt.

Ein Körper heißt *algebraisch abgeschlossen*, wenn jedes nicht-konstante Polynom eine Nullstelle besitzt. Der Körper $\mathbb{C}$ der komplexen Zahlen ist algebraisch abgeschlossen. Jeder Körper läßt sich in einen algebraisch abgeschlossenen Körper einbetten.

Folgerung 5.3.7. *In einem algebraisch abgeschlossenen Körper zerfällt jedes normierte Polynom in Linearfaktoren.*

5.4 Die Jordansche Normalform

Wir wollen die Matrixdarstellung (5.5) (Seite 78) eines Endomorphismus auf seinen Haupträumen weiter vereinfachen. Dazu nehmen wir an, daß V selbst ein Hauptraum von ϕ ist. Weiter können wir annehmen, daß V Hauptraum zum Eigenwert 0 ist. Sonst ersetzen wir ϕ durch $\phi - \lambda$, wenn λ der Eigenwert von ϕ ist.

ϕ ist dann *nilpotent* im Sinne der folgenden Definition:

Definition. *Eine Endomorphismus ϕ heißt nilpotent, wenn*

$$\phi^m = 0$$

für ein genügend großes m. Eine quadratische Matrix A heißt nilpotent, wenn f_A nilpotent ist.

Das nächste Lemma ist uns im wesentlichen schon bekannt.

Lemma 5.4.1. *Sei ϕ Endomorphismus eines n–dimensionalen Vektorraums V. Dann sind äquivalent:*

a) Jedes $x \in V$ wird durch eine genügend große Potenz von ϕ annulliert.

b) ϕ *ist nilpotent*

c) $\phi^n = 0$

d) $\mathrm{P}_\phi(x) = x^n$

e) Bei geeigneter Basiswahl läßt sich ϕ *durch eine obere Dreiecksmatrix darstellen, in deren Diagonale nur Nullen stehen.*

Beweis.
Für die Äquivalenz von a), b) und c) genügt es zu zeigen, daß

a)$\to$c): Es ist leicht zu sehen, daß für beliebige Endomorphismen ϕ und ψ

$$\mathrm{Ker}(\phi) \subset \mathrm{Ker}(\psi \circ \phi) = \phi^{-1}(\mathrm{Ker}(\psi).$$

Daraus folgt, daß die Kerne der Potenzen ϕ^i eine aufsteigende Folge bilden. Also gibt es ein $m \leq n$ mit $\mathrm{Ker}\,\phi^m = \mathrm{Ker}\,\phi^{m+1}$ oder, anders ausgedrückt, $\mathrm{Ker}\,\phi^m = \phi^{-1}(\mathrm{Ker}\,\phi^m)$. $\phi \restriction \mathrm{Ker}\,\phi^m$ ist also regulär. Aus a) folgt jetzt, daß $\mathrm{Ker}\,\phi^m = V$. Also ist $\phi^m = 0$ und daher auch $\phi^n = 0$.

Die Äquivalenz von d) und e) folgt sofort aus Folgerung 5.3.4 und der Formel für das charakteristische Polynom einer oberen Dreiecksmatrix (Seite 71).

e)$\to$c): Sei $b_1, \ldots, b_n$ eine Basis bezüglich der ϕ durch eine obere Dreiecksmatrix dargestellt wird, in deren Diagonale nur Nullen stehen. Dann ist

$$\phi\langle b_1, \ldots, b_i\rangle \subset \langle b_1, \ldots, b_{i-1}\rangle \tag{5.6}$$

(siehe das Beispiel auf Seite 76). Daraus folgt $\phi^n\langle b_1, \ldots, b_n\rangle = 0$. Es sei bemerkt, daß d)$\to$c) auch direkt aus dem Satz von Hamilton (5.2.5) folgt.

a)$\to$e): Wenn wir schon wissen, daß das charakteristische Polynom von ϕ in Linearfaktoren zerfällt (zum Beispiel, weil K algebraisch abgeschlossen ist), folgt die Behauptung sofort aus Folgerung 5.3.4, weil a) bedeutet, daß V ein Hauptraum von 0 ist. Wenn wir das nicht voraussetzen, müssen wir den Beweis von Satz 5.2.3 wiederholen:

Sei ϕ nilpotent. Wir wollen eine Basis konstruieren die (5.6) erfüllt. Wenn $n = 0$, ist nichts zu zeigen. Nehmen wir an, daß $n > 0$ und die Behauptung für Dimension $n-1$ gilt. Dann gibt es ein $b_1 \neq 0$ mit $\phi(b_1) = 0$. Sei $\overline{\phi}$ der von ϕ induzierte Endomorphismus von $V/\langle b_1\rangle$. Dann ist $\overline{\phi}$ wieder nilpotent und wir finden nach Annahme eine Basis $b_2 + \langle b_1\rangle, \ldots, b_n + \langle b_1\rangle$ von $V/\langle b_1\rangle$, die (5.6) für $\overline{\phi}$ erfüllt. $b_1, \ldots, b_n$ ist dann wie gewünscht. □

Die Eigenschaft c) und Lemma 5.3.1 implizieren, daß $(\phi - \lambda)^m x = 0$ für alle $x \in V'_\lambda$, wenn m die Vielfachheit von λ in P_ϕ ist. Also ist

$$V'_\lambda = \mathrm{Ker}(\phi - \lambda)^m.$$

Lemma 5.4.2. *Sei ϕ nilpotent und $v \in V$. Sei k die kleinste natürliche Zahl, für die $\phi^k(v) = 0$. Dann ist die Folge*

$$v, \phi^1(v), \phi^2(v), \ldots, \phi^{k-1}(v)$$

Basis eines ϕ–invarianten Unterraums U von V.

Man nennt U den von v *erzeugten zyklischen* Unterraum (der Ordnung k).

Beweis. Daß U ϕ–invariant ist, ist klar. Wir zeigen die lineare Unabhängigkeit durch Induktion über k. Für $k = 0$ ist nichts zu zeigen ($U = \mathbf{0}$). Für $k > 0$ nehmen wir an, daß

$$\alpha_0 v + \alpha_1 \phi(v) + \cdots \alpha_{k-2}\phi^{k-2}(v) + \alpha_{k-1}\phi^{k-1}(v) = 0.$$

Wir multiplizieren mit ϕ und erhalten

$$\alpha_0 \phi(v) + \alpha_1 \phi^2(v) + \cdots \alpha_{k-2}\phi^{k-1}(v) = 0.$$

Wir wenden die Induktionsvoraussetzung auf $\phi(v)$ an und erhalten

$$\alpha_0 = \alpha_1 = \cdots = \alpha_{k-2} = 0.$$

Daraus folgt auch $\alpha_{k-1} = 0$, weil $\phi^{k-1}(v) \neq 0$. □

Auf U wird ϕ bezüglich der im Lemma angegeben Basis (in umgekehrter Reihenfolge) durch die k–k–Matrix

$$\mathrm{J}_0^k = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \end{pmatrix} \tag{5.7}$$

dargestellt, in der in der *Nebendiagonalen* Einsen stehen, und sonst Nullen. Man beachte, daß $\mathrm{J}_0^1 = \mathbf{0}$.

Satz 5.4.3. *Sei ϕ ein nilpotenter Endomorphismus von V. Dann ist V direkte Summe von zyklischen Unterräumen.*

Beweis. Die Räume $U^i = \mathrm{Ker}(\phi) \cap \mathrm{Im}(\phi^i)$ bilden eine absteigende Folge

$$\mathrm{Ker}(\phi) = U^0 \supset U^1 \supset \ldots \supset U^n = 0$$

von ϕ–invarianten Unterräumen. Wir wählen eine Basis $b_1^n, \ldots, b_{m_n}^n$ von U^{n-1}. Dann erweitern wir diese Basis um $b_1^{n-1}, \ldots, b_{m_{n-1}}^{n-1}$ zu einer Basis von U^{n-2} und fahren so fort. Schließlich erhalten wir eine Basis (b_j^i), $(i = 1, \ldots, n;\ j = 1, \ldots, m_i)$ von $\mathrm{Ker}(\phi)$, sodaß jeweils $\left(b_j^{i'} \mid i < i';\ j < m_{i'}\right)$ eine Basis von U^i ist.

Dann wählen wir zu jedem b^i_j einen Vektor v^i_j mit $\phi^{i-1}(v^i_j) = b^i_j$. Jedes dieser v^i_j erzeugt einen zyklischen Unterraum V^i_j der Ordnung i. Wir werden zeigen, daß V die direkte Summe der V^i_j ist.

Die V^i_j erzeugen V: Sei $v \in V$ und m minimal mit $\phi^m v = 0$. Wir zeigen durch Induktion über m, daß v im Erzeugnis W der V^i_j liegt. Für $m = 0$ ist nichts zu zeigen. Wenn $m > 0$, liegt $\phi^{m-1}v$ in U^{m-1}. W ist so konstruiert, daß U^{m-1} in $\phi^{m-1}(W)$ liegt. Also gibt es ein $w \in W$ mit

$$\phi^{m-1}w = \phi^{m-1}v.$$

Dann ist aber $\phi^{m-1}(v-w) = 0$ und nach Induktionsvoraussetzung ist $v-w \in W$. Es folgt $v \in W$.

Die V^i_j sind unabhängig: Nehmen wir an, es gäbe Elemente $w^i_j \in V^i_j$, nicht alle Null, deren Summe verschwindet. Wir multiplizieren die w^i_j mit einer geeigneten Potenz ϕ^m, so daß alle $\phi^m w^i_j$ im Kern von ϕ liegen, aber nicht alle gleich Null sind. Dann ist auch die Summe der $\phi^m w^i_j$ Null. Das ist unmöglich, weil jetzt

$$\phi^m w^i_j \in V^i_j \cap \mathrm{Ker}(\phi) = K b^i_j$$

und die Kb^i_j unabhängig sind. □

Folgerung 5.4.4. *Ein nilpotenter Endomorphismus ϕ läßt sich bezüglich einer geeigneten Basis durch eine aus Matrizen J^k_0 zusammengesetzte Blockmatrix beschreiben.* □

Beispiel Betrachte den durch die Blockmatrix

$$\begin{pmatrix} \mathrm{J}^3_0 & 0 & 0 & 0 \\ 0 & \mathrm{J}^2_0 & 0 & 0 \\ 0 & 0 & \mathrm{J}^1_0 & 0 \\ 0 & 0 & 0 & \mathrm{J}^1_0 \end{pmatrix}$$

gegeben Endomorphismus ϕ des K^7. Die im Beweis von 5.4.3 angegebene Konstruktion wird im folgenden Bild veranschaulicht.

$\mathrm{Ker}\,\phi^3 = V$	$v^3_1 = \mathrm{e}_3$		
$\mathrm{Ker}\,\phi^2$	$\phi(v^3_1) = \mathrm{e}_2$	$v^2_1 = \mathrm{e}_5$	
$\mathrm{Ker}\,\phi$	$b^3_1 = \mathrm{e}_1$	$b^2_1 = \mathrm{e}_4$	$v^1_1 = b^1_1 = \mathrm{e}_6,\ v^1_2 = b^1_2 = \mathrm{e}_7$
	$\mathrm{Im}\,\phi^2$	$\mathrm{Im}\,\phi$	$\mathrm{Im}\,\phi^0 = V$

In einem zyklischen Raum der Ordnung k hat $\mathrm{Ker}(\phi) \cap \mathrm{Im}(\phi^i)$ die Dimension 1, wenn $i < k$, und sonst die Dimension 0. Wenn daher J_0^k in der Blockmatrix von ϕ genau μ^k–mal vorkommt, ist

$$\mu^k = \dim\bigl(\mathrm{Ker}(\phi) \cap \mathrm{Im}(\phi^{k-1})\bigr) - \dim\bigl(\mathrm{Ker}(\phi) \cap \mathrm{Im}(\phi^k)\bigr).$$

Die zu ϕ gehörende Blockmatrix ist also bis auf eine Permutation der Blöcke eindeutig bestimmt[5].

Daraus ergibt sich die Eindeutigkeitsaussage des nächsten Satzes.

Satz 5.4.5 (Jordansche Normalform)**.** *Ein Endomorphismus eines endlichdimensionalen Vektorraums, dessen charakteristisches Polynom in Linearfaktoren zerfällt, läßt sich bezüglich einer geeigneten Basis durch eine Blockmatrix beschreiben, die sich aus Jordanmatrizen* J_λ^k *zusammengesetzt. Dabei ist*

$$\mathrm{J}_\lambda^k = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix}.$$

Diese sogenannte Jordanform ist bis auf die Reihenfolge der Jordanblöcke eindeutig bestimmt.

Beweis. Aus 5.3.3 und 5.4.4. □

Eine andere Formulierung des Satzes ist: *Jede quadratische Matrix A, deren charakteristisches Polynom in Linearfaktoren zerfällt, ist einer Matrix J in Jordanform ähnlich. J ist bis auf die Reihenfolge der Jordanblöcke eindeutig durch A bestimmt.*

[5] Im Beweis von 5.4.3 ist $m_i = \mu^i$.

Kapitel 6

Dualität

6.1 Der Dualraum

Definition. *Der Dualraum V^* eines K–Vektorraums V ist der Vektorraum aller Linearformen $\lambda : V \to K$.*

In früherer Notation ist also $V^* = \mathrm{L}(V, K)$.

Lineare Abbildungen $\lambda : K^n \to K$ werden beschrieben durch Zeilenvektoren

$$\lambda = (\alpha_1 \dots \alpha_n) \in \mathrm{M}_{1,n} \, .$$

Identifiziert man K^n mit dem Raum $\mathrm{M}_{n,1}$ der n–dimensionalen Spaltenvektoren

$$x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

operiert λ als Matrizenmultiplikation:

$$\lambda(x) = \alpha_1 \xi_1 + \dots + \alpha_n \xi_n = (\alpha_1 \dots \alpha_n) \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} .$$

Sei $\mathfrak{B} = (b_1, \dots, b_n)$ eine Basis von V. Die Koordinatenfunktionale

$$\lambda_i(\xi_1 b_1 + \cdots + \xi_n b_n) = \xi_i$$

bilden dann eine Basis von V^*. Denn, daß λ bezüglich $\mathfrak{B} = (b_1, \ldots, b_n)$ durch die Zeile $(\alpha_1 \cdots \alpha_n)$ beschrieben wird:

$$\lambda(\xi_1 b_1 + \cdots + \xi_n b_n) = \alpha_1 \xi_1 + \cdots + \alpha_n \xi_n,$$

drückt sich aus als

$$\lambda = \alpha_1 \lambda_1 + \cdots + \alpha_n \lambda_n.$$

Anders gesagt:

Definition. *Sei $b_1, \ldots, b_n$ eine Basis von V. Die Linearformen $\lambda_1, \ldots, \lambda_n$, definiert durch*

$$\lambda_i(b_j) = \delta_{ij},$$

bilden die duale *Basis von V^*.*

Nehmen wir an, daß die Spalten $b_1, \ldots, b_n$ der Matrix B eine Basis des K^n bilden (das heißt, daß B regulär ist.) Die Zeilen $\lambda_1, \ldots, \lambda_n$ der Matrix Λ bilden genau dann die duale Basis, wenn $\Lambda B = \mathbf{I}$.

Die duale Basis zur kanonischen Basis $\mathrm{e}_1, \ldots, \mathrm{e}_n$ des K^n sind die *Einheitszeilen* $\mathrm{e}_1^\top, \ldots, \mathrm{e}_n^\top$

Folgerung 6.1.1. *Wenn V endlichdimensional ist, hat V^* die gleiche Dimension wie V.* □

V und V^* sind also isomorph. Ein Isomorphismus f läßt sich zum Beispiel dadurch angeben, daß man die Elemente einer Basis $b_1, \ldots, b_n$ auf die Elemente $\lambda_1, \ldots, \lambda_n$ der dualen Basis abbildet. f hängt allerdings von der Wahl der Basis $b_1, \ldots, b_n$ ab. Man sagt, daß V und V^* nicht *kanonisch* isomorph sind.

Sei nun V beliebig (nicht notwendig endlichdimensional). Jedes Element x von V definiert vermöge $\lambda \mapsto \lambda(x)$ eine lineare Abbildung $\Phi(x)$ von V^* nach K.

$$\Phi : V \to V^{**}$$

ist die *kanonische* Abbildung von V nach V^{**}.

Lemma 6.1.2. *Die kanonische Abbildung $\Phi : V \to V^{**}$ ist injektiv und linear.*

Wenn V endlichdimensional ist, sind also V und V^{**} kanonisch isomorph.

Beweis. Φ ist linear, weil

$$\Phi(x+y)\lambda = \lambda(x+y) = \lambda(x)+\lambda(y) = \Phi(x)\lambda + \Phi(y)\lambda = (\Phi(x)+\Phi(y))\lambda$$

und

$$\Phi(\alpha x)\lambda = \lambda(\alpha x) = \alpha\lambda(x) = \alpha(\Phi(x)\lambda) = (\alpha\Phi(x))\lambda$$

Wenn x ungleich Null ist, gibt es eine Linearform λ mit $\Phi(x)\lambda = \lambda(x) \neq 0$. Also ist $\Phi(x) \neq 0$ und Φ ist injektiv. □

Lemma 6.1.3. *Wechselt man eine Basis von V mittels der Matrix B, so wechseln die dualen Basen von V^* mit $(B^\top)^{-1}$.*

Beweis. $b_1, \dots, b_n$ und $b'_1, \dots, b'_n$ seien zwei Basen von V, deren Übergang von der Matrix B vermittelt wird. Der Übergang der dualen Basen $(\lambda_1, \dots, \lambda_n)$ und $(\lambda'_1, \dots, \lambda'_n)$ sei vermittelt durch X.

Daß $\lambda_1, \dots, \lambda_n$ dual zu $b_1, \dots, b_n$ ist, schreiben wir als

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} (b_1 \cdots b_n) = \mathbf{I}.$$

Wir haben dann

$$\mathbf{I} = \begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_n \end{pmatrix} (b'_1 \cdots b'_n) = X^\top \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} (b_1 \cdots b_n) B = X^\top B.$$

Also ist $X = (B^\top)^{-1}$. □

6.2 Duale Abbildungen

Eine lineare Abbildung $f : V \to W$ induziert vermöge $f^*(\lambda) = \lambda \circ f$ eine lineare Abbildung

$$f^* : W^* \to V^*,$$

die zu f *duale* Abbildung.

Lemma 6.2.1. *Dualisieren ist ein kontravarianter, linearer Funktor: Es gilt*

a) $(\mathrm{id}_V)^* = \mathrm{id}_{V^*}$

b) Wenn $f : U \to V$ und $g : V \to W$ linear sind, ist $(g \circ f)^* = f^* \circ g^*$.

c) Wenn f und g lineare Abbildungen von V nach W sind, ist $(f+g)^* = f^* + g^*$

d) Für alle $\alpha \in K$ ist $(\alpha f)^* = \alpha f^*$.

Außerdem ist der Funktor treu*:*

$$f^* = 0 \Rightarrow f = 0.$$

Für endlichdimensionale Vektorräume ist der Funktor voll*: Für jedes $g : W^* \to V^*$ gibt es ein $f : V \to W$ mit $f^* = g$.*

Beweis. Die ersten vier Behauptungen ergeben sich aus einfachen Rechnungen. Für c) und d) verwenden wir Lemma 2.2.2:

a)

$$(\mathrm{id}_V)^*(\lambda) = \lambda \circ \mathrm{id}_V = \lambda = \mathrm{id}_{V^*}(\lambda)$$

b)

$$(g \circ f)^*(\lambda) = \lambda \circ g \circ f = g^*(\lambda) \circ f = f^*(g^*(\lambda)) = (f^* \circ g^*)(\lambda)$$

c)

$$(f + g)^*(\lambda) = \lambda \circ (f + g) = \lambda \circ f + \lambda \circ g = f^*(\lambda) + f^*(\lambda) = (f^* + g^*)(\lambda)$$

d)

$$(\alpha f)^*(\lambda) = \lambda \circ (\alpha f) = \alpha(\lambda \circ f) = \alpha f^*(\lambda) = (\alpha f^*)(\lambda)$$

Wenn $f^* = 0$ und $x \in V$, ist für alle $\lambda \in W^*$

$$\lambda(f(x)) = f^*(\lambda)(x) = 0.$$

Es folgt $f(x) = 0$. Also ist $f = 0$, wenn $f^* = 0$.

Die lineare Abbildung

$$* : \mathrm{L}(V, W) \to \mathrm{L}(W^*, V^*)$$

ist injektiv. Wenn V und W endlichdimensional sind, haben beide Räume die Dimension $\dim(V) \cdot \dim(W)$ und $*$ muß auch surjektiv sein.[1]

□

Identifiziert man wie oben die Elemente von K^m und K^n mit Spaltenvektoren und die Elemente der Dualräume mit Zeilenvektoren, so wird für eine m–n–Matrix A die Abbildung $f = f_A : K^n \to K^m$ gegeben durch

$$f : x \mapsto Ax$$

[1]Man sieht leicht, daß es genügt, anzunehmen, daß W endlichdimensional ist.

und f^* durch

$$f^* : \lambda \mapsto \lambda A,$$

weil

$$(\lambda A)x = \lambda(Ax).$$

Daraus ergibt sich:

Satz 6.2.2. *V und W seien endlichdimensional und $f : V \to W$ linear. Wenn f bezüglich der Basen $\mathfrak{B}$ und $\mathfrak{C}$ durch die Matrix A gegeben ist, ist f^* bezüglich der dualen Basen durch die Transponierte $A^\top$ gegeben.* □

Man kann den letzten Satz auch so beweisen: Die dualen Basen seien $\mathfrak{B}^*$ und $\mathfrak{C}^*$, als *Spalten* geschrieben:

$$\mathfrak{B}^*\mathfrak{B} = \mathbf{I} \quad \text{und} \quad \mathfrak{C}^*\mathfrak{C} = \mathbf{I}.$$

Und f^* sei durch X gegeben (vgl. (2.1), Seite 27):

$$f^*(\mathfrak{C}^*) = X^\top \mathfrak{B}^*$$

Dann folgt

$$X^\top = X^\top \mathfrak{B}^*\mathfrak{B} = f^*(\mathfrak{C}^*)\mathfrak{B} = \mathfrak{C}^* f(\mathfrak{B}) = \mathfrak{C}^*\mathfrak{C}A = A.$$

Folgerung 6.2.3. *Für Endomorphismen von endlichdimensionalen Vektorräumen gilt* $\det(\phi^*) = \det(\phi)$.

Beweis. Satz 4.3.5. □

Lemma 6.2.4. *Der Dualraum einer direkten Summe läßt sich auf natürliche Weise mit der direkten Summe der Dualräume identifizieren:*

$$(V_1 \oplus V_2)^* = V_1^* \oplus V_2^*.$$

Das Duale der Einbettung $V_1 \to V_1 \bigoplus V_2$ ist die Projektion $V_1^* \bigoplus V_2^* \to V_1^*$. Das Duale der Projektion $V_1 \bigoplus V_2 \to V_1$ ist die Einbettung $V_1^* \to V_1^* \bigoplus V_2^*$.

Beweis. Die erste Behauptung ist klar, weil eine lineare Abbildung $\lambda : V_1 \oplus V_2 \to K$ bestimmt ist durch die beiden Einschränkungen $\lambda_i = \lambda \restriction V_i$. Die Abbildung

$$\lambda \to \lambda_1 \oplus \lambda_2$$

ist der gewünschte Isomorphismus.

Ein anderer Beweis ergibt sich aus den funktoriellen Eigenschaften von $*$. Ein Isomorphismus von W mit $V_1 \oplus V_2$ ist gegeben durch zwei Projektionen $\pi_i : W \to V_i$ und Einbettungen $\epsilon_i : V_i \to W$, die den Bedingungen

$$\begin{aligned} \pi_i \epsilon_i &= \mathrm{id}_{V_i} \quad (i = 1, 2) \\ \epsilon_1 \pi_1 + \epsilon_2 \pi_2 &= \mathrm{id}_W \end{aligned}$$

Wenn man dualisiert, ergibt sich

$$\begin{aligned} \epsilon_i^* \pi_i^* &= \mathrm{id}_{V_i^*} \quad (i = 1, 2) \\ \pi_1^* \epsilon_1^* + \pi_2^* \epsilon_2^* &= \mathrm{id}_{W^*} \end{aligned}$$

und W^* ist isomorph zu $V_1^* \oplus V_2^*$ vermöge der Projektionen $\epsilon_i^* : W^* \to V_i^*$ und der Einbettungen $\pi_i^* : V_i^* \to W^*$. □

Folgerung 6.2.5. *Sei $f : V \to W$ linear. Dann gilt:*

1. *f ist genau dann injektiv, wenn f^* surjektiv ist.*
2. *f ist genau dann surjektiv, wenn f^* injektiv ist.*
3. *Wenn W endlichdimensional ist, haben f und f^* den gleichen Rang.*

Beweis. Zuerst bemerken wir, daß f^* ein Isomorphismus ist, wenn f eine Isomorphismus ist. Wenn nämlich g ein Inverses von f ist, ist g^* ein Inverses von f^*.

Sei X ein Komplement von $\ker(f)$ in V und Y ein Komplement von $\mathrm{Im}(f)$ in W. Weil f einen Isomorphismus zwischen X und $\mathrm{Im}(f)$ induziert, erhalten wir f als die Komposition:

$$\ker(f) \oplus X \to X \to \mathrm{Im}(f) \to \mathrm{Im}(f) \oplus Y$$

und f^* als

$$\mathrm{Im}(f)^* \oplus Y^* \to \mathrm{Im}(f)^* \to X^* \to \ker(f)^* \oplus X^*,$$

wobei der mittlere Pfeil wieder ein Isomorphismus ist. Also ist

$$f \text{ injektiv} \Leftrightarrow \ker(f) = 0 \Leftrightarrow \ker(f)^* = 0 \Leftrightarrow f^* \text{ surjektiv}$$

und

$$f \text{ surjektiv} \Leftrightarrow Y = 0 \Leftrightarrow Y^* = 0 \Leftrightarrow f^* \text{ injektiv}.$$

Die letzte Behauptung folgt aus 6.1.1, weil der Rang von f die Dimension von X ist und der Rang von f^* die Dimension von X^* ist.

□

Eine Matrix A hat also immer den gleichen Rang wie ihre Transponierte $A^\top$. Weil der Rang von $A^\top$ die Maximalzahl linear unabhängiger Zeilen von A ist, läßt sich unser Ergebnis formulieren als

$$\textit{Spaltenrang} = \textit{Zeilenrang}.$$

Einen Alternativbeweis liefert Folgerung 3.3.5. Wenn für reguläre B und B' das Produkt $B'AB$ die Gestalt (3.4) auf Seite 50 hat, ist

$$\mathrm{Rang}(A) = \mathrm{Rang}(B'AB) = \mathrm{Rang}(B'AB)^\top = \mathrm{Rang}(B^\top A^\top B'^\top) = \mathrm{Rang}(A^\top).$$

6.3 Duale Paare

V und W seien endlichdimensionale K–Vektorräume und

$$(\ ,\ ) : V \times W \to K$$

bilinear.

Sei $\mathfrak{B} = (b_1, \ldots, b_m)$ eine Basis von V und $\mathfrak{C} = (c_1, \ldots, c_n)$ eine Basis von W. Wie in Lemma 2.6.1 sieht man, daß $(\ ,\ )$ durch die m–n–Matrix

$$B = (b_i, c_j)$$

eindeutig bestimmt ist: Wenn wir die Koordinaten von v und w als Spaltenvektoren x und y schreiben, ergibt sich (in äußerst gewagter Notation):

$$(v, w) = (x^\top \mathfrak{B}^\top, \mathfrak{C}y) = x^\top(\mathfrak{B}^\top, \mathfrak{C})y = x^\top By.$$

Die auf $K^m \times K^n$ bezüglich der kanonischen Basen durch B gegebene Bilinearform bezeichnen wir mit $(\ ,\ )_B$. Faßt man die Elemente von K^m und K^n als Spaltenvektoren auf, ist

$$(x, y)_B = x^\top By$$

Lemma 6.3.1. *Wenn man mit der Matrix D von $\mathfrak{B}$ zur Basis $\mathfrak{B}'$ wechselt (siehe Abschnitt 3.3) und mit E von $\mathfrak{C}$ zur Basis $\mathfrak{C}'$, wird $(\ ,\ )$ bezüglich der neuen Basen durch*

$$B' = D^\top BE$$

beschrieben.

Beweis.

$$(\mathfrak{B}'^\top, \mathfrak{C}') = ((\mathfrak{B}D)^\top, \mathfrak{C}E) = (D^\top\mathfrak{B}^\top, \mathfrak{C}E) = D^\top(\mathfrak{B}^\top, \mathfrak{C})E = D^\top BE$$

□

Der Rang der darstellenden Matrix von $(\ ,\ )$ hängt also von der Basiswahl nicht ab. Man nennt ihn den *Rang von* $(\ ,\ )$.

Definition. *Das Tripel* $V, W, (\ ,\ )$ *heißt duales Paar, wenn*

$$\dim(V) = \dim(W) = \textit{Rang von } (\ ,\ ).$$

Wenn V, W ein duales Paar bilden, dann auch W, V. Dabei muß man natürlich von $(\ ,\ )$ zur Bilinearform $(y, x)' = (x, y)$ übergehen.

Setzt man für endlichdimensionales W $\quad V = W^*$ und $(\lambda, x) = \lambda(x)$, bilden V und W ein duales Paar. Das nächste Lemma zeigt, daß jedes duale Paar so aussieht.

Lemma 6.3.2. *Bilineare Abbildungen* $(\ ,\ ) : V \times W \to K$ *und lineare Abbildungen* $\Phi : V \to W^*$ *entsprechen einander vermöge*

$$\Phi(v)(w) = (v, w) \tag{6.1}$$

Wenn W endlichdimensional ist, bilden V und W genau dann ein duales Paar, wenn Φ ein Isomorphismus ist.

Beweis. Es ist klar, daß (6.1) eine Bijektion zwischen allen Abbildungen $(\ ,\ ) : V \times W \to K$ und allen Abbildungen $\Phi : V \to \{\lambda \mid \lambda : W \to K\}$ definiert. Daß $(\ ,\ )$ im zweiten Argument linear ist, bedeutet, daß alle $\Phi(v)$ in W^* liegen. $(\ ,\ )$ ist genau dann linear im ersten Argument, wenn Φ linear ist.

Zum Beweis des zweiten Teils der Behauptung, können wir annehmen, daß $V = K^m$, $W = K^n$ und $(\ ,\ ) = (\ ,\ )_B$. Dann ist

$$(v, w) = v^\top B w = (B^\top v)^\top w.$$

Die Linearform $\Phi(v)$ wird also durch die Zeile $(B^\top v)^\top$ dargestellt. Daraus folgt, daß $\Phi : K^m \to K^n \cong (K^n)^*$ die Matrix $B^\top$ hat.

Φ ist genau dann ein Isomorphismus, wenn $n = m$ und $B^\top$ (das heißt B) regulär ist. Das war zu zeigen. $\square$

Folgerung 6.3.3. *V und W seien endlichdimensional und $(\ ,\ ) : V \times W \to K$ bilinear. Dann bilden V und W genau dann ein duales Paar, wenn $(\ ,\ )$* nicht–ausgeartet *ist. Das heißt, daß für alle v und w*

$$\begin{aligned} \forall w'\,(v, w') = 0 \;&\Rightarrow\; v = 0 \\ \forall v'\,(v', w) = 0 \;&\Rightarrow\; w = 0 \end{aligned}$$

Beweis. Die erste Bedingung bedeutet, daß Φ injektiv ist, die zweite, daß die duale Abbildung $\Psi : W \to V^*$, definiert durch $\Psi(w)(v) = (v, w)$, injektiv ist. Wenn (V, W) ein duales Paar ist, sind Φ und Ψ Isomorphismen. Wenn umgekehrt Φ und Ψ injektiv sind, folgt $\dim(V) \leq \dim(W^*) = \dim(W)$ und $\dim(W) \leq \dim(V^*) = \dim(V)$. Also ist $\dim(V) = \dim(W)$ und Φ ist als injektive lineare Abbildung zwischen zwei Räumen gleicher endlicher Dimension ein Isomorphismus.

Mit Spalten und Matrizen kann man auch so argumentieren: Die Bilinearform sei durch B gegeben. Sei v eine Spalte. Dann gilt $\forall w'\ v^\top Bw' = 0$ genau dann, wenn $v^\top B = 0$. Die erste Bedingung bedeutet also

$$v^\top B = 0 \ \Rightarrow\ v = 0$$

und die zweite Bedingung

$$Bw = 0 \ \Rightarrow\ w = 0.$$

Anders ausgedrückt: Zeilen und Spalten von B sind jeweils linear unabhängig. Das ist genau dann der Fall, wenn B quadratisch und regulär ist. □

Sei nun V und W ein duales Paar. Mit Hilfe des Isomorphismus Φ lassen sich die Begriffsbildungen der Abschnitte 6.1 und 6.2 übertragen:

Wir nennen zwei Basen $\mathfrak{B} = (b_1, \ldots, b_m)$ und $\mathfrak{C} = (c_1, \ldots, c_n)$ von Vund W *dual*, wenn

$$(b_i, c_j) = \delta_{ij}.$$

Wenn $\mathfrak{C}$ vorgegeben ist und $(\lambda_1, \ldots, \lambda_n)$ die zu $\mathfrak{C}$ duale Basis von W^*, gewinnt man $\mathfrak{B}$ durch $b_i = \Phi^{-1}(\lambda_i)$. Man sieht leicht, daß ein duales Basenpaar nur existieren kann, wenn V und W ein duales Paar bilden.

Sei nun f ein Endomorphismus von W. Dann überträgt sich f^* zu einem Endomorphismus f^t von V durch

$$f^t = \Phi^{-1} \circ f^* \circ \Phi.$$

Man nennt f^t den zu f *adjungierte* Endomorphismus. Wie in 6.2.2 wird f^t bzgl. der dualen Basis durch die transponierte Matrix dargestellt.

Lemma 6.3.4. *f^t ist bestimmt durch die Gültigkeit der Gleichung*

$$(f^t(v), w) = (v, f(w)).$$

Beweis. Die angegebene Bedingung ist nichts anderes als die Übertragung der Formel $f^*(\lambda)(w) = \lambda(f(w))$. Die genaue Rechnung ist

$$(f^t(v), w) = (\Phi^{-1} f^* \Phi(v), w) = (f^* \Phi(v))(w) = \Phi(v)(f(w)) = (v, f(w)).$$

□

Für Endomorphismen f von V definiert man f^t analog durch Vertauschen von Rechts und Links. Man hat dann

$$f^{tt} = f.$$

Wenn man in Matrizen denkt, bedeutet das nichts anderes als $(A^\top)^\top = A$. Man kann das aber auch matrizenfrei ausrechnen:

$$(f^{tt}(v), w) = (v, f^t(w)) = (f(v), w).$$

Definition. *Für ein duales Paar V, W und einen Unterraum U von V definieren wir den zu U orthogonalen Unterraum*

$$U^\perp = \{w \in W \mid \forall u \in U \ (u, w) = 0\}$$

Lemma 6.3.5.

1. $(U^\perp)^\perp = U$
2. *Mit der Bilinearform*
$$[v + U, w] = (v, w)$$
wird $V/U, U^\perp$ zu einem dualen Paar. Es folgt $\dim U + \dim(U^\perp) = \dim V$.

Beweis.
1): Es ist klar, daß U in $(U^\perp)^\perp$ enthalten ist. Wenn v nicht zu U gehört, gibt es eine Linearform, die auf U verschwindet und auf v den Wert 1 annimmt. (Wähle eine Linearform auf V/U, die auf $v + U$ den Wert 1 annimmt und komponiere sie mit der Projektion $V \to V/U$.) Dann gibt es auch ein $w \in W$ mit $(U, w) = 0$ und $(v, w) = 1$. w gehört also $U^\perp$ und daher v nicht zu $(U^\perp)^\perp$.

2): Wenn v und v' zur gleichen Nebenklasse von U gehören, ist für alle $w \in U^\perp$

$$(v', w) = (v, w) + (v' - v, w) = (v, w).$$

$[\,,\,] : V/U \times U^\perp \to K$ ist also wohldefiniert. Wir zeigen, daß $[\,,\,]$ nicht ausgeartet ist. Wenn $[v' + U, w] = 0$ für alle $v' \in V$, ist sofort klar, daß $w = 0$. Wenn $[v + U, w'] = 0$ für alle $w' \in U^\perp$, gehört v zu $(U^\perp)^\perp = U$. Also ist $v + U = 0$. $\square$

Kapitel 7

Symmetrische Bilinearformen

7.1 Bilinearformen

Wir betrachten Bilinearformen auf einem endlichdimensionalen Vektorraum:

$$(\ , \ ) : V \times V \to K$$

Die Bilinearform $(\ , \ )$ heißt *regulär* oder *nicht ausgeartet*, wenn sie (V, V) zu einem dualen Paar macht.

Wenn $(\ , \ )$ regulär ist, hat (wie in 6.3.4) jeder Endomorphismus f einen *adjungierten* Endomorphismus f^t, der gegeben ist durch

$$(f^t(u), v) = (u, f(v)). \tag{7.1}$$

Definition. *Ein Automorphismus von $(V, (\ , \ ))$ ist ein Automorphismus f von V, der mit $(\ , \ )$ verträglich ist, für den also*

$$(f(u), f(v)) = (u, v)$$

für alle $u, v \in V$.

Lemma 7.1.1. *Sei $(\ , \ )$ eine reguläre Bilinearform auf V. Dann ist f genau dann ein Automorphismus von $(V, (\ , \ ))$, wenn*

$$f^t \circ f = \ id_V.$$

Beweis. $(f(u), f(v)) = (f^t(f(u)), v) = (u, v)$ für alle $u, v \in V$ bedeutet $f^t \circ f = \mathrm{id}_V$. □

Bilinearformen auf K^n werden (wie auf Seite 90) gegeben durch quadratische Matrizen B als

$$(x,y)_B = x^\top By.$$

$(x,y)_B$ ist genau dann regulär, wenn B regulär ist.

Wenn $f = f_A$, hat f^t die Matrix ${B^\top}^{-1}A^\top B^\top$, weil

$$({B^\top}^{-1}A^\top B^\top x)^\top By = x^\top BAB^{-1}By = x^\top BAy.$$

Wir nennen

$$(x,y)_{\mathbf{I}} = x^\top y$$

die *Standardbilinearform* auf K^n. Die Standardbilinearform ist regulär. Sie stimmt auf $\mathbb{R}^n$ mit dem Skalarprodukt (Abschnitt 1.4) überein. Für die Standardbilinearform auf K^n ist

$$\big(f_A\big)^t = f_{A^\top}.$$

Lemma 7.1.2. *Sei $(\ ,\ )$ eine reguläre Bilinearform auf V. Dann hat jede Bilinearform $[\ ,\ ]$ auf V die Gestalt*

$$[u,v] = (u, f(v))$$

für einen eindeutig bestimmten Endomorphismus f.

Beweis. Die Behauptung gilt für beliebige duale Paare $(\ ,\ ) : V \times W \to K$. Sei $[\ ,\ ] : V \times W \to K$ bilinear und $\Phi : W \to V^*$ definiert durch $\Phi(w)v = (v,w)$ und $\Psi : W \to V^*$ durch $\Psi(w)v = [v,w]$. Weil $(\ ,\ )$ regulär ist, ist Φ ein Isomorphismus (6.3.2). Dann ist

$$[v,w] = \Psi(w)v = \Phi\Phi^{-1}\Psi(w)(v) = (v, \Phi^{-1}\Psi(w)).$$

Man setzt also $f = \Phi^{-1} \circ \Psi$. □

7.2 Symmetrische Bilinearformen

Im Folgenden sei K ein Körper, der nicht die Charakteristik 2 hat und V ein K–Vektorraum.

Definition. *Eine Bilinearform $(\ ,\ ) : V \times V \to K$ heißt symmetrisch, wenn $(v,w) = (w,v)$ für alle $v, w \in V$.*

Die Standardbilinearform des K^n ist symmetrisch.

$(\ ,\ )_B$ ist genau dann symmetrisch, wenn $\beta_{i,j} = \beta_{j,i}$ für alle i,j. Oder, anders ausgedrückt, wenn

$$B^\top = B.$$

Quadratische Matrizen mit dieser Eigenschaft heißen *symmetrisch*.

Lemma 7.2.1. *Sei $(\ ,\ )$ eine symmetrische reguläre Bilinearform auf V und f ein Endomorphismus. Dann ist die Bilinearform $(u, f(v))$ genau dann symmetrisch, wenn $f = f^t$.*

Beweis. Die gespiegelte Bilinearform $[u, v] = (v, f(u))$ gehört zu f^t, weil

$$\big(v, f(u)\big) = \big(f^t(v), u\big) = \big(u, f^t(v)\big).$$

□

Endomorphismen f mit $f^t = f$ heißen *selbstadjungiert*[1] .

Sei $(\ ,\ )$ eine symmetrische Bilinearform auf V. Die Funktion $q : V \to K$, definiert durch

$$q(u) = (u, u),$$

nennt man die zu $(\ ,\ )$ gehörende *quadratische Form*. $(\ ,\ )$ läßt sich durch q ausdrücken:

$$(u, v) = \frac{q(u + v) - q(u - v)}{4}. \tag{7.2}$$

Das folgt leicht aus der "binomischen" Formel

$$q(u \pm v) = q(u) \pm 2(u, v) + q(v).$$

Folgerung 7.2.2. *Eine symmetrische Bilinearform ist durch ihre quadratische Form eindeutig bestimmt.* □

Satz 7.2.3. *Symmetrische Bilinearformen auf endlichdimensionalen Vektorräumen lassen sich bei geeigneter Basiswahl durch Diagonalmatrizen darstellen.*

Eine andere Formulierung ist: Jeder endlichdimensionale Vektorraum mit einer symmetrischen Bilinearform hat eine *Orthogonalbasis*. Das ist eine Basis $a_1, \dots, a_n$, deren Elemente paarweise *orthogonal* sind:

$$(a_i, a_j) = 0, \text{ wenn } \quad i \neq j.$$

[1]Ob f selbstadjungiert ist, hängt natürlich von $(\ ,\ )$ ab.

Beweis. Sei n die Dimension von V. Wenn $q(v) = 0$ für alle $v \in V$, verschwindet die Bilinearform und jede Basis ist orthogonal. Wenn nicht, finden wir einen Vektor b_1 mit $q(b_1) \neq 0$. Sei

$$V_1 = \{v \in V \mid (v, b_1) = 0\}$$

der Kern der Linearform $v \to (v, b_1)$. V_1 hat die Dimension $n - 1$ (wegen 3.1.3). Weil b_1 nicht zu V_1 gehört, ist $V = Kb_1 \oplus V_1$. Wenn wir so fortfahren, erhalten wir eine orthogonale Basis $b_2, \ldots, b_n$ von V_1. $b_1, b_2, \ldots, b_n$ ist dann eine orthogonale Basis von V. □

Folgerung 7.2.4. *Wenn in K jedes Element Quadrat ist (wie zum Beispiel in $\mathbb{C}$), dann wird bei geeigneter Basiswahl jede symmetrische Bilinearform durch eine Matrix der Form*

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

dargestellt.

Die Zahl der vorkommenden Einsen ist der Rang der Bilinearform und also von der Wahl der Basis unabhängig.

Beweis. Sei $b_1, \ldots, b_n$ eine Basis, für die die Bilinearform Diagonalgestalt hat. Wir setzen dann

$$c_i = \begin{cases} \dfrac{b_i}{\sqrt{q(b_i)}} & \text{wenn } q(b_i) \neq 0 \\ b_i & \text{sonst} \end{cases}$$

Wenn wir die c_i noch so umsortieren, daß zuerst die Vektoren kommen, für die der erste Fall eintritt, haben wir die gewünschte Basis. □

Folgerung 7.2.5. *Über $\mathbb{R}$ läßt sich jede symmetrische Bilinearform bei geeigneter Basiswahl immer durch eine Diagonalmatrix*

$$\begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix} \tag{7.3}$$

darstellen, in deren Diagonalen nur 1, -1 und 0 vorkommen.

Beweis. Wir setzen jetzt

$$c_i = \begin{cases} \dfrac{b_i}{\sqrt{q(b_i)}} & \text{wenn } q(b_i) > 0 \\ \dfrac{b_i}{\sqrt{-q(b_i)}} & \text{wenn } q(b_i) < 0 \\ b_i & \text{sonst} \end{cases}$$

□

Sei p die Zahl der vorkommenden Einsen, q die Zahl der Minuseinsen, und r die Zahl der Nullen, die in der Diagonalen dieser Matrix vorkommen.

Satz 7.2.6 (Sylvester)**.** *p, q und r sind von der Wahl der diagonalisierenden Basis unabhängig.*

Man nennt $p - q$ die Signatur der Bilinearform. Weil $p + q$ der Rang der Bilinearform ist, ist der Satz von Sylvester äquivalent zur Invarianz der Signatur.

Beweis. Wir brauchen ein Definition:

Definition. *Eine symmetrische Bilinearform* (,) *heißt* positiv semidefinit*, wenn $q(x) = (x, x)$ niemals negativ wird. Wenn sogar $q(x)$ für alle von 0 verschiedenen x positiv ist, heißt* (,) positiv definit*. Analog definiert man* negativ (semi)definit*. Symmetrische Bilinearformen, die nicht semidefinit sind, nennt man* indefinit*.*

Für eine Diagonalmatrix B ist $(\ ,\ )_B$ genau dann positiv (semi)definit, wenn alle Diagonalelemente von B positiv (nicht-negativ) sind.

Sei nun (,) bezüglich der Basis $b_1, \ldots, b_n$ durch die Matrix (7.3) dargestellt. Dann ist p die größte Dimension eines Unterraums von V, auf dem (,) positiv definit ist. Denn einerseits ist (,) positiv definit auf dem p–dimensionalen Raum $\langle b_1, \ldots, b_p \rangle$. Wenn andererseits (,) positiv definit auf U ist, kann die Dimension von U nicht größer als p sein. Denn sonst hätte U nicht–trivialen Durchschnitt mit $\langle b_{p+1}, \ldots, b_n \rangle$, was nicht sein kann, weil dieser Raum negativ semidefinit ist. □

7.3 Euklidische Räume

Definition. *Eine euklidische Bilinearform ist eine positiv definite symmetrische Bilinearform auf einem endlichdimensionalen reellen Vektorraum. Ein euklidischer Raum ist ein reeller Vektorraum mit einer euklidischen Bilinearform.*

Das in 1.4 definierte Standardskalarprodukt auf dem $\mathbb{R}^n$ ist euklidisch. Eine euklidische Bilinearform ist immer regulär. Das folgt aus 6.3.3 wegen

$$\forall w' \, (v, w') = 0 \;\Rightarrow\; (v, v) = 0 \;\Rightarrow v = 0.$$

Aus 7.2.5 folgt:

Satz 7.3.1. *Jeder euklidische Vektorraum hat eine* Orthonormalbasis. *Das ist eine Basis $b_1, \ldots, b_m$, die zu sich selbst dual ist, für die also*

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

□

Es folgt, daß jeder euklidische Vektorraum V zu einem $\mathbb{R}^n$ mit dem Standardskalarprodukt isomorph ist.

Ausgehend von einer Basis $a_1, \ldots, a_n$ von V läßt sich mit dem *Schmidtschen Orthonormalisierungsverfahren* eine Orthonormalbasis $b_1, \ldots, b_n$ konstruieren: Wir setzen $b_1 = \dfrac{a_1}{\|a_1\|}$ und für $i = 2, \ldots, n$

$$b'_i = a_i - (a_i, b_1)b_1 - \ldots - (a_i, b_{i-1})b_{i-1} \tag{7.4}$$

und

$$b_i = \frac{b'_i}{\|b'_i\|}. \tag{7.5}$$

Wie in 1.4 ist dabei $\|a\|$ die Norm oder Länge $\sqrt{(a, a)}$ von a .

Man sieht durch Induktion sofort, daß für alle i

$$\langle b_1, \ldots, b_i \rangle = \langle a_1, \ldots, a_i \rangle.$$

Die b_i bilden also eine Basis. Daß die b_i die Länge 1 haben, folgt aus (7.5). Die Orthogonalität induktiv aus (7.4): Wenn $j < i$ ist

$$\begin{aligned}(b'_i, b_j) &= (a_i, b_j) - (a_i, b_1)(b_1, b_j) \cdots - (a_i, b_j)(b_j, b_j) \cdots - (a_i, b_{i-1})(b_{i-1}, b_j)\\ &= (a_i, b_j) - (a_i, b_j)\|b_j\| = 0.\end{aligned}$$

Lemma 7.3.2. *Sei $b_1, \ldots, b_n$ eine Orthonormalbasis. Dann ist für alle $a \in V$*

$$a = (a, b_1)b_1 + \ldots + (a, b_n)b_n.$$

Beweis. Die Diskussion des Schmidtschen Orthonormalisierungsverfahrens zeigt, daß alle b_i senkrecht auf $b' = a - (a, b_1)b_1 - \ldots - (a, b_n)b_n$ stehen. b' muß also Null sein. Einen alternativen Beweis erhält man durch den Ansatz

$$a = \alpha_1 b_1 + \cdots + \alpha_n b_n.$$

Man multipliziert mit b_j und erhält

$$(a, b_j) = \alpha_1(b_1, b_j) + \cdots + \alpha_n(b_n, b_j) = \alpha_j(b_j, b_j) = \alpha_j$$

$\square$

Lemma 7.3.3. *Sei U ein Unterraum des euklidischen Raums V. Dann ist V direkte Summe von U und dem* orthogonalen Komplement

$$U^\perp = \{w \in V \mid \forall u \in U\ (u, w) = 0\}.$$

Beweis. Sei v_0 ein beliebiger Vektor aus V. Weil $(\ ,\ )$ auf U regulär ist, gibt es ein eindeutig bestimmtes $u_0 \in U$ mit

$$(v_0, u) = (u_0, u)$$

für alle $u \in U$. Es folgt $(v_0 - u_0, u) = 0$ für alle $u \in U$. Das heißt $v_0 - u_0 \in U^\perp$. $\square$

Das Lemma folgt auch leicht aus dem Schmidtschen Verfahren. Wir wählen eine Basis $a_1, \ldots, a_n$, deren erste m Elemente U aufspannen. Wenn wir zu $b_1, \ldots, b_n$ orthonormalisieren, wird U von $b_1, \ldots, b_m$ und $U^\perp$ von $b_{m+1}, \ldots, b_n$ aufgespannt.

Eine weitere Beweisalternative ist die Verwendung der Gleichung[2]

$$\dim U + \dim(U^\perp) = \dim V$$

aus Lemma 6.3.5. Weil $U^\perp \cap U = 0$, folgt daraus die Behauptung.

Sei f ein Endomorphismus von V. Wir erinnern an die Definition (7.1) (Seite 94) des adjungierten Endomorphismus f^t.

Lemma 7.3.4. *Sei f ein Endomorphismus von V, der bezogen auf die Orthonormalbasis $\mathfrak{B}$ durch die Matrix A dargestellt wird. Dann wird der adjungierte Endomorphismus f^t durch die transponierte Matrix $A^\top$ dargestellt.*

Beweis. Das folgt aus der Bemerkung vor Lemma 6.3.4 oder direkt aus 6.2.2. Wir können auch Satz 7.3.1 verwenden und die hinter Lemma 7.1.1 gemachte Bemerkung, nach der die Behauptung für den $\mathbb{R}^n$ mit dem Standardskalarprodukt gilt.

Wir geben noch einmal einen direkten Beweis. Sei $\mathfrak{B}$ eine Orthonormalbasis, f und f^t bezüglich dieser Basis dargestellt durch A und A'. Nehmen wir an, daß die beiden Vektoren u und v in dieser Basis die Koeffizientenvektoren x und y haben[3]. Dann ist

$$(u, f(v)) = x^\top A y.$$

[2]Eigentlich brauchen wir nur die Ungleichung $U + \dim(U^\perp) \geq \dim V$, die für beliebige bilineare Abbildungen $(\ ,\ ) : V \times U \to K$ richtig ist.

[3]Das heißt $u = \mathfrak{B}x$ und $v = \mathfrak{B}y$

und

$$(f^t(u), v) = (A'x)^\top y = x^\top {A'}^\top y$$

Aus $x^\top {A'}^\top y = x^\top Ay$ für alle Spalten x, y folgt $A' = A^\top$. □

Automorphismen eines euklidischen Vektorraums heißen *orthogonale* Abbildungen.

Lemma 7.3.5. *Sei f ein Endomorphismus des euklidischen Vektorraums V, der bezüglich einer Orthonormalbasis $b_1, \dots, b_m$ durch die Matrix A dargestellt wird. Dann sind äquivalent:*

a) f ist orthogonal.

b) $f^t \circ f = id_V$.

c) $A^\top A = \mathbf{I}$.

d) $f(b_1), \dots, f(b_m)$ ist eine Orthonormalbasis.

Beweis. Die Äquivalenz von a), b) und c) folgt aus 7.1.1 und 7.3.4. c) besagt, daß die Spalten von A orthonormal sind. c) und d) sind also auch äquivalent. □

Für die Äquivalenz von a) und d) gebe ich noch einen eigenen Beweis: f ist genau dann ein Automorphismus, wenn die Bilinearform

$$(u, v)^f = (f(u), f(v))$$

mit (x, y) übereinstimmt[4]. Bezüglich $b_1, \dots, b_n$ hat aber $(\ ,\ )^f$ die Matrix

$$(f(b_i), f(b_j)).$$

Also ist f genau dann orthogonal, wenn $(f(b_i), f(b_j)) = \delta_{ij}$. Wir überlegen uns noch, daß eine orthonormale Familie (c_i) immer linear unabhängig ist:

$$0 = \sum_i \xi_i c_i \quad \Rightarrow \quad 0 = \Big(\sum_i \xi_i c_i, c_j\Big) = \xi_j.$$

Folgerung 7.3.6. *Orthogonale Abbildungen haben die Determinante 1 oder -1.*

[4] Wenn $(\ ,\ )^f = (\ ,\ )$, ist f injektiv, weil

$$u \neq 0 \Rightarrow (u, u) \neq 0 \Rightarrow (f(u), f(u)) \neq 0 \Rightarrow f(u) \neq 0.$$

Beweis. Aus 7.3.4 und 4.3.5 folgt $\det(f^t) = \det(f)$. Also ist mit 7.3.5

$$\det(f)^2 = \det(f^t)\det(f) = \det(f^t \circ f) = \det(\mathrm{id}_V) = 1.$$

□

Die *reelle orthogonale Gruppe*

$$\mathrm{O}_n(\mathbb{R}) = \left\{A \in \mathrm{M}_n(\mathbb{R}) \mid A^\top A = \mathbf{I}\right\}$$

ist die Gruppe aller orthogonalen n-n–Matrizen. Die orthogonalen Abbildungen mit Determinante 1 bilden die *spezielle* orthogonale Gruppe $\mathrm{SO}_n(\mathbb{R})$. $\mathrm{SO}_n(\mathbb{R})$ und das Komplement $\mathrm{O}_n(\mathbb{R}) \setminus \mathrm{SO}_n(\mathbb{R})$ sind Zusammenhangskomponenten von $\mathrm{O}_n(\mathbb{R})$. Weil sich jedes Element von $\mathrm{SO}_n(\mathbb{R})$ mit der Einheitsmatrix durch eine stetige Familie von orthogonalen Matrizen verbinden läßt, nennt man die Elemente von $\mathrm{SO}_n(\mathbb{R})$ *Drehungen*. Im Abschnitt 1.5 (ab Seite 15) haben wir die Drehungen der euklidischen Ebene $\mathbb{R}^2$ bestimmt.

Sei V ein n–dimensionaler euklidischer Vektorraum. Eine n–Form μ heißt *normiert*, wenn

$$\mu(b_1, \dots, b_n) \in \{-1, 1\}$$

für alle Orthonormalbasen $b_1, \dots, b_n$.

Lemma 7.3.7. *Ein euklidischer Vektorraum der Dimension n hat genau zwei normierte n–Formen. Die beiden Formen gehen durch Multiplikation mit -1 auseinander hervor.*

Beweis. Jede Volumenform μ, die auf (nur) einer Orthonormalbasis $b_1, \dots, b_n$ den Wert ± 1 hat, ist schon normiert. Denn, wenn $c_1, \dots, c_n$ eine andere Orthonormalbasis ist, ist die lineare Abbildung $f(b_i) = c_i$ orthogonal. Also ist (vgl. 4.5.2)

$$\mu(c_1, \dots, c_n) = \det(f)\,\mu(b_1, \dots, b_n) = \pm\mu(b_1, \dots, b_n) = \pm 1.$$

Die Behauptung folgt jetzt aus 4.3.1. □

Man nennt (für eine normierte Volumenform μ_0)

$$|\,\mu_0(a_1, \dots, a_n)|$$

das *Volumen* des Parallelepipeds $\mathrm{PE}(a_1, \dots, a_n)$. Für den $\mathbb{R}^n$ mit Skalarprodukt ist die Standardvolumenform normiert. Also stimmt unsere Definition mit 4.5.1 überein.

Man beachte, daß die Bilinearform eines euklidischen Raumes keine Orientierung festlegt, aber auf jedem Unterraum U wieder eine euklidische Struktur [5]

[5] mit der auf $U \times U$ eingeschränkten Bilinearform

induziert. Eine Volumenform auf V liefert uns eine Orientierung, kann aber auf Unterräumen keine Struktur induzieren.

Satz 7.3.8 (Gramsche Determinante)**.** *Das Volumen des von $a_1, \dots, a_n$ aufgespannten Parallelepipeds ist die Quadratwurzel der* Gramschen *Determinante*

$$\det \begin{pmatrix} (a_1, a_1) & (a_1, a_2) & \cdots & (a_1, a_n) \\ (a_2, a_1) & (a_2, a_2) & \cdots & (a_2, a_n) \\ \vdots & \vdots & & \vdots \\ (a_n, a_1) & (a_n, a_2) & \cdots & (a_n, a_n) \end{pmatrix}$$

Beweis. Wir können annehmen, daß wir uns im $\mathbb{R}^n$ mit dem Standardskalarprodukt befinden. Die Spalten a_i fassen wir zu einer Matrix A zusammen. Die Gramsche Determinante ist die Determinante von

$$\left(a_i^\top a_j\right) = A^\top A,$$

also gleich dem Quadrat von $\det(A)$. Der Betrag von $\det(A)$ ist aber das Volumen von $\mathrm{PE}(a_1, \dots, a_n)$. □

Zum Beispiel ist die Fläche des von $a_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ und $a_2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$ aufspannten Parallelogramms[6] die Quadratwurzel von

$$\begin{vmatrix} (a_1, a_1) & (a_1, a_2) \\ (a_2, a_1) & (a_2, a_2) \end{vmatrix} = \begin{vmatrix} 14 & 32 \\ 32 & 77 \end{vmatrix} = 54$$

7.4 Die Hauptachsentransformation

Sei V ein euklidischer Vektorraum. Der Satz von der Hauptachsentransformation erscheint in drei äquivalenten Formen.

Satz 7.4.1 (Hauptachsen einer symmetrischen Bilinearform)**.** *Sei* [,] *eine symmetrische Bilinearform auf V. Dann gibt es eine Orthonormalbasis von V, bezüglich der* [,] *durch eine Diagonalmatrix dargestellt wird.*

Man nennt die Elemente der diagonalisierenden Orthonormalbasis die *Hauptachsen* von [,].

Weil sich vermöge

$$[x, y] = (x, f(y))$$

symmetrische Bilinearformen und selbstadjungierte Endomorphismen entsprechen (7.1.2 und 7.2.1), ist der Satz gleichbedeutend mit

[6] Also sein Volumen in dem Unterraum, den a_1 und a_2 aufspannen.

Satz 7.4.2 (Hauptachsen eines selbstadjungierten Endomorphismus)**.** *Sei f ein selbstadjungierter Endomorphismus. Dann gibt es eine Orthonormalbasis von V aus Eigenvektoren von f.*

Für Matrizen formuliert:

Satz 7.4.3 (Hauptachsen einer symmetrischen Matrix)**.** *Zu jeder symmetrischen reellen Matrix A gibt es eine orthogonale Matrix D, sodaß $D^\top AD$ Diagonalgestalt hat.*

Im $\mathbb{R}^n$ werden symmetrische Bilinearformen durch symmetrische Matrizen A beschrieben. Wenn man mit D einen Basiswechsel vornimmt, ändert sich A (nach 6.3.1) zu $D^\top AD$. Die neue Basis ist genau dann orthonormal, wenn D orthogonal ist. Damit ist die Äquivalenz von 7.4.1 und 7.4.3 gezeigt.

7.4.2 bedeutet im $\mathbb{R}^n$, daß für symmetrische A der Endomorphismus f_A durch einen orthonormalen Basiswechsel in Diagonalgestalt gebracht werden kann. Wenn D den Basiswechsel beschreibt, bedeutet das, daß $D^{-1}AD$ Diagonalgestalt hat. Weil $D^{-1} = D^\top$, ist das wieder äquivalent zu 7.4.3.

Beweis. Ich gebe zwei unabhängige Beweise für 7.4.1 und 7.4.2. Wir können dabei ohne weiteres annehmen, daß V mindestens die Dimension 2 hat.

Sei $[\ ,\ ]$ eine eine symmetrische Bilinearform auf $\mathbb{R}^n$.

Behauptung 1: Es gibt einen Vektor x_1 der Länge 1, für den

$$(x_1, x) = 0 \ \Rightarrow\ [x_1, x] = 0.$$

Beweis: Die stetige Funktion $[x, x]$ nimmt auf der Einheitskugel

$$S = \{x \in \mathbb{R}^n \mid (x, x) = 1\}$$

ein Maximum an. Sagen wir bei x_1. Dann ist das Differential

$$\mathrm{d}[x, x] = [\mathrm{d}x, x] + [x, \mathrm{d}x] = 2[x, \mathrm{d}x]$$

Null an jedem Tangentenvektor X von S an der Stelle x_1. Das heißt, daß für alle X

$$(x_1, X) = 0 \ \Rightarrow\ [x_1, X] = 0. \tag{7.6}$$

Wenn man Differentialrechnung vermeiden will, kann man so argumentieren: Betrachte die symmetrische Bilinearform $[x, y]' = \lambda \cdot (x, y) - [x, y]$, wobei $\lambda = [x_1, x_1]$. $[\ ,\ ]'$ ist positiv semidefinit; nach Folgerung 7.2.5 gibt es eine Basis $b_1, \dots, b_n$ mit $[b_i, b_j]' = 0$ für $i \neq j$, $[b_1, b_1]' = \cdots = [b_p, b_p]' = 1$ und $[b_{p+1}, b_{p+1}]' = \cdots = [b_n, b_n]' = 0$. Weil $[x_1, x_1]' = 0$, liegt x_1 in $\langle b_{p+1}, \dots b_n\rangle$. Damit ist $[x_1, y]' = 0$ für alle y, woraus (7.6) folgt.

Wir können induktiv annehmen, daß das orthogonale Komplement

$$x_1^\perp = \{x \mid (x_1, x) = 0\}$$

eine Orthonormalbasis $x_2, \ldots, x_n$ hat, für die $[\ ,\ ]$ Diagonalgestalt hat. Weil $[x_1, x_2] = \ldots = [x_1, x_n] = 0$, ist $x_1, x_2, \ldots, x_n$ die gesuchte Orthonormalbasis von $\mathbb{R}^n$. Damit ist 7.4.1 bewiesen.

Sei nun f ein selbstadjungierter Endomorphismus von V.

Behauptung 2: f hat einen (reellen) Eigenwert.

Beweis: Sei $V = \mathbb{R}^n$ und $f = f_A$. Weil $\mathbb{C}$ algebraisch abgeschlossen ist, hat das charakteristische Polynom von A eine komplexe Nullstelle λ. λ ist ein Eigenwert der *komplexen* Matrix A. Also gibt es ein $x_1 \in \mathbb{C}^n \setminus \mathbf{0}$, mit $Ax_1 = \lambda x_1$. Durch Konjugation (Seite 112) erhalten wir $A\overline{x_1} = \overline{\lambda}\overline{x_1}$. Weil

$$(x_1, \overline{x_1}) = \sum_{i=1}^{n} \xi_i \overline{\xi_i}$$

als Summe der Betragsquadrate[7] der Koeffizienten ξ_i von x_1 ungleich Null ist, und weil A symmetrisch ist, liefert das nächste Lemma $\lambda = \overline{\lambda}$.

Lemma (A)**.** *α sei Eigenwert von f und β Eigenwert von f^t. Wenn $\alpha \neq \beta$, stehen die jeweiligen Eigenvektoren aufeinander senkrecht.*

Beweis. Das Lemma gilt für beliebige duale Paare $(\ ,\ ) : V \times W \to K$ und Endomorphismen $f : V \to V$: Sei $f(v) = \alpha v$ und $f^t(w) = \beta w$. Dann ist

$$\alpha(v, w) = (\alpha v, w) = (f(v), w) = (v, f^t(w)) = (v, \beta w) = \beta(v, w).$$

Wenn $\alpha \neq \beta$, folgt $(v, w) = 0$. □

Damit ist die Behauptung 2 bewiesen. Sei x_1 ein Eigenvektor von f der Länge 1. $\mathbb{R}x_1$ ist invariant unter f, also ist nach dem nächsten Lemma $x_1^\perp$ invariant unter $f^t = f$. Wir können wieder annehmen, daß $x_1^\perp$ eine Orthonormalbasis $x_2, \ldots, x_n$ aus Eigenvektoren von f hat. $x_0, x_1, \ldots, x_n$ ist dann die gesuchte Orthonormalbasis von $\mathbb{R}^n$. Damit ist 7.4.2 bewiesen. □

Lemma (B)**.** *Wenn U invariant ist unter f, ist $U^\perp$ invariant unter f^t.*

Beweis. Das Lemma gilt allgemein für duale Paare. Sei $w \in U^\perp$. Dann ist für alle $u \in U$

$$(u, f^t(w)) = (f(u), w) = 0.$$

Also $f^t(w) \in U^\perp$. □

[7] Der Betrag einer komplexe Zahl $z = \alpha + \beta\mathrm{i}$ ist die Wurzel aus $z\overline{z} = \alpha^2 + \beta^2$.

Eine Folgerung ist zum Beispiel:

Satz 7.4.4. *Für jeden Endomorphismus f gibt es eine Orthonormalbasis $b_1, \ldots, b_n$ von V, sodaß $f(b_1), \ldots, f(b_n)$ paarweise orthogonal sind.*

Beweis. Wir wählen einfach einen Orthonormalbasis, die die symmetrische Bilinearform

$$[u, v] = (f(u), f(v))$$

diagonalisiert. □

Folgerung 7.4.5 (Polarisationssatz)**.** *Jeder Endomorphismus von V ist ein Produkt einer orthogonalen Abbildung und einer selbstadjungierten Abbildung.*

Beweis. Sei f ein Endomorphismus und $b_1, \ldots, b_n$ eine Orthonormalbasis, die von f auf ein orthogonales System abgebildet wird. Wir finden dann eine Orthonormalbasis $c_1, \ldots, c_n$ und reelle Zahlen λ_i, für die

$$f(b_i) = \lambda_i c_i.$$

Sei nun g die (orthogonale) lineare Abbildung, die b_i auf c_i abbildet und h die (selbstadjungierte) Abbildung, die c_i auf $\lambda_i c_i$ abbildet. Dann ist $f = hg$. □

Nehmen wir an, daß bezüglich einer Orthonormalbasis die symmetrische Bilinearform $[u, v] = (u, f(v))$ durch die Diagonalmatrix

$$\begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

gegeben ist. Durch Permutation der Basisvektoren können wir immer erreichen, daß $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$. Die λ_i sind dann als die Eigenwerte von f eindeutig bestimmt. Der nächste Satz zeigt die Bedeutung der λ_i für die Bilinearform $[\ ,\ ]$.

Satz 7.4.6.

$$\lambda_m = \min\big\{\max\{[x, x] \mid x \in U,\ \|x\| = 1\} \ \big|\ U \leq V,\ \dim U = m\big\}$$

Beweis. Der Satz wird ähnlich bewiesen wie der Satz von Sylvester (7.2.6). Wir können annehmen, daß $V = \mathbb{R}^n$ und daß $[\ ,\ ]$ in der kanonischen Basis die angegebene Diagonalgestalt hat. Sei U_0 aufgespannt von der ersten m Basisvektoren und $x = (\xi_1, \ldots, \xi_m, 0, \ldots, 0)$ ein Element von U_0 der Länge 1. Dann ist

$$[x, x] = \lambda_1 \xi_1^2 + \ldots + \lambda_m \xi_m^2 \leq \lambda_m (\xi_1^2 + \ldots + \xi_m^2) = \lambda_m.$$

Weil andererseits $[\mathrm{e}_m, \mathrm{e}_m] = \lambda_m$, folgt

$$\lambda_m = \max\{[x, x] \mid x \in U_0,\ \|x\| = 1\}.$$

Wir müssen noch zeigen, daß dieses Maximum für keinen m–dimensionalen Unterraum U kleiner werden kann: Ein solches U schneidet den $(n-m+1)$-dimensionalen Raum $U_1 = \langle \mathrm{e}_m, \ldots, \mathrm{e}_n \rangle$ immer nicht–trivial. Es gibt also in $U \cap U_1$ ein x der Länge 1. Weil

$$\lambda_m = \min\{[x,x] \mid x \in U_1,\ \|x\| = 1\},$$

folgt $[x,x] \geq \lambda_m$ und daraus die Behauptung. $\square$

Quadriken

Eine *Quadrik* im $\mathbb{R}^n$ ist die Lösungsmenge

$$Q = \{(\xi_1, \ldots, \xi_n) \mid q(\xi_1, \ldots, \xi_n) = 0\}$$

einer quadratischen Gleichung. Dabei ist q ein Polynom zweiten Grades

$$q(x_1, \ldots, x_n) = \sum_{i,j} \alpha_{ij} x_i x_j + \sum_i \alpha_i x_i + \alpha.$$

Wir können ohne weiteres annehmen, daß $A = (\alpha_{ij})$ symmetrisch ist. Wir setzen $a = (\alpha_1 \ldots \alpha_n)^\top$. Dann schreibt sich unsere quadratische Gleichung in $x = (\xi_1 \ldots \xi_n)^\top$ als

$$x^\top A x + a^\top x + \alpha = 0$$

oder, in einem abstrakten euklidischen Raum für ein selbstadjungiertes f, als

$$(x, f(x)) + (a, x) + \alpha = 0. \tag{7.7}$$

Wir wollen alle Quadriken klassifizieren, wobei wir Quadriken, die durch eine *euklidische affine* Abbildung[8]

$$x \mapsto Bx + b$$

(B orthogonal) ineinander übergehen, nicht unterscheiden.

Die Hauptunterscheidung ist, ob a im Bild von f liegt oder nicht.

<u>Fall 1</u> $a \in \operatorname{Im} f$.

Die Verschiebung des Nullpunkts $x = b + y$ transformiert (7.7) nach

$$(y, f(y)) + (2f(b) + a, y) + (b, f(b)) + (a, b) + \alpha = 0.$$

Wir wählen b so, daß $2f(b) + a = 0$. Der lineare Term fällt jetzt weg und wir erhalten eine Gleichung der Form $(x, f(x)) + \alpha = 0$. Wir können die Gleichung noch mit einem geeigneten Faktor multiplizieren und bekommen $\alpha = 0$ oder $\alpha =$

[8] *Affine* Abbildungen sind Abbildungen der Form $x \mapsto Bx + b$ für beliebige Matrizen B.

-1. Wenn wir eine geeignete Orthonormalbasis wählen, hat f Diagonalgestalt. Wir schreiben die positiven Eigenwerte als $\frac{1}{a_i^2}$ und die negativen als $-\frac{1}{b_j^2}$. Die Gleichungen in den neuen Koordinaten sind dann die Normalformen

$$\left(\frac{\xi_1}{a_1}\right)^2 + \cdots + \left(\frac{\xi_p}{a_p}\right)^2 - \left(\frac{\zeta_1}{b_1}\right)^2 - \cdots - \left(\frac{\zeta_q}{b_q}\right)^2 = 1 \tag{7.8}$$

$$\left(\frac{\xi_1}{a_1}\right)^2 + \cdots + \left(\frac{\xi_p}{a_p}\right)^2 - \left(\frac{\zeta_1}{b_1}\right)^2 - \cdots - \left(\frac{\zeta_q}{b_q}\right)^2 = 0 \tag{7.9}$$

Fall 2 $a \notin \operatorname{Im} f$.
Wir wählen b so, daß $2f(b) + a$ senkrecht auf dem Bild von f steht. Wir können dadurch annehmen, daß in (7.7) a senkrecht auf dem Bild von f steht[9]. Durch Multiplikation erreichen wir, daß a die Länge 1 hat. Wir verschieben (noch einmal) um $-\alpha a$:

$$(y, f(y)) + (-\alpha 2f(a) + a, y) + \alpha^2(a, f(a)) - \alpha(a, a) + \alpha = 0.$$

Weil $f(a) = 0$, ergibt sich die Gleichung

$$(y, f(y)) + (a, y) = 0.$$

Wir diagonalisieren f mit einer Orthonormalbasis, die wir so wählen können, daß sie $-a$ enthält. Die resultierende Gleichung ist die dritte Normalform

$$\left(\frac{\xi_1}{a_1}\right)^2 + \cdots + \left(\frac{\xi_p}{a_p}\right)^2 - \left(\frac{\zeta_1}{b_1}\right)^2 - \cdots - \left(\frac{\zeta_q}{b_q}\right)^2 = \eta. \tag{7.10}$$

Der projektive Raum

Sei K Körper und V ein Vektorraum über K. Der zugeordnete *projektive* Raum $\mathrm{P}(V)$ ist die Menge aller eindimensionalen Unterräume von V:

$$\mathrm{P}(V) = \{\langle v\rangle \mid v \in V \setminus \mathbf{0}\}$$

Wir definieren

$$\dim(\mathrm{P}(V)) = \dim(V) - 1.$$

Unterräume von $\mathrm{P}(V)$ sind Teilmengen der Form $\mathrm{P}(U)$ für Unterräume $U \leq V$.

[9] Das ist gleichbedeutend mit $a \in \operatorname{Ker} f$. Man rechnet leicht nach, daß in dualen Paaren

$$(\operatorname{Im} f)^{\perp} = \operatorname{Ker}(f^t).$$

$$
\begin{array}{lrl}
n=1: & x^2 = 0 & \text{Ein Punkt } (x=0) \\
 & \left(\frac{x}{a}\right)^2 = 1 & \text{Zwei Punkte } (x=\pm a) \\
n=2: & \left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2 = 0 & \text{Zwei Geraden } \left(\frac{x}{a} = \frac{y}{b} \text{ und } \frac{x}{a} = -\frac{y}{b}\right) \\
 & \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1 & \text{Ellipse} \\
 & \left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2 = 1 & \text{Hyperbel} \\
 & \left(\frac{x}{a}\right)^2 = y & \text{Parabel} \\
n=3: & \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1 & \text{elliptischer Zylinder} \\
 & \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 - \left(\frac{z}{c}\right)^2 = 0 & \text{elliptischer Kegel} \\
 & \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 + \left(\frac{z}{c}\right)^2 = 1 & \text{Ellipsoid} \\
 & \left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2 + \left(\frac{z}{c}\right)^2 = 1 & \text{einschaliges Hyperboloid} \\
 & \left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2 - \left(\frac{z}{c}\right)^2 = 1 & \text{zweischaliges Hyperboloid} \\
 & \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = z & \text{Paraboloid} \\
 & \left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2 = z & \text{Sattelfläche}
\end{array}
$$

Abbildung 7.1: Quadriken im $\mathbb{R}^n$ für $n = 1, 2, 3$

Der leere Unterraum $\mathrm{P}(0)$ hat Dimension -1. Die Elemente (oder Punkte) von $\mathrm{P}(V)$ sind Unterräume der Dimension 0. Geraden sind Unterräume der Dimension 1, Ebenen haben die Dimension 2.

Aus 2.5.9 folgt

Bemerkung. *Für Unterräume Q, R eines endlichdimensionalen projektiven Raumes P gilt:*

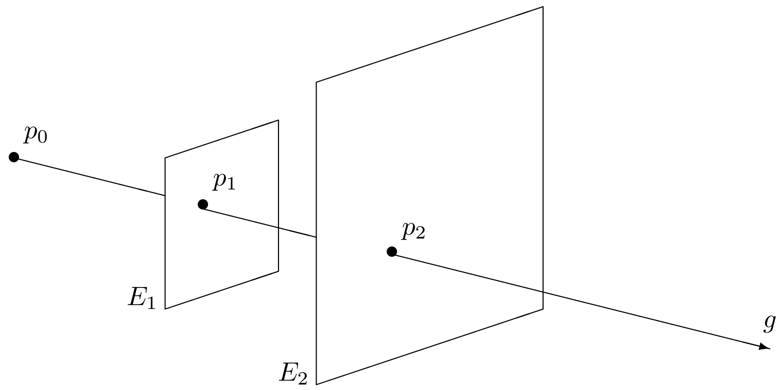
$$\dim(Q \cap R) \geq \dim(Q) + \dim(R) - \dim(P).$$

Insbesondere schneiden sich alle Geraden in einer projektiven Ebene. Und im dreidimensionalen projektiven Raum schneiden sich Ebenen und Geraden.

Lineare Abbildungen zwischen projektiven Räumen $\mathrm{P}(f) : \mathrm{P}(U) \to \mathrm{P}(V)$ werden induziert von injektiven linearen Abbildungen $f : \mathrm{U} \to V$ als

$$\mathrm{P}(f)\big(\langle u \rangle\big) = \langle f(u) \rangle.$$

Seien E_1 und E_2 zwei Ebenen in einem dreidimensionalen projektiven Raum P und p_0 ein Punkt, der zu keiner der beiden Ebenen gehört. Wir können von p_0 aus E_1 auf E_2 in folgender Weise *projizieren*: Wir verbinden einen beliebigen Punkt p_1 von E_1 durch eine Gerade g mit p_0. Dann schneidet g die Ebene E_2 in einem Punkt p_2. Man rechnet leicht nach, daß diese Projektion $p_1 \longmapsto p_2$ linear ist.



Betrachten wir zum Beispiel $\mathrm{P}^n(\mathbb{R}) = \mathrm{P}(\mathbb{R}^{n+1})$. Wir beschreiben die Punkte $\langle(\xi_0, \ldots, \xi_n)\rangle$ durch ihre *homogenen* Koordinaten

$$(\xi_0 : \ldots : \xi_n).$$

Die ξ_i dürfen nicht simultan verschwinden. Es ist $(\xi_0 : \ldots : \xi_n) = (\zeta_0 : \ldots : \zeta_n)$ genau dann wenn $\xi_i = \lambda\zeta_i \;\; (i = 0, \ldots, n)$ für ein $\lambda \neq 0$. Wenn man den $n-1$–dimensionalen Unterraum

$$H_0 = \{(0 : \xi_1 : \ldots : \xi_n) \mid (\xi_1 : \ldots : \xi_n) \in \mathrm{P}^{n-1}(\mathbb{R})\}$$

der *unendlich fernen* Punkte entfernt, erhält man die Menge

$$U_0 = \mathrm{P}^n(\mathbb{R}) \setminus H_0 = \{(1 : \xi_1 : \ldots : \xi_n) \mid (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n\},$$

die wir mit $\mathbb{R}^n$ identifizieren wollen. Man kann sich vorstellen, daß auf diese Weise $\mathrm{P}^n(\mathbb{R})$ aus $\mathbb{R}^n$ entsteht, indem man den Raum der *unendlich fernen* Punkte $\mathrm{P}^{n-1}(\mathbb{R})$ hinzufügt.

Die Unterräume von $\mathrm{P}^n(\mathbb{R})$, die U_0 schneiden, schneiden U_0 in *affinen Unterräumen* von $\mathbb{R}^n$. Das sind Teilmengen der Form $a + U$ $(U \leq \mathbb{R}^n)$. Jeder affine Unterraum vom U_0 entsteht auf diese Weise.

Die affine Gerade

$$g = (\alpha_1, \ldots, \alpha_n) + \mathbb{R} \cdot (\xi_1, \ldots, \xi_n)$$

zum Beispiel ist der Schnitt von U_0 mit der projektiven Geraden

$$\overline{g} = \mathrm{P}\big(\mathbb{R} \cdot (1, \alpha_1, \ldots, \alpha_n) + \mathbb{R} \cdot (0, \xi_1, \ldots, \xi_n)\big),$$

die aus g durch Hinzufügen des unendlich fernen Punktes $(0 : \xi_1 : \ldots : \xi_n)$ entsteht.

In ähnlicher Weise lassen sich affine Quadriken

$$Q = \Big\{(\xi_1, \ldots, \xi_n) \;\Big|\; \sum_{i,j=1}^{n} \alpha_{ij}\xi_i\xi_j + \sum_{i=1}^{n} \alpha_i\xi_i + \alpha = 0\Big\}$$

durch Hinzufügen unendlich ferner Punkte fortsetzen zu *projektiven* Quadriken

$$\overline{Q} = \Big\{(\xi_0 : \xi_1 : \ldots : \xi_n) \;\Big|\; \sum_{i,j=1}^{n} \alpha_{ij}\xi_i\xi_j + \sum_{i=1}^{n} \alpha_i\xi_0\xi_i + \alpha\xi_0^2 = 0\Big\}.$$

Allgemein werden Projektive Quadriken definiert durch eine homogene quadratische Gleichung:

$$\Big\{(\xi_0 : \xi_1 : \ldots : \xi_n) \;\Big|\; \sum_{i,j=0}^{n} \alpha_{ij}\xi_i\xi_j = 0\Big\}.$$

Identifiziert man Quadriken, die sich durch Automorphismen von $\mathrm{P}^n(\mathbb{R})$ ineinander überführen lassen, ergibt sich nach 7.2.5 die folgende Normalform:

$$\xi_1^2 + \cdots + \xi_p^2 - \zeta_1^2 - \cdots - \zeta_q^2 = 0. \tag{7.11}$$

7.5 Unitäre Räume

Wir betrachten in diesem Abschnitt Vektorräume $U, V, \ldots$ über dem Körper $\mathbb{C}$ der komplexen Zahlen.

$n = 1:$	$x_0^2 = 0$	Ein Punkt $((0:1))$
	$x_0^2 - x_1^2 = 0$	Zwei Punkte $((1:1)$ und $(1:-1))$
$n = 2:$	$x_0^2 - x_1^2 = 0$	Zwei Geraden
	$x_0^2 - x_1^2 - x_2^2 = 0$	Kreis
$n = 3:$	$x_0^2 - x_1^2 - x_2^2 = 0$	Zylinder
	$x_0^2 - x_1^2 - x_2^2 - x_3^2 = 0$	Kugel
	$x_0^2 + x_1^2 - x_2^2 - x_3^2 = 0$	Hyperboloid

Abbildung 7.2: Quadriken im $\mathrm{P}^n(\mathbb{R})$ für $n = 1, 2, 3$

Wir erinnern uns an die Definition der Konjugationsabbildung

$$\overline{a + bi} = a - bi$$

Die Konjugation ist ein Automorphismus der $\mathbb{R}$–Algebra $\mathbb{C}$.

Definition. *Eine Abbildung $f : U \to V$ heißt antilinear, wenn*

1. $f(u + v) = f(u) + f(v)$
2. $f(\lambda u) = \overline{\lambda} f(u)$

Eine sesquilineare[10] *Form ist eine Abbildung $(\ ,\ ) : U \times U \to \mathbb{C}$, die antilinear im ersten und linear im zweiten Argument ist.*

Sei $b_1 \dots b_n$ eine Basis von U. Eine Sesquilinearform $(\ ,\ )$ ist bestimmt durch ihre Matrix $B = \big((b_i, b_j)\big)$. Denn wenn x und y die Koeffizientenspalten von $u = \sum_i \xi_i b_i$ und $v = \sum_j \zeta_j b_j$ sind, ist

$$(u, v) = \sum_{i,j} \overline{\xi_i} \zeta_j (b_i, b_j) = x^* B y.$$

Hier steht x^* für den Zeilenvektor $\overline{x}^\top$. Diese Schreibweise verwenden wir für beliebige Matrizen:

$$A^* = \overline{A}.^\top$$

A^* heißt die zu A *adjungierte* Matrix.

Mit Hilfe des *konjugierten* Vektorraums lassen sich die beiden neuen Begriffe auf Bekanntes zurückführen. Sei V ein Vektorraum über $\mathbb{C}$. Der *konjugierte* Vektorraum $V^{(c)}$ hat die gleichen Elemente wie V und die gleiche Addition. Die Multiplikation $\cdot^{(c)}$ mit Elementen von $\mathbb{C}$ ist aber definiert durch

$$\alpha \cdot^{(c)} v = \overline{\alpha}\, v$$

[10] Die Hörer der Vorlesung haben mir erklärt, daß sesquilinear *eineinhalblinear* bedeutet.

Man überlegt sofort, daß $V^{(c)}$ die gleichen Basen, Unterräume und Endomorphismen wie V hat.

Lemma 7.5.1.

1. *$f : U \to V$ ist genau dann antilinear, wenn $f : U^{(c)} \to V$ linear ist.*
2. *$(\ ,\ ) : U \times U \to \mathbb{C}$ ist genau dann sesquilinear, wenn $(\ ,\ ) : U^{(c)} \times U \to \mathbb{C}$ bilinear ist.*

Beweis. $f : U^{(c)} \to V$ ist genau dann linear, wenn für alle $u \in U$ und $\alpha \in \mathbb{C}$

$$f(\overline{\alpha}\, u) = f(\alpha \cdot^{(c)} u) = \alpha f(u).$$

Für $\alpha = \overline{\lambda}$ bedeutet das $f(\lambda u) = \overline{\lambda} f(u)$. □

Wenn $(\ ,\ )$ eine Sesquilinearform ist, ist auch

$$[u, v] = \overline{(v, u)}$$

sesquilinear. Zum Beispiel ist

$$[\lambda u, v] = \overline{(v, \lambda u)} = \overline{\lambda (v, u)} = \overline{\lambda}[u, v].$$

Definition. *Eine Sesquilinearform heißt hermitesch, wenn*

$$(u, v) = \overline{(v, u)}$$

für alle u, v.

Eine bezüglich einer Basis $b_1, \dots, b_n$ durch die Matrix B gegebene Sesquilinearform ist genau dann hermitesch, wenn B *hermitesch* oder *selbstadjungiert* ist. Das heißt, wenn

$$B^* = B.$$

Beweis: Die Matrizen (b_i, b_j) und $[b_i, b_j] = \overline{(b_j, b_i)}$ sind adjungiert.

Wenn $(\ ,\ )$ hermitesch ist, ist $\overline{(v, v)} = (v, v)$ für alle v. Das bedeutet $(v, v) \in \mathbb{R}$. Darum ist die nächste Definition sinnvoll:

Definition. *Eine hermitesche Sesquilinearform*[11] *heißt positiv definit, wenn $(v, v) > 0$ für alle von Null verschiedenen v. Ein unitärer Vektorraum ist ein endlichdimensionaler $\mathbb{C}$–Vektorraum zusammen mit einer positiv definiten hermiteschen Sesquilinearform.*

[11] Ich danke F.Windisch für den Hinweis auf einen früheren Fehler in der Definition.

$\mathbb{C}^n$ mit der Standardsesquilinearform x^*y ist unitär.

Wir setzen wieder $\|a\| = \sqrt{(a,a)}$.

Es ist $\|\lambda a\| = |\lambda|\|a\|$. Die Dreiecksungleichung $\|a+b\| \leq \|a\| + \|b\|$ folgt wie in 1.4.3 aus dem nächsten Lemma.

Lemma 7.5.2 (Cauchy-Schwarzsche Ungleichung)**.** *In unitären Räumen ist*

$$|(a,b)| \leq \|a\|\|b\|$$

Gleichheit gilt genau dann, wenn a und b linear abhängig sind.

Beweis. Wenn $a = 0$, sind beide Seiten der Ungleichung Null. Wenn nicht, können wir annehmen, daß $\|a\| = 1$. (Sonst dividieren wir durch $\|a\|$.) Jetzt ist für $\alpha = (a,b)$

$$0 \leq (b - \alpha a, b - \alpha a) = (b,b) - \alpha(b,a) - \overline{\alpha}(a,b) + |\alpha|^2(a,a) = \|b\|^2 - |\alpha|^2.$$

Daraus folgt $|\alpha| \leq \|b\|$. Wenn Gleichheit eintritt, ist $b - \alpha a = 0$. □

Lemma 7.5.3. *V sei ein unitärer Raum mit Sesquilinearform $(\ ,\ )$.*

1. *$(\ ,\ )$ macht $(V^{(c)}, V)$ zu einem dualen Paar.*
2. *Sei U ein Unterraum von V. Dann ist V direkte Summe von U und $U^\perp$.*
3. *V hat eine Orthonormalbasis. (Also ist V zum unitären Raum $\mathbb{C}^n$ isomorph.)*
4. *Sei $b_1, \ldots, b_n$ eine Orthonormalbasis. Dann ist für alle $a \in V$*

$$a = (b_1, a)b_1 + \ldots + (b_n, a)b_n.$$

Beweis. 1): folgt sofort aus 6.3.3, weil $(v,v) = 0 \Rightarrow v = 0$

2): Wir wiederholen den Beweis von 7.3.3. Sei Sei v_0 ein beliebiger Vektor aus V. Weil $(U^{(c)}, U)$ ein duales Paar bilden ist, gibt es nach 6.3.2 ein eindeutig bestimmtes $u_0 \in U$ mit

$$(v_0, u) = (u_0, u)$$

für alle $u \in U$. Es folgt $(v_0 - u_0, u) = 0$ für alle $u \in U$. Das heißt $v_0 - u_0 \in U^\perp$.

3): Durch Induktion über die Dimension von V. Wenn $V = \mathbf{0}$, ist nichts zu zeigen. Sonst gibt es einen Vektor b_1 der Länge 1. V ist direkte Summe von $U = \mathbb{C}b_1$ und $U^\perp$. Die Induktionsvoraussetzung liefert uns eine Orthonormalbasis $b_2, \ldots, b_n$ von $U^\perp$. $b_1, b_2, \ldots, b_n$ ist die gesuchte Orthonormalbasis von V.

4): Sei $a = \alpha_1 b_1 + \cdots + \alpha_n b_n$. Linksmultiplikation mit b_j ergibt

$$(b_j, a) = \alpha_1(b_j, b_1) + \cdots + \alpha_n(b_j, b_n) = \alpha_j(b_j, b_j) = \alpha_j.$$

□

Satz 7.5.4. *Sei V ein unitärer Vektorraum. Zu jedem Endomorphismus f von V gibt es einen durch $(f^t(u), v) = (u, f(v))$ eindeutig bestimmten* adjungierten *Endomorphismus f^t. Wenn f bezüglich einer Orthonormalbasis durch die Matrix A dargestellt wird, wird f^t durch die adjungierte Matrix A^* dargestellt.*

Beweis. Die eindeutige Existenz von f^t folgt sofort aus 6.3.4. Die zweite Behauptung kann man wie 7.3.4 beweisen. Eine Alternative ist folgende Überlegung: Sei $b_1, \ldots, b_n$ eine Orthonormalbasis und f durch $A = (\alpha_{ij})$ dargestellt. Dann folgt aus 7.5.3.(4), daß

$$\alpha_{ij} = (b_i, f(b_j)).$$

Die Matrix von f^t berechnet sich dann als

$$(b_i, f^t(b_j)) = (f(b_i), b_j) = \overline{(b_j, f(b_i))} = \overline{\alpha_{ji}}.$$

□

Sei V unitär. Der Beweis von 7.1.2 zeigt, daß sich Sesquilinearformen und Endomorphismen von V vermöge

$$[u, v] = (u, f(v))$$

entsprechen. $[\ ,\ ]$ ist genau dann hermitesch, wenn f selbstadjungiert ist, weil

$$\overline{[v, u]} = \overline{(v, f(u))} = (f(u), v) = (u, f^t(v)).$$

Automorphismen eines unitären Vektorraums heißen *unitär*.

Lemma 7.5.5. *Sei f ein Endomorphismus des unitären Vektorraums V, der bezüglich einer Orthonormalbasis $b_1, \ldots, b_m$ durch die Matrix A dargestellt wird. Dann sind äquivalent:*

1. *f ist unitär.*
2. *$f^t \circ f = \mathrm{id}_V$.*
3. *$A^* A = \mathbf{I}$.*
4. *$f(b_1), \ldots, f(b_m)$ ist eine Orthonormalbasis.*

Beweis. f ist ein Automorphismus, wenn

$$(u, v) = (f(u), f(v)) = (f^t f(u), v).$$

Damit ist die Äquivalenz von 1) und 2) klar. Die Äquivalenz von 2) und 3) folgt aus 7.5.4. 3) besagt, daß die Spalten von A orthonormal sind. 3) und 4) sind also auch äquivalent. □

Folgerung 7.5.6. *Der Absolutbetrag einer Determinante einer unitären Abbildung ist* 1.

Beweis. Aus 7.5.4 und 4.3.5 folgt $\det(f^t) = \overline{\det(f)}$. Also ist

$$|\det(f)|^2 = \det(f^t)\det(f) = \det(f^t \circ f) = \det(\mathrm{id}_V) = 1.$$

□

Die *unitäre Gruppe*

$$\mathrm{U}_n = \{A \in \mathrm{M}_n(\mathbb{C}) \mid A^* A = \mathbf{I}\}$$

ist die Gruppe aller unitären n-n–Matrizen. Die unitären Abbildungen mit Determinante 1 bilden die *spezielle* Gruppe SU_n.

Beispiel

Man rechnet leicht nach, daß

$$\mathrm{SU}_2 = \left\{ \begin{pmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{pmatrix} \;\middle|\; |\alpha|^2 + |\beta|^2 = 1 \right\}.$$

Setzt man $\alpha = x_1 + x_2\mathrm{i}$ und $\beta = x_3 + x_4\mathrm{i}$, ergibt sich eine (in beiden Richtungen stetige) Bijektion von SU_2 mit der 3–Sphäre

$$\mathrm{S}^3 = \left\{ (x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \;\middle|\; x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1 \right\}.$$

Definition. *Ein Endomorphismus eines unitären Vektorraums heißt normal, wenn* $f \circ f^t = f^t \circ f$.

Selbstadjungierte und unitäre Endomorphismen sind normal.

Satz 7.5.7. *Ein Endomorphismus f eines unitären Vektorraums V ist genau dann normal, wenn V eine Orthonormalbasis aus Eigenvektoren von f hat.*

Beweis. Wenn f in einer Orthonormalbasis durch die Diagonalmatrix A dargestellt wird, wird f^t durch die Matrix $A^* = \overline{A}$ dargestellt (7.5.4), die natürlich mit A kommutiert.

Sei umgekehrt f normal. Wir zeigen die Behauptung durch Induktion über die Dimension von V. Wenn $\dim(V) = 0$, ist nichts zu zeigen. Sonst hat f, weil $\mathbb{C}$ algebraisch abgeschlossen ist, einen Eigenwert λ. Wenn ψ mit ϕ kommutiert, ist der Eigenraum $U = \{v \mid \phi(v) = \lambda x\}$ invariant unter ψ, weil

$$\phi(v) = \lambda v \Rightarrow \phi\psi(v) = \psi\phi(v) = \psi(\lambda v) = \lambda\psi(v).$$

Insbesondere ist $U = \{v \mid f(v) = \lambda x\}$ invariant unter f^t. Lemma B (Seite 105), das für duale Paare bewiesen wurde, impliziert, daß $U^\perp$ invariant unter $f^{tt} = f$ ist. Wir können jetzt die Induktionsvoraussetzung auf die Einschränkung von f auf $U^\perp$ anwenden und erhalten eine Orthonormalbasis von $U^\perp$, die aus Eigenvektoren von f besteht. Wenn wir diese Basis mit irgendeiner Orthonormalbasis von U zusammensetzen, erhalten wir das Gewünschte. □

Ein normaler Endomorphismus ist selbstadjungiert, wenn seine Eigenwerte reell sind. Und unitär, wenn die Eigenwerte den Betrag 1 haben.

Beispiel

Die Drehmatrix

$$D = \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix}$$

ist normal. Ihr charakteristisches Polynom

$$\mathrm{P}_D = (x - \cos(\phi))^2 + \sin^2(\phi) = x^2 - 2\cos(\phi)x + 1$$

hat die Nullstellen

$$\lambda_1 = \cos(\phi) + \mathrm{i}\sin(\phi) \quad \text{und} \quad \lambda_2 = \cos(\phi) - \mathrm{i}\sin(\phi)$$

Die zugehörigen Eigenvektoren

$$v_1 = \begin{pmatrix} \mathbf{1} \\ -\mathrm{i} \end{pmatrix} \quad \text{und} \quad v_2 = \begin{pmatrix} \mathbf{1} \\ \mathrm{i} \end{pmatrix}$$

sind orthogonal. Normiert bilden sie die Spalten einer unitären Matrix

$$B = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-\mathrm{i}}{\sqrt{2}} & \frac{\mathrm{i}}{\sqrt{2}} \end{pmatrix}.$$

und wir haben

$$B^{-1}DB = B^*DB = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Kapitel 8

Multilineare Algebra

In diesem Kapitel betrachten wir Vektorräume über einem beliebigen Körper K.

8.1 Tensorprodukt

U und V seien zwei Vektorräume. Ein Vektorraum T zusammen mit einer bilinearen Abbildung $\otimes : U \times V \to T$ heißt *Tensorprodukt* von U und V wenn man die folgende *universelle Eigenschaft* hat: Jede bilineare Abbildung $(\ ,\ ) : U \times V \to W$ läßt sich als Komposition von $\otimes$ mit einer eindeutig bestimmten linearen Abbildung $f : T \to W$ schreiben.

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

Satz 8.1.1. *Je zwei Vektorräume U und V haben ein Tensorprodukt*

$$\otimes : U \times V \to T.$$

T ist eindeutig bestimmt in folgendem Sinn: Wenn $\otimes' : U \times V \to T'$ ein zweites Tensorprodukt ist, gibt es genau einen Isomorphismus $t : T \to T'$, der das Diagramm

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

kommutativ macht.

Beweis. Wir wählen Basen $(a_i)_{i\in I}$ und $(b_j)_{j\in J}$ von U und V. Für T nehmen wir einen Vektorraum mit einer nach $I \times J$ indizierten Basis $(c_{i,j})$. Schließlich definieren wir die bilineare Abbildung $\otimes : U \times V \to T$ durch $a_i \otimes b_j = c_{i,j}$ (wie in 2.6.1). Sei die Abbildung $\otimes' : U \times V \to T'$ bilinear. Wenn $t : T \to T'$ linear ist, ist $t \circ \otimes$ bilinear und stimmt mit $\otimes'$ genau dann überein, wenn für alle i und j

$$t(a_i \otimes b_j) = t(c_{i,j}) = a_i \otimes' b_j.$$

Man kann aber Werte von t auf der Basis $(c_{i,j})$ beliebig vorschreiben und t ist dadurch eindeutig bestimmt. □

Wir wählen für alle Paare U und V ein Tensorprodukt aus und bezeichnen es mit $U \bigotimes V$.

Aus dem Beweis des Satzes ergibt sich sofort Teil 1) der Folgerung:

Folgerung 8.1.2. *Sei $(a_i)_{i\in I}$ eine Basis von U.*

1. *Wenn $(b_j)_{j\in J}$ eine Basis von V ist, ist $(a_i \otimes b_j)_{i\in I,\, j\in J}$ eine Basis von $U \bigotimes V$.*
2. *Jedes Element von $U \bigotimes V$ läßt sich schreiben als*

$$\sum_i a_i \otimes v_i$$

für eindeutig bestimmte $v_j \in V$.

Beweis. Man hat $\sum_{i,j} \alpha_{i,j}(a_i \otimes b_j) = \sum_i a_i \otimes v_i$ für

$$v_i = \sum_j \alpha_{i,j} b_j.$$

Das beweist Teil 2). □

Aus 1) folgt, daß für endlichdimensionale Vektorräume

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

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

Lemma 8.1.3. *Die kanonische Abbildung $K \bigotimes V \to V$, definiert durch*

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

ist ein Isomorphismus.

Beweis. Die Umkehrabbildung ist $v \mapsto 1 \otimes v$, weil

$$\alpha \otimes v = \alpha(1 \otimes v) = 1 \otimes (\alpha v).$$

□

$f : U \to U'$ und $g : V \to V'$ seien lineare Abbildungen. Weil

$$(u, v) \mapsto f(u) \otimes g(v)$$

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

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

mit $(f \otimes g)(u \otimes v) = f(u) \otimes g(v)$.

$\bigotimes$ wird auf diese Weise zu einem linearen Bifunktor:

Lemma 8.1.4.

1. $\mathrm{id}_U \otimes \mathrm{id}_V = \mathrm{id}_{U \bigotimes V}$
2. $(f \circ f') \otimes (g \circ g') = (f \otimes g) \circ (f' \otimes g')$
3. $(f + f') \otimes g = (f \otimes g) + (f' \otimes g)$
4. $(\alpha f) \otimes g = \alpha(f \otimes g)$

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

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

1): Klar

2):

$$\begin{aligned}(ff' \otimes gg')(u \otimes v) &= ff'u \otimes gg'v \\ &= (f \otimes g)(f'u \otimes g'v) \\ &= (f \otimes g)(f' \otimes g')(u \otimes v)\end{aligned}$$

3):

$$\begin{aligned}((f + f') \otimes g)(u \otimes v) &= (f + f')u \otimes gv \\ &= (fu + f'u) \otimes gv \\ &= fu \otimes gv + f'u \otimes gv \\ &= (f \otimes g)(u \otimes v) + (f' \otimes g)(u \otimes v) \\ &= (f \otimes g + f' \otimes g)(u \otimes v)\end{aligned}$$

4):

$$\begin{aligned}(\alpha f \otimes g)(u \otimes v) &= (\alpha fu) \otimes gv \\ &= \alpha(fu \otimes gv) \\ &= \alpha(f \otimes g)(u \otimes v)\end{aligned}$$

□

Beispiel

Sei V ein $\mathbb{R}$–Vektorraum. Dann ist $\mathbb{C} \bigotimes V$ ein $\mathbb{C}$–Vektorraum mit der Skalarmultiplikation

$$\lambda(\alpha \otimes v) = (\lambda\alpha) \otimes v. \tag{8.1}$$

Man nennt $\mathbb{C} \bigotimes V$ die Komplexifizierung von V. Daß (8.1) eine $\mathbb{C}$–Vektorraumstruktur definiert, kann man schnell so einsehen: Sei $\mu_\lambda : \mathbb{C} \to \mathbb{C}$ die Multiplikation mit λ. Dann ist $\lambda(\alpha \otimes v) = (\mu_\lambda \otimes \mathrm{id}_V)(\alpha \otimes v)$. Jetzt kann man das letzte Lemma anwenden.

Definition. *Das Tensorprodukt einer endlichen Familie $V_1, \ldots, V_n$ von Vektorräumen ist ein Vektorraum*

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

zusammen mit einer multilinearen Abbildung

$$\otimes : V_1 \times \cdots \times V_n \to V_1 \bigotimes \cdots \bigotimes V_n,$$

so daß sich jede multilineare Abbildung $\mu : V_1 \times \cdots \times V_n \to W$ eindeutig in der Form $\mu = f \circ \otimes$ für ein lineares $f : V_1 \bigotimes \cdots \bigotimes V_n \to W$ schreiben läßt.

Wir schreiben immer $v_1 \otimes \cdots \otimes v_n$ statt $\otimes(v_1, \ldots, v_n)$. f ist dann gegeben durch

$$f(v_1 \otimes \cdots \otimes v_n) = \mu(v_1, \ldots, v_n).$$

Man sagt, daß μ *durch* $\otimes$ *faktorisiert.*

Existenz und Eindeutigkeit zeigt man wie in 8.1.1. Wenn $(b_i^1)_{i \in I_1}, \ldots, (b_i^n)_{i \in I_n}$ Basen der $V_1, \ldots, V_n$ sind, ist

$$\left(b_{i_1}^1 \otimes \cdots \otimes b_{i_n}^n\right)_{i_1 \in I_1, \ldots, i_n \in I_n}$$

eine Basis von $V_1 \bigotimes \cdots \bigotimes V_n$.

Für das Tensorprodukt eines einzigen Vektorraums V (der Fall $n = 1$) kann man offensichtlich V selbst nehmen (mit der identischen Abbildung $V \to V$).

Satz 8.1.5.

1. *Durch* $(u_1 \otimes \cdots u_m) \otimes (v_1 \otimes \cdots v_n) \mapsto u_1 \otimes \cdots \otimes u_m \otimes v_1 \otimes \cdots v_n$ *wird ein Isomorphismus*

$$(U_1 \bigotimes \cdots U_m) \bigotimes (V_1 \bigotimes \cdots V_n) \xrightarrow{\simeq} U_1 \bigotimes \cdots U_m \bigotimes V_1 \bigotimes \cdots V_n$$

   definiert.

2. $U \bigotimes V$ *und* $V \bigotimes U$ *sind kanonisch isomorph. Der Isomorphismus wird durch* $u \otimes v \mapsto v \otimes u$ *gegeben.*

Beweis.

1): Wir verwenden die Abkürzungen $U = U_1 \bigotimes \cdots U_m$, $V = V_1 \bigotimes \cdots V_n$ und $W = U_1 \bigotimes \cdots U_m \bigotimes V_1 \bigotimes \cdots V_n$.

Für jedes festgehaltene m–Tupel $\overline{u} = (u_1, \ldots, u_m)$ ist die Abbildung

$$(v_1, \ldots, v_n) \mapsto u_1 \otimes \cdots \otimes u_m \otimes v_1 \otimes \cdots v_n$$

multilinear und läßt sich daher schreiben als

$$u_1 \otimes \cdots \otimes u_m \otimes v_1 \otimes \cdots v_n = f_{\overline{u}}(v_1 \otimes \cdots v_n)$$

für ein lineares $f_{\overline{u}} : V \to W$.

Die Abbildung $(u_1, \ldots, u_m) \mapsto f_{\overline{u}}$ ist multilinear und daher von der Form

$$f_{\overline{u}} = f(u_1 \otimes \cdots u_m)$$

für eine lineare Abbildung $f : U \to \mathrm{L}(V, W)$.

Weil $(u, v) \mapsto f(u)(v)$ bilinear ist, gibt es ein lineares $g : U \bigotimes V \to W$ mit $g(u \otimes v) = f(u)(v)$. Es gilt offenbar

$$g((u_1 \otimes \cdots u_m) \otimes (v_1 \otimes \cdots v_n)) = u_1 \otimes \cdots \otimes u_m \otimes v_1 \otimes \cdots v_n.$$

Die in der Behauptung angegebene Abbildung ist also tatsächlich wohldefiniert.

Um zu zeigen, daß g ein Isomorphismus ist, geben wir eine Umkehrabbildung an.

$$(u_1, \ldots, u_m, v_1, \ldots, v_n) \mapsto (u_1 \otimes \cdots u_m) \otimes (v_1 \otimes \cdots v_n)$$

ist multilinear. Also gibt es lineares $h : W \to U \bigotimes V$ mit

$$h(u_1 \otimes \cdots u_m \otimes v_1 \otimes \cdots v_n) = (u_1 \otimes \cdots u_m) \otimes (v_1 \otimes \cdots v_n).$$

h ist offensichtlich die Umkehrabbildung von g.

2): Weil $(u, v) \mapsto v \otimes u$ multilinear ist, gibt es ein lineares $g : U \bigotimes V \to V \bigotimes U$ mit $g(u \otimes v) = v \otimes u$. g ist ein Isomorphismus: Man kann sofort die Umkehrabbildung $h(v \otimes u) = u \otimes v$ angeben, □

Sei $\bigotimes^0 = K$ und (für positive p) $\bigotimes^p V$ das Tensorprodukt von p Kopien von V. Dann ist nach dem vorhergehenden Satz für alle p und q eine bilineare Abbildung $\otimes : \bigotimes^p V \times \bigotimes^q V \to \bigotimes^{p+q} V$ definiert. Wenn p oder q Null sind, soll diese Abbildung Multiplikation mit Elementen von K sein.

Satz 8.1.6. *Die Abbildungen $\otimes : \bigotimes^p V \times \bigotimes^q V \to \bigotimes^{p+q} V$ machen*

$$\mathrm{T}(V) = \bigoplus_{p \in \mathbb{N}} \left(\bigotimes^p V \right)$$

zu einer assoziativen Algebra mit Einselement, der Tensoralgebra von V.

Beweis. Das Einselement von $\bigotimes^0 V = K$ ist natürlich Einselement. Es bleibt zu zeigen, daß die auf $\mathrm{T}(V)$ definierte bilineare Operation $\otimes$ assoziativ ist. Es genügt, das für Elemente der direkten Summanden $\bigotimes^p V$ zu zeigen. Sei also $v = v_1 \otimes \cdots$ aus $\bigotimes^p V$, $u = u_1 \otimes \cdots$ aus $\bigotimes^q V$ und $w = w_1 \otimes \cdots$ aus $\bigotimes^r V$. Dann ist

$$\begin{aligned}(u \otimes v) \otimes w &= (u_1 \otimes \cdots v_1 \otimes \cdots) \otimes w \\ &= u_1 \otimes \cdots v_1 \otimes \cdots w_1 \otimes \cdots \\ &= u \otimes (v_1 \otimes \cdots w_1 \otimes \cdots) = u \otimes (v \otimes w).\end{aligned}$$

□

<u>Beispiel</u>

Wir wollen $\mathrm{T}(K)$ berechnen. Alle Tensorpotenzen von K sind kanonisch isomorph zu K. Wir bezeichnen das Einselement von $\bigotimes^1 K$ mit x. Dann ist für jedes p

$$x^p = \underbrace{x \otimes \cdots \otimes x}_{p-\mathrm{mal}}$$

Basis von $\bigotimes^p V$. Die Elemente $1, x, x^2, \ldots$ bilden also eine Basis von $\mathrm{T}(K)$ und $\mathrm{T}(V)$ ist kanonisch isomorph zum Polynomring $K[X]$.

8.2 Tensorprodukt und Dualität

Definition. *Die durch die bilineare Abbildung* $(\lambda, x) \mapsto \lambda(x)$ *gegebene lineare Abbildung*

$$\mathrm{v} : V^* \bigotimes V \to K$$

heißt Verjüngung.

Allgemeiner definiert man die Verjüngung über V_i und V_j

$$\mathrm{v} : \bigotimes_{k=1}^{p} V_k \to \bigotimes_{k \neq i,j}^{p} V_k,$$

wenn für ein Indexpaar $V_i = V_j^*$, durch[1]

$$\mathrm{v}(v_1 \otimes \cdots v_i \cdots v_j \cdots \otimes v_p) = v_i(v_j) \cdot v_1 \otimes \cdots \widehat{v_i} \cdots \widehat{v_j} \cdots \otimes v_p.$$

Die Elemente von

$$V_p^q = (\bigotimes^{p} V) \bigotimes (\bigotimes^{q} V^*)$$

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

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

Verjüngen über der i–ten Kopie von V und der j–ten Kopie von V^* liefert eine Abbildung

$$\Gamma_i^j : V_p^q \to V_{p-1}^{q-1}.$$

Die Verjüngung

$$\mathrm{v} : U \bigotimes V^* \bigotimes V \to U$$

liefert für jedes $a \in U \bigotimes V^*$ eine lineare Abbildung $f_a : V \to U$, definiert durch

$$f_a(v) = \mathrm{v}(a \otimes v).$$

Die Abbildung $a \mapsto f_a$ ist offensichtlich linear und es gilt

$$f_{u \otimes \lambda}(v) = \lambda(v) u.$$

[1]Die Schreibweise bedeutet, daß im rechten Produkt die Faktoren v_i und v_j weggelassen werden.

Satz 8.2.1. *Die Abbildung $a \mapsto f_a$ ist ein Isomorphismus zwischen $U \bigotimes V^*$ und dem Raum $\mathrm{L_e}(V, U)$ der lineare Abbildungen $V \to U$, die endlichen Rang haben.*

Beweis. Wenn $a = u_1 \otimes \lambda_1 + \cdots + u_m \otimes \lambda_m$, liegt das Bild von f_a in dem von den u_i aufgespannten Raum. f_a hat also höchstens den Rang m.

Sei $u_1, \ldots, u_m$ eine Basis von $\mathrm{Im}(f)$. Jedes $f(v)$ läßt sich schreiben als

$$f(v) = \mu_1(v)u_1 + \cdot + \mu_m(v)u_m$$

mit eindeutig bestimmten Linearformen μ_i. Für

$$a = u_1 \otimes \mu_1 + \cdots + u_m \otimes \mu_m$$

ist dann $f = f_a$. Daraus folgt die Behauptung. □

Man kann daher (im endlichdimensionalen) $U \bigotimes V^*$ und $\mathrm{L}(V, U)$ identifizieren. Dabei ergeben sich folgende Entsprechungen:

1. Matrizen:
   Sei (v_j) eine Basis von V, (u_i) eine Basis von U und die lineare Abbildung $f : V \to U$ gegeben durch die Matrix $(\alpha_{i,j})$. Sei (λ_j) die zu (v_j) duale Basis von V^*. Dann ist (bei der obigen Identifizierung)

   $$f = \sum_{i,j} \alpha_{i,j}(u_i \otimes \lambda_j).$$

   Beweis:
   $$\Big(\sum_{i,j} \alpha_{i,j} u_i \otimes \lambda_j\Big)(v_k) = \sum_{i,j} \alpha_{i,j}\lambda_j(v_k)u_i = \sum_i \alpha_{ik}u_i = f(v_k)$$

2. Komposition:
   Die Komposition von linearen Abbildungen ist eine bilineare Abbildung

   $$\circ : \mathrm{L}(V, U) \times \mathrm{L}(W, V) \to \mathrm{L}(W, U).$$

   Sie entspricht der Verjüngung

   $$(U \bigotimes V^*) \bigotimes (V \bigotimes W^*) \to U \bigotimes W^*.$$

   Beweis: Einerseits verjüngt sich $(u \otimes \lambda) \otimes (v \otimes \mu)$ zu $\lambda(v)(u \otimes \mu)$. Andererseits berechnet sich die Komposition der Abbildungen $u \otimes \lambda$ und $v \otimes \mu$ als

   $$(u \otimes \lambda)(v \otimes \mu)(w) = (u \otimes \lambda)(\mu(w)v) = \lambda(v)\mu(w)u = \lambda(v)(u \otimes \mu)(w).$$

3. Duale Abbildung:
   Die Dualisierung $\mathrm{L}(V, U) \to \mathrm{L}(U^*, V^*)$, definiert durch $f \mapsto f^*$ entspricht der Abbildung

   $$U \bigotimes V^* \to V^* \bigotimes U = V^* \bigotimes U^{**}$$

aus Satz 8.1.5(2).

Beweis: Die zu $(u \otimes \lambda)$ duale Abbildung f^* ist charakterisiert durch

$$f^*(\chi)(v) = \chi\big((u \otimes \lambda)(v)\big).$$

Das rechnen wir für $\lambda \otimes u^{**}$ nach. Dabei ist u^{**} das Bild von u unter dem kanonischen Isomorphismus $U \xrightarrow{\simeq} U^{**}$.

$$(\lambda \otimes u^{**})(\chi)(v) = \big(u^{**}(\chi)\lambda\big)(v) = \chi(u)\lambda(v) = \chi(\lambda(v)u) = \chi\big((u \otimes \lambda)(v)\big)$$

4. Die Spurabbildung:
   $\mathrm{Tr} : \mathrm{End}(V) \to K$ ist nichts anderes als die Verjüngung
   $$V \bigotimes V^* \to K.$$
   Beweis: Wenn man
   $$f = \sum_{i,j} \alpha_{i,j}(u_i \otimes \lambda_j)$$
   verjüngt, ergibt sich
   $$\mathrm{v}(f) = \sum_{i,j} \alpha_{i,j}\lambda_j(u_i) = \sum_i \alpha_{i,i}\lambda_i(u_i) = \sum_i \alpha_{i,i} = \mathrm{Tr}(f).$$

Zweimalige Verjüngung

$$(\lambda \otimes \mu) \otimes (u \otimes v) \mapsto \lambda(u)\mu(v)$$

liefert eine bilineare Abbildung $(U^* \bigotimes V^*) \times (U \bigotimes V) \to K$.

Satz 8.2.2. *Wenn U und V endlichdimensional sind, bilden $U^* \bigotimes V^*$ und $U \bigotimes V$ mit der eben definierten bilinearen Abbildung ein duales Paar.*

Beweis. Sei (u_i) ein Basis von U mit dualer Basis (λ_i). (v_j) sei eine Basis von V und (μ_j) die duale Basis. Dann ist die Basis $(\lambda_i \otimes \mu_j)$ *dual* zu $(u_i \otimes v_j)$, denn

$$\big(\lambda_k \otimes \mu_l, u_i \otimes v_j\big) = \lambda_k(u_i)\mu_l(v_j) = \begin{cases} 1 & \text{wenn } (i,j) = (k,l) \\ 0 & \text{sonst} \end{cases}$$

Daraus folgt die Behauptung. □

$(U \bigotimes V)^*$ kann also mit $U^* \bigotimes V^*$ identifiziert werden.[2]

8.3 Die äußere Algebra

Definition. *Eine k–stellige multilineare Abbildung $\mu : V^k \to W$ heißt alternierend, wenn*

$$\mu(a_1, \ldots, a_k) = 0$$

für alle k–Tupel $a_1, \ldots, a_k$, in denen ein Vektor zweimal vorkommt.

[2]Dafür genügt es, wenn einer der beiden Vektorräume endlichdimensional ist.

Alle Sätze, die im Abschnitt 4.2 über k–Formen, also für alternierende Abbildungen nach K, bewiesen worden sind, gelten für beliebige alternierende Abbildungen.

Sei V ein K–Vektorraum und p eine natürliche Zahl. Ein Vektorraum A zusammen mit einer (p–stelligen) multilinearen alternierenden Abbildung $\wedge^p : V^p \to A$ heißt *p-fache äußere Potenz* von V, wenn sich jede alternierende multilineare Abbildung $a : V^p \to W$ als Komposition von $\wedge^p$ mit einer eindeutig bestimmten linearen Abbildung $f : A \to W$ schreiben läßt.

$$\begin{tikzcd} V^p \arrow[r, "\wedge^p"] \arrow[dr, "a"'] & A \arrow[d, "f"] \\ & W \end{tikzcd}$$

Wir schreiben

$$v_1 \wedge v_2 \wedge \cdots \wedge v_p$$

für $\wedge^p(v_1, v_2, \cdots, v_p)$.

Satz 8.3.1. *V hat für alle p eine p–fache äußere Potenz $\wedge^p : V^p \to A$. A ist eindeutig bestimmt in folgendem Sinn: Wenn $\wedge' : V^p \to A'$ eine zweite p–fache äußere Potenz von V ist, gibt es genau einen Isomorphismus $a : A \to A'$, der das Diagramm*

$$\begin{tikzcd} V^p \arrow[r, "\wedge^p"] \arrow[dr, "\wedge'"'] & A \arrow[d, "a"] \\ & A' \end{tikzcd}$$

kommutativ macht.

Beweis. Sei $(e_i)_{i \in \mathcal{I}}$ eine Basis von V. Für A wählen wir einen Vektorraum mit einer Basis (e_I), die mit den p–elementigen Teilmengen I von $\mathcal{I}$ indiziert ist.

Um $\wedge^p$ zu definieren, fixieren wir eine lineare Ordnung[3] von $\mathcal{I}$ und definieren

$$e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_p}$$

als 0, wenn $I = \{i_1, \ldots, i_p\}$ weniger als p Elemente hat. Und sonst als

$$\epsilon_{i_1 \cdots i_p} e_I,$$

wobei $\epsilon_{i_1 \cdots i_p}$ die Signatur der Permutation $\sigma \in \operatorname{Sym}(I)$ ist, für die $\sigma(i_1) < \sigma(i_2) < \cdots < \sigma(i_p)$. Wir verifizieren, wie im Beweis von 4.3.1, daß die multilineare Abbildung $\wedge^p$ alterniert. Der Einfachheit halber zeigen wir nur, daß $\wedge^p$ Null ist, wenn die ersten beiden Argumente gleich sind. Dazu genügt es zu zeigen, daß $a \wedge a \wedge e_{i_3} \wedge \cdots \wedge e_{i_p} = 0$ für alle $i_3, \ldots, i_p$. Wenn $a = \sum_j \xi_j e_j$, ist

$$a \wedge a \wedge e_{i_3} \wedge \cdots \wedge e_{i_p} = \sum_{j<k} \xi_j \xi_k \big((e_j \wedge e_k \wedge \cdots) + (e_k \wedge e_j \wedge \cdots) \big)$$

Dabei summieren wir über alle i, k, die in $\{i_3, \ldots, i_p\}$ nicht vorkommen. Weil

$$\epsilon_{jki_3 \cdots i_p} = -\epsilon_{kji_3 \cdots i_p}$$

verschwinden alle Summanden, wie gewünscht. □

Wir wählen für alle V und p eine p–fache äußere Potenz aus und bezeichnen sie mit $\bigwedge^p V$. Für $p = 0$ ist der Sinn der Definition nicht ganz klar; man vereinbart $\bigwedge^0 V = K$. $\bigwedge^1 V$ kann mit V identifiziert werden. Die Elemente von $\bigwedge^p V$ heißen die *p–Vektoren* von V.

Betrachtet man die Definition der äußeren Potenz im Fall $W = K$, sieht man, daß sich der Raum $\mathrm{A}^p V$ aller p–Formen auf V mit dem Dualraum $(\bigwedge^p V)^*$ identifizieren läßt: Jedes $\mu \in \mathrm{A}^p V$ läßt sich schreiben als

$$\mu(v_1, \ldots, v_p) = \lambda(v_1 \wedge \cdots \wedge v_p)$$

für ein $\lambda \in (\bigwedge^p V)^*$.

Sei $e_1, e_2, \ldots e_n$ eine Basis von V. Für alle p– Teilmengen I von $\{1, \ldots, n\}$ sei

$$e_I = e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_p},$$

wobei $i_1 < i_2 < \ldots < i_p$ eine monotone Aufzählung von I ist. Aus dem Beweis des Satzes ergibt sich:

Folgerung 8.3.2. *Die e_I bilden eine Basis von $\bigwedge^p V$.*

Wenn $\dim V = n$, ist also

$$\dim(\bigwedge^p V) = \binom{n}{p}.$$

[3]Zorns Lemma liefert die Existenz einer maximalen partiellen Ordnung auf $\mathcal{I}$. Maximale partielle Ordnungen sind linear.

Insbesondere ist $\dim(\bigwedge^n V) = 1$ und $\dim(\bigwedge^p V) = 0$ für alle $p > n$.

Nehmen wir, an die Vektoren $v_1, \ldots, v_p$ seien vermöge

$$(v_1 \ldots v_p) = (e_1 \ldots e_n)A$$

durch die n–p–Matrix A gegeben. Dann können wir das äußere Produkt der v_i folgendermaßen berechnen. (I läuft über alle p–elementigen Teilmengen von $\{1, \ldots, n\}$.)

$$\begin{aligned}
v_1 \wedge \cdots \wedge v_p &= \sum_I \sum_{\{i_1,\ldots,i_p\}=I} \alpha_{i_1,1} e_{i_1} \wedge \cdots \wedge \alpha_{i_p,p} e_{i_p} \\
&= \sum_I \Big(\sum_{\{i_1,\ldots,i_p\}=I} \epsilon_{i_1 \cdots i_p} \prod_k \alpha_{i_k,k} \Big) e_I \\
&= \sum_I \det(A_I) e_I
\end{aligned}$$

Dabei besteht die p–p–Matrix A_I aus den Zeilen von A, deren Zeilennummern zu I gehören.

Lemma 8.3.3. *Zu jeder linearen Abbildung $\phi : V \to U$ gibt es eine lineare Abbildung $\wedge^p \phi : \bigwedge^p V \to \bigwedge^p U$, die eindeutig bestimmt ist durch*

$$\wedge^p \phi(v_1 \wedge \cdots \wedge v_p) = \phi(v_1) \wedge \cdots \wedge \phi(v_p).$$

$\bigwedge^p$ wird dadurch zu einem Funktor.

Beweis. Existenz und Eindeutigkeit folgen aus der universellen Eigenschaft der äußeren Potenz, weil $\phi(v_1) \wedge \cdots \wedge \phi(v_p)$ ein alternierender Ausdruck in den v_i ist. Daß $\wedge^p \mathrm{id} = \mathrm{id}$ ist klar. Schließlich folgt $\wedge^p(\phi\psi) = \wedge^p(\phi) \wedge^p (\psi)$ aus

$$\begin{aligned}
\wedge^p(\phi\psi)(v_1 \wedge \cdots) &= (\phi\psi v_1) \wedge \cdots \\
&= \wedge^p \phi(\psi v_1 \wedge \cdots) \\
&= \wedge^p \phi \wedge^p \psi(v_1 \wedge \cdots)
\end{aligned}$$

□

Sei $e_1, \ldots, e_m$ eine Basis von U, $f_1, \ldots, f_n$ eine Basis von V und $\phi : V \to U$ durch die m–n–Matrix A gegeben. Wir wollen die Matrixdarstellung von $\wedge^p \phi$ berechnen.

Sei $J = \{j_1 < \ldots < j_p\}$ eine p–elementige Teilmenge von $\{1, \ldots, n\}$ und A^J die Matrix, die aus den Spalten von A besteht, deren Index in J liegt. Dann ist $\big(\phi(f_{j_1}) \cdots \phi(f_{j_p})\big) = (e_1 \cdots e_m)A^J$. Also

$$\wedge^p \phi(f_J) = \phi(f_{j_1}) \wedge \cdots \wedge \phi(f_{j_p}) = \sum_I \det(A_I^J) e_I$$

Daraus folgt sofort

Lemma 8.3.4. *Wenn* $\dim V = n$ *und* ϕ *ein Endomorphismus von* V *ist, ist* $\wedge^n \phi = \det(\phi)$. □

Man beachte, daß ein Endomorphismus eines eindimensionalen Raumes Multiplikation mit einem Körperelement ist.

Satz 8.3.5. *Für alle* p *und* q *gibt es eine bilineare Abbildung*

$$\wedge : \bigwedge^p V \times \bigwedge^q V \to \bigwedge^{p+q} V,$$

gegeben durch $(v_1 \wedge \cdots \wedge v_p) \wedge (v_{p+1} \wedge \cdots \wedge v_{p+q}) = v_1 \wedge \cdots \wedge v_{p+q}$.

Wenn p oder q gleich Null sind, nehmen wir für $\wedge$ die Multiplikation mit Elementen von K.

Beweis. Der Beweis folgt dem ersten Teil des Beweis von 8.1.5(1): Für jedes festgehaltene p–Tupel $\overline{v} = (v_1, \ldots, v_p)$ ist die Abbildung

$$(v_{p+1}, \ldots, v_{p+q}) \mapsto v_1 \wedge \cdots \wedge v_{p+q}$$

alternierend und läßt sich daher schreiben als

$$v_1 \wedge \cdots \wedge v_{p+q} = f_{\overline{v}}(v_{p+1} \wedge \cdots \wedge v_{p+q})$$

für ein lineares $f_{\overline{v}} : \bigwedge^q V \to \bigwedge^{p+q} V$.

Die Abbildung $(v_1, \ldots, v_p) \mapsto f_{\overline{v}}$ ist alternierend und daher von der Form

$$f_{\overline{v}} = f(v_1 \wedge \cdots v_p)$$

für eine lineare Abbildung $f : \bigwedge^p V \to \mathrm{L}(\bigwedge^q V, \bigwedge^{p+q} V)$. $u \wedge v = f(u)(v)$ ist also wohldefiniert und bilinear. □

Sei $e_1, e_2, \ldots e_n$ eine Basis von V. Wenn I eine p–elementige Teilmenge und J eine q–elementige Teilmenge von $\{1, \ldots n\}$ ist, ist

$$e_I \wedge e_J = \begin{cases} 0 & , \text{ wenn } \quad I \cap J \neq 0 \\ \epsilon_{I,J} e_{I \cup J} & , \text{ sonst.} \end{cases}$$

Dabei ist, wenn $i_1 < \ldots < i_p$, $\;j_1 < \ldots < j_q$ und $k_1 < \ldots < k_{p+q}$ die aufsteigenden Aufzählungen von I, J und $I \cup J$ sind, $\epsilon_{I,J}$ die Signatur der Permutation

$$\begin{pmatrix} i_1 \ldots i_p j_1 \ldots j_q \\ k_1 \ldots k_{p+q} \end{pmatrix},$$

die $i_1 \ldots i_p j_1 \ldots j_q$ der Reihe nach auf $k_1 \ldots k_{p+q}$ abbildet.

Folgerung 8.3.6. *Die direkte Summe*

$$\bigwedge V = \bigoplus_{p\in\mathbb{N}} \left(\bigwedge^p V\right)$$

wird durch die Abbildungen $\wedge$ zu einer assoziativen Algebra mit Einselement. Wenn $v \in \bigwedge^p V$ und $w \in \bigwedge^q V$, ist

$$v \wedge w = (-1)^{pq}\, w \wedge v. \tag{8.2}$$

$\bigwedge V$ heißt die *äußere Algebra* von V.

Beweis. Wie in 8.1.6 beweist man die Assoziativität. $v \wedge w = (-1)^{pq} w \wedge v$ folgt aus $\epsilon_{J,I} = (-1)^{pq}\epsilon_{I,J}$, was wiederum aus

$$\operatorname{sign}\begin{pmatrix} 1 & \cdots & p & p+1 & \cdots & p+q \\ p+1 & \cdots & p+q & 1 & \cdots & p \end{pmatrix} = (-1)^{pq}$$

folgt. □

Satz 8.3.7. *Sei $n = \dim(V)$ und $p + q = n$. Wenn wir einen Isomorphismus $\mu : \bigwedge^n V \xrightarrow{\cong} K$ fixieren, werden $\bigwedge^p V$ und $\bigwedge^q V$ vermöge*

$$\wedge : \bigwedge^p V \times \bigwedge^q V \to \bigwedge^n V \xrightarrow{\mu} K$$

zu einem dualen Paar.

Beweis. Sei e_i eine Basis von V und $\mu(e_{\{1,\ldots,n\}}) = \alpha$. Dann sind die e_I eine Basis von $\bigwedge^p V$ und die

$$f_I = \alpha^{-1}\epsilon_{I,\overline{I}}\, e_{\overline{I}}$$

die dazu duale Basis von $\bigwedge^q V$. Die I laufen dabei über alle p–elementigen Teilmengen von $\{1,\ldots,n\}$ und $\overline{I} = \{1,\ldots,n\} \setminus I$ bezeichnet die Komplementmenge. □

Folgerung 8.3.8 (Cramersche Regel)**.** *Sei $n = \dim(V)$ und μ eine n–Form, dann ist für alle $v_1,\ldots,v_n$ und alle $x \in V$*

$$\mu(v_1,\ldots,v_n)x = \sum_{i=1}^{n} \mu(v_1,\ldots,v_{i-1},x,v_{i+1},\ldots,v_n)v_i$$

Beweis. Zuerst überlegt man sich, daß für festes x die beide Seiten der behaupteten Gleichung alternierend in $v_1,\ldots,v_n$ sind. Wenn zum Beispiel $v_1 = v_2$, verschwinden alle Summanden der rechten Seite bis auf $\mu(x,v_2,\ldots)v_1$ und $\mu(v_1,x,\ldots)v_2$, die sich aber gegenseitig aufheben.

Das bedeutet, daß wir annehmen können, daß $v_1, \ldots, v_n$ eine Basis ist. Wir können jetzt dem Beweis von 4.4.1 folgen, oder dem folgenden Gedankengang. Wenn $\mu = 0$, ist alles klar. Sonst definiert μ eine Isomorphismus zwischen $\bigwedge^n V$ und K, den wir ebenfalls mit μ bezeichnen. Sei wieder $\alpha = \mu(v_1, \ldots, v_n)$. Für die duale Basis

$$w_i = \alpha^{-1}\epsilon_{i,\bar{i}}\, v_{\bar{i}}$$

müssen wir zeigen, daß

$$\mu(v_1, \ldots, v_{i-1}, x, v_{i+1}, \ldots, v_n) = \alpha\, \mu(x \wedge w_i).$$

In der Tat

$$\begin{aligned}
\mu(v_1, \ldots, v_{i-1}, x, v_{i+1}, \ldots, v_n) &= \epsilon_{i,\bar{i}}\, \mu(x, v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n) \\
&= \epsilon_{i,\bar{i}}\, \mu(x \wedge v_{\bar{i}}) \\
&= \alpha\, \mu(x \wedge w_i)
\end{aligned}$$

□

Sei ϕ ein Endomorphismus von V und $\mathrm{adj}(\phi) : V \to V$ der zu $\wedge^{n-1}\phi : \bigwedge^{n-1} V \to \bigwedge^{n-1} V$ adjungierte Endomorphismus. $\mathrm{adj}(\phi)$ wird also bestimmt durch die Gültigkeit von

$$\mathrm{adj}(\phi)(v_1) \wedge v_2 \wedge \cdots \wedge v_n = v_1 \wedge \phi(v_2) \wedge \cdots \wedge \phi(v_n)$$

für alle $v_1, \ldots, v_n$ aus V.

Lemma 8.3.9. *Wenn ϕ (bezüglich einer Basis von V) durch die Matrix A dargestellt wird, wird* $\mathrm{adj}(\phi)$ *durch die Adjunkte* $\mathrm{adj}(A)$ *dargestellt.*

Man nennt daher $\mathrm{adj}(\phi)$ die Adjunkte von ϕ.

Beweis. Sei $e_1, \ldots, e_n$ eine Basis von V. ϕ sei in dieser Basis durch die Matrix A dargestellt und $\mathrm{adj}(\phi)$ durch B. Wir fixieren zwei Indizes i und j. Dann ist einerseits[4]

$$\begin{aligned}
\mathrm{adj}(\phi)(e_j) \wedge e_1 \wedge \cdots \widehat{e_i} \cdots \wedge e_n &= (\beta_{ij} e_i) \wedge e_1 \wedge \cdots \widehat{e_i} \cdots \wedge e_n \\
&= (-1)^{i-1}\beta_{ij}(e_1 \wedge \cdots \wedge e_n)
\end{aligned}$$

und andererseits, wenn $a_1, \ldots, a_n$ die Spalten von A sind,

$$\begin{aligned}
e_j \wedge \phi(e_1) \wedge \cdots \widehat{\phi(e_i)} \cdots \wedge \phi(e_n) &= \det(\mathrm{e}_j\, a_1 \cdots \widehat{a_i} \cdots a_n)(e_1 \wedge \cdots \wedge e_n) \\
&= (-1)^{j-1} \det(A_{ji})(e_1 \wedge \cdots \wedge e_n).
\end{aligned}$$

Daraus folgt $\beta_{ij} = (-1)^{i+j} \det(A_{ji})$. Das ist die Behauptung. □

[4] "$\widehat{e_i}$" bedeutet, daß dieser Term *fehlt*

8.4 Äußere Algebra und Dualität

V sei diesem Abschnitt ein n–dimensionaler K–Vektorraum.

Lemma 8.4.1. *Durch*

$$\langle \lambda_1 \wedge \cdots \wedge \lambda_p, v_1 \wedge \cdots \wedge v_p \rangle = \det \begin{pmatrix} \lambda_1(v_1) & \cdots & \lambda_1(v_p) \\ \vdots & \ldots & \vdots \\ \lambda_p(v_1) & \cdots & \lambda_p(v_p) \end{pmatrix}$$

wird eine bilineare Abbildung

$$\langle\ ,\ \rangle : \bigwedge^p V^* \times \bigwedge^p V \to K$$

definiert.

Wenn $p = 0$, nehmen wir für $\langle\ ,\ \rangle$ die Multiplikation $K \times K \to K$.

Beweis. Wir müssen zeigen, daß $\det(\lambda_i(v_j))$ alternierend in $\lambda_1, \ldots, \lambda_p$ und alternierend in $v_1, \ldots, v_p$ ist. Das folgt aber sofort aus der Tatsache, daß die Determinante eine alternierende Funktion sowohl ihrer Zeilen (4.3.5) als auch ihrer Spalten ist. □

Der nächste Satz zeigt, daß man für endlichdimensionale V den Raum A^pV der p–Formen auf V mit $\bigwedge^p(V^*)$ identifizieren kann.

Satz 8.4.2. *Wenn V endlichdimensional ist, bilden $\bigwedge^p V^*$ und $\bigwedge^p V$ vermöge der eben definierten bilinearen Abbildung ein duales Paar.*

Beweis. Sei $e_1, \ldots, e_n$ eine Basis von V und $e_1^*, \ldots, e_n^*$ die duale Basis von V^*. Dann sind die Basen (e_I) und (e_I^*) von $\bigwedge^p V$ und $\bigwedge^p V^*$ dual zueinander. Wenn nämlich $I = \{k_1, \ldots, k_p\}$, ist

$$\langle e_I^*, e_I \rangle = \det(e_{k_i}^*(e_{k_j})) = \det(\mathbf{I}) = 1.$$

Wenn $J \neq I$, ist $\langle e_I^*, e_J \rangle$ Determinante einer Matrix, die eine Nullspalte (und eine Nullzeile) hat, also gleich Null. □

Mit Hilfe der kanonischen Isomorphismen

$$\mathrm{A}^pV \cong (\bigwedge^p V)^* \cong \bigwedge^p(V^*)$$

übertragen wir die Algebrastruktur von $\bigwedge(V^*)$ auf

$$\mathrm{A}V = \bigoplus_{p \geq 0} \mathrm{A}^pV,$$

die *Algebra der alternierenden Formen.*

Satz 8.4.3. *Sei V endlichdimensional, μ eine p–Form und ν eine q–Form auf V. Dann berechnet sich die $p+q$–Form $\mu \wedge \nu$ durch*

$$(\mu \wedge \nu)(v_1, \ldots, v_{p+q}) = \sum \epsilon_{I,J}\, \mu(v_{i_1}, \ldots, v_{i_p})\, \nu(v_{j_1}, \ldots, v_{j_q}). \tag{8.3}$$

Dabei wird über alle Zerlegungen von $\{1, \ldots, p+q\}$ in eine Menge $I = \{i_1 < \ldots < i_p\}$ und eine Menge $J = \{j_1 < \ldots < j_q\}$ summiert.

Beweis. Wenn wir die Summanden der rechten Seite schreiben als

$$RS_{I,J} = \operatorname{sign}\begin{pmatrix} i_1 & \cdots & i_p & j_1 & \cdots & j_q \\ 1 & \cdots & p & p+1 & \cdots & p+q \end{pmatrix} \mu(v_{i_1}, \ldots, v_{i_p})\, \nu(v_{j_1}, \ldots, v_{j_q})$$

sehen wir, daß der Wert nur von $I = \{i_1, \ldots, i_p\}$ und $J = \{j_1, \ldots, j_q\}$ abhängt, nicht aber von der richtigen Anordnung der $i_1, \ldots, i_p$ und $j_1, \ldots, j_q$. Denn, wenn wir zum Beispiel die Reihenfolge der $i_1, \ldots$ ändern, ändert sich die Signatur der Permutation in der gleichen Weise wie das Vorzeichen von μ. Daraus folgt nun leicht, daß die rechte Seite von (8.3) alternierend in $v_1, \ldots, v_{p+q}$ ist. Um das einzusehen, nehmen wir an, daß $v_i = v_j$ für zwei verschieden i und j. Dann ist $RS_{I,J} = 0$, wenn i und j beide in I oder beide in J liegen. Wenn $i \in I$ und $j \in J$, heben sich $RS_{I,J}$ und $RS_{I\cup\{j\}\setminus\{i\}, J\cup\{i\}\setminus\{j\}}$ gegenseitig weg, weil

$$\begin{pmatrix} i_1 & \cdots i \cdots & i_p & j_1 & \cdots j \cdots & j_q \\ 1 & \cdots & p & p+1 & \cdots & p+q \end{pmatrix}$$

und

$$\begin{pmatrix} i_1 & \cdots j \cdots & i_p & j_1 & \cdots i \cdots & j_q \\ 1 & \cdots & p & p+1 & \cdots & p+q \end{pmatrix}$$

verschiedene Vorzeichen haben.

Wir können also annehmen, daß die v_i die ersten $p+q$ Elemente einer Basis $e_1, \ldots, e_n$ sind. Weil beide Seiten von (8.3) bilinear in μ und ν sind, können wir annehmen, daß μ und ν zwei Basiselemente e_{I_0} und e_{J_0} von $\bigwedge^p V$ und $\bigwedge^q V$ sind. Dann ist, für aufsteigend geordnete Indizes, $\mu(e_{i_1}, \ldots, e_{i_p})$ (bzw. $\nu(e_{j_1}, \ldots, e_{j_q})$) gleich 1, wenn $I = I_0$ (bzw. $J = J_0$), und gleich 0 sonst.

Beide Seiten von (8.3) sind Null, außer wenn I_0 und J_0 eine Partition von $\{1, \ldots, p+q\}$ bilden. Dann ist aber $\mu \wedge \nu = \epsilon_{I_0,J_0} e_{\{1,\ldots,p+q\}}$ und

$$(\mu \wedge \nu)(e_1, \ldots, e_{p+q}) = \epsilon_{I_0,J_0}.$$

Auf der rechten Seite ist der einzige nicht verschwindende Summand

$$R_{I_0,J_0} = \epsilon_{I_0,J_0}.$$

□

Wenn $p = 1$, d.h. $\mu \in V^*$, hat die Formel die einfache Gestalt:

$$(\mu \wedge \nu)(v_1, \ldots, v_{q+1}) = \sum_{i=1}^{q+1} (-1)^{i-1} \mu(v_i) \nu(v_1, \ldots, \hat{v}_i, \ldots, v_{q+1}).$$

Dabei bedeutet $\hat{v}$, daß die Variable v ausgelassen wird.

Folgerung 8.4.4. *Wenn K die Charakteristik 0 hat, ist*

$$(\mu \wedge \nu)(v_1, \dots, v_{p+q}) = \sum_{\pi \in \mathrm{S}_n} \frac{\mathrm{sign}(\pi)}{p!\, q!}\, \mu(v_{\pi(1)}, \dots, v_{\pi(p)})\, \nu(v_{\pi(p+1)}, \dots, v_{\pi(p+q)}).$$

8.5 Die äußere Algebra eines euklidischen Raumes

Sei V ein n–dimensionaler euklidischer Vektorraum. Jede Linearform hat die Form $x \mapsto (a, x)$ für ein eindeutig bestimmtes $a \in V$. Wir können auf diese Weise V^* mit V identifizieren. Daraus ergibt sich auf natürliche Weise eine Identifizierung von $\bigwedge^p V^*$ und $\bigwedge^p V$. Es ist dann für $x_1, \dots, x_p, y_1, \dots, y_p \in V$

$$\langle x_1 \wedge \dots \wedge x_p, y_1 \wedge \dots \wedge y_p \rangle = \det \begin{pmatrix} (x_1, y_1) & \cdots & (x_1, y_p) \\ \vdots & \dots & \vdots \\ (x_p, y_1) & \cdots & (x_p, y_p) \end{pmatrix}.$$

Lemma 8.5.1. *Die $\bigwedge^p V$ mit der Bilinearform $\langle\ ,\ \rangle$ sind euklidische Vektorräume.*

Beweis. Wenn $e_1, \dots, e_n$ eine Orthonormalbasis von V ist, ist (e_I) eine Orthonormalbasis von $\bigwedge^p V$. □

Lemma 8.5.2. *Eine n–Form $\mu \in \bigwedge^n V^* = \bigwedge^n V$ ist genau dann normiert (siehe Seite 102), wenn $\|\mu\| = 1$.*

Beweis. Sei $e_1, \dots, e_n$ eine Orthonormalbasis und $\mu = e_1 \wedge \dots \wedge e_n$ als n-Form aufgefaßt. Die Gleichung $\langle e_1 \wedge \dots \wedge e_n, e_1 \wedge \dots \wedge e_n \rangle = \det(e_i, e_j) = 1$ hat dann eine doppelte Bedeutung:

1. $\mu(e_1, \dots, e_n) = 1$, dh. μ ist normiert.
2. $\langle \mu, \mu \rangle = 1$, d.h. $\|\mu\| = 1$.

Weil alle n–Formen Vielfache von μ sind, folgt die Behauptung. □

Wir fixieren eine Orientierung von V, indem wir eine der zwei normierten n–Formen festhalten. Nennen wir sie μ_0. Nach 8.3.7 werden, wenn $p+q=n$, $\bigwedge^p V$ und $\bigwedge^q V$ vermöge $x, y \mapsto \mu_0(x \wedge y)$ zu einem dualen Paar. Es gibt also einen Vektorraumisomorphismus (die *Hodge–Abbildung*) $\overline{} : \bigwedge^p V \to \bigwedge^q V$, definiert[5]

$$\mu_0(x \wedge y) = \langle \overline{x}, y \rangle$$

(für alle $y \in \bigwedge^q V$).

[5]Die Hodgeabbildung setzt sich also zusammen aus zwei Isomorphismen $\bigwedge^p V \to \bigwedge^q V^* \to \bigwedge^q V$. Der erste Isomorphismus kommt von der Paarung $\mu_0(x, y)$ zwischen $\bigwedge^p V$ und $\bigwedge^q V$, der zweite durch die euklidische Bilinearform $\langle x, y \rangle$ auf $\bigwedge^q V$.

Satz 8.5.3.

1. $\overline{\mu_0} = 1$

2. *Wenn* $e_1, \dots, e_n$ *eine positiv orientierte Orthonormalbasis von* V *ist*[6], *ist für alle* $I \subset \{1, \dots, n\}$
$$\overline{e_I} = \epsilon_{I,\overline{I}}\, e_{\overline{I}}.$$
Dabei ist $\overline{I} = \{1, \dots, n\} \setminus I$.

3. $\overline{\overline{x}} = (-1)^{pq}\, x$.

4. $\overline{}$ *respektiert die euklidische Bilinearform:* $\langle \overline{x}, \overline{z} \rangle \; = \; \langle x, z \rangle$

Beweis.

1) Eigentlich meinen wir statt μ_0 den entsprechenden n–Vektor f_0. Es gilt $\mu_0(f_0) = \langle f_0, f_0 \rangle = 1$. Also
$$\mu_0(f_0 \wedge y) = y = \langle 1, y \rangle.$$

2) Bezüglich der Paarung $\mu_0(x \wedge y)$ ist die Basis $\left(\epsilon_{I,\overline{I}}\, e_{\overline{I}}\right)$ dual zur Basis (e_I) (Beweis von 8.3.7). Bezüglich der Paarung $\langle \;,\; \rangle$ ist (e_I) zu sich selbst dual. Daraus folgt die Behauptung.

3) Das folgt aus 2) und $\epsilon_{J,I} = (-1)^{pq}\, \epsilon_{I,J}$.

4) Folgt aus 3).
$$\begin{aligned} \langle \overline{x}, \overline{z} \rangle \; &= \mu_0(x \wedge \overline{z}) = (-1)^{pq} \mu_0(\overline{z} \wedge x) \\ &= (-1)^{pq} \langle \overline{\overline{z}}, x \rangle \; = \; \langle z, x \rangle \; = \; \langle x, z \rangle \end{aligned}$$

□

Sei schließlich V ein dreidimensionaler euklidischer Vektorraum mit einer einer ausgezeichneten normierten 3–Form μ. (Bis auf Isomorphie ist das der $\mathbb{R}^3$ mit der Standardbilinearform und der Standarddeterminantenform mit $\mu(\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3) = 1$ für die kanonische Basis.)

Definition. *Das Kreuzprodukt* $\times : V \times V \to V$ *ist definiert durch* $x \times y = \overline{x \wedge y}$.

Für den $\mathbb{R}^3$ mit der Standardbilinearform und der Standarddeterminantenform stimmt das jetzt definierte Kreuzprodukt mit dem auf Seite 39 definierten Kreuzprodukt überein. Aus der Definition folgt, daß $\times$ unter der Wirkung aller $\phi \in \mathrm{SO}_3(\mathbb{R})$ invariant bleibt, daß also $\phi(x \times y) = \phi(x) \times \phi(y)$ für alle $x, y \in \mathbb{R}^3$.

$x \times y$ ist eindeutig bestimmt durch die Gültigkeit von
$$\mu(x, y, z) = (x \times y, z)$$
für alle $z \in V$.

[6] Das heißt, wenn $\mu_0(e_1, \dots, e_n) = 1$

Satz 8.5.4. *Für das Kreuzprodukt gelten folgende Rechenregeln:*

$$
\begin{array}{lrcl}
(1) & a \times b & = & -b \times a \\
(2) & (a \times b, c) & = & \mu(a, b, c) \\
(3) & ((a \times b), (c \times d)) & = & (a, c)(b, d) - (a, d)(b, c) \\
(4) & (a \times b) \times c & = & (a, c)b - (b, c)a \\
(5) & (a \times b) \times c + (b \times c) \times a + (c \times a) \times b & = & 0
\end{array}
$$

Bemerkung. *Eine Algebra, in der* (1) *und die Jacobiidentität* (5) *gelten, heißt Liealgebra.*[7]

Beweis.

(1) Das folgt aus $a \wedge b = -b \wedge a$ (Gleichung (8.2), Seite 131).

(2) Das ist die Definition des Kreuzprodukts.

(3) Weil $^{-}$ das Skalarprodukt erhält, ist

$$((a \times b), (c \times d)) = \langle a \wedge b, c \wedge d \rangle = (a, c)(b, d) - (a, d)(b, c).$$

(4) Wir multiplizieren $(a \times b) \times c$ mit einem beliebigen Vektor x.

$$
\begin{aligned}
((a \times b) \times c\, ,\ x) &= \mu(a \times b, c, x) = \mu(c, x, a \times b) \\
&= (c \times x, a \times b) = (c, a)(x, b) - (c, b)(x, a) \\
&= ((a, c)b - (b, c)a\, ,\ x)
\end{aligned}
$$

Daraus folgt die Behauptung.

(5) Folgt sofort aus (4). □

[7] Wenn der Grundkörper die Charakteristik 2 hat, fordert man $a \times a = 0$. statt (1), vgl. Lemma 4.2.3.

Index