Professor Dr. Jürgen Ritter
Lehrstuhl für Algebra und Zahlentheorie
Institut für Mathematik · Universität Augsburg

Vorlesung Lineare Algebra 2, Wintersemester 2005/06

12. Kapitel: Geometrie II; Längen- und winkeltreue Abbildungen des 2- und 3-dimensionalen affinen Euklidischen Raumes in koordinatenfreier Darstellung. Das Vektorprodukt.

Wir erinnern an die Beschreibung der orthogonalen Matrizen A (vgl. Kapitel 9)

$$A \in \mathbb{R}_{2\times 2} \quad : \quad \begin{cases} \det A = 1 \quad \Longrightarrow \quad A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \text{ mit } 0 \le \varphi < 2\pi \\[2ex] \det A = -1 \quad \Longrightarrow \quad A \overset{o}{\sim} \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \end{cases}$$

$$A \in \mathbb{R}_{3\times 3} \quad : \quad A \overset{o}{\sim} \begin{pmatrix} 1 & & \\ & 1 & \\ & & -1 \end{pmatrix}, \begin{pmatrix} 1 & & \\ & \cos\varphi & \sin\varphi \\ & -\sin\varphi & \cos\varphi \end{pmatrix}, \begin{pmatrix} -1 & & \\ & \cos\varphi & \sin\varphi \\ & -\sin\varphi & \cos\varphi \end{pmatrix}$$

$$A \in \mathbb{C}_{n\times n} \quad : \quad A \overset{o}{\sim} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \ |\lambda_i| = 1 \ ;$$

hier bedeutet $\overset{o}{\sim}$ orthogonale Konjugation.

Sei nun $A \in \mathfrak{O}_n(\mathbb{R})$, n beliebig. Das reelle Polynom $m_A(x)$ zerfällt in lineare und quadratische irreduzible Faktoren; entsprechend seien $\lambda_1, \ldots, \lambda_r$ die reellen und $\lambda_{r+1}, \overline{\lambda}_{r+1}, \ldots, \lambda_{r+s}, \overline{\lambda}_{r+s}$ die nichtreellen komplexen Eigenwerte von A (also $r + 2s = n$). Nun gilt: Ist v_1 Eigenvektor zum reellen Eigenwert λ_1, so ist $\langle v_1\rangle^\perp$ stabil unter A und

$$A \overset{o}{\sim} \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & A_2 & \\ 0 & & & \end{pmatrix}$$

mit $A_2 \in \mathfrak{O}_{n-1}(\mathbb{R})$; des weiteren hat A_2 die Eigenwerte $\lambda_2, \ldots, \lambda_r, \lambda_{r+1}, \overline{\lambda}_{r+1}, \ldots, \lambda_{r+s}, \overline{\lambda}_{r+s}$. Mittels Induktion erreicht man

$$A \overset{o}{\sim} \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_r & \\ & & & B \end{pmatrix}, \ B \in \mathfrak{O}_{n-r}(\mathbb{R}).$$

Natürlich hat B die Eigenwerte $\lambda_{r+1}, \overline{\lambda}_{r+1}, \ldots, \lambda_{r+s}, \overline{\lambda}_{r+s}$.

Fasse B als komplexe (orthogonale) Matrix auf und wähle einen Eigenvektor w zu λ_{r+1}. Dann ist dasjenige $\overline{w}$ Eigenvektor zu $\overline{\lambda}_{r+1}$, dessen i-te Koordinate konjugiert komplex zur i-ten Koordinate von w ist. Insbesondere sind $w+\overline{w}$, $i(w-\overline{w})$ reelle, linear unabhängige Vektoren, deren Aufspann unter B stabil ist und auf dem B die Matrix $\begin{pmatrix} \cos\varphi_1 & \sin\varphi_1 \\ -\sin\varphi_1 & \cos\varphi_1 \end{pmatrix}$ mit $\lambda_{r+1} = \cos\varphi_1 + i\sin\varphi_1$ induziert (beachte $|\lambda_{r+i}| = 1$). Also

$$A \overset{o}{\sim} \begin{pmatrix} \lambda_1 &&&&& \\ & \ddots &&&& \\ && \lambda_r &&& \\ &&& K_1 && \\ &&&& \ddots & \\ &&&&& K_s \end{pmatrix} \quad , \quad K_\nu = \begin{pmatrix} \cos\varphi_\nu & \sin\varphi_\nu \\ -\sin\varphi_\nu & \cos\varphi_\nu \end{pmatrix}$$

mit $\lambda_{r+\nu} = \cos\varphi_\nu + i\sin\varphi_\nu$.

Sämtliche obigen orthogonalen Konjugationen sind abhängig von der jeweiligen Matrix A; die gleichzeitige Konjugation von mehreren orthogonalen Matrizen A in die gezeigte einfache Form ist i.A. nicht möglich. Für $\mathfrak{O}_2(\mathbb{R})$ und $\mathfrak{O}_3(\mathbb{R})$ helfen hier jedoch spezielle Realisierungen des V_2 bzw. V_3 weiter.

Im folgenden sei V_2 der 2-dimensionale reelle Standardvektorraum mit dem symmetrischen Skalarprodukt $(\cdot,\cdot)$, bezüglich dem e_1, e_2 eine Orthonormalbasis ist.
Auf $\mathbb{C}$, dem Körper der komplexen Zahlen, den wir als zweidimensionalen reellen Vektorraum mit Basis $1, i$ auffassen, definieren wir ein Skalarprodukt durch

$$(z_1, z_2) = 1/2(z_1\overline{z}_2 + \overline{z}_1 z_2).$$

Dann sind $1, i$ ein Paar orthonormaler Vektoren ($\overline{z}$ bezeichnet, wie üblich, die zu z konjugiert komplexe Zahl). Über $1 \leftrightarrow e_1, i \leftrightarrow e_2$ wird nun V_2 mit $\mathbb{C}$ identifiziert; diese Identifikation respektiert die jeweiligen Skalarprodukte. Man beachte, daß mit iz für $0 \neq z \in \mathbb{C}$ ein eindeutig bestimmter zu z senkrechter Vektor der gleichen Länge wie z bestimmt ist.

Sei schließlich f eine orthogonale lineare Abbildung von V_2 in sich, die wir als orthogonale Abbildung von $\mathbb{C}$ auffassen.

Satz 12.A. *Falls* $\det f = 1$*, ist* f *von der Form* $f(z) = uz$ *mit einem* $u \in \mathbb{C}$ *der Länge* 1*; falls* $\det f = -1$*, ist* f *von der Form* $f(z) = u\overline{z}$ *mit* u *wie oben.*

Bemerkung Ist $\det f = 1$, so ist f die Drehung um den Winkel φ, falls $u = \cos\varphi + i\sin\varphi$. Ist $\det f = -1$, so ist f die Hintereinanderausführung der Spiegelung an der reellen Achse und obiger Drehung.

Der Fall der Dimension 3

Den 3-dimensionalen reellen Vektorraum V_3 mit Orthonormalbasis e_1, e_2, e_3 identifizieren wir mit

$$\mathbb{H}' = \{\alpha_1 i + \alpha_2 j + \alpha_3 k \;:\; \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}\} \subset \mathbb{H}$$

über $e_1 \leftrightarrow i, e_2 \leftrightarrow j, e_3 \leftrightarrow k$; $\mathbb{H}$ bezeichnet dabei wieder den Schiefkörper der Quaternionen. Für $h = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k \in \mathbb{H}$ setze $\overline{h} = \alpha_0 - \alpha_1 i - \alpha_2 j - \alpha_3 k$. Auf dem reellen Vektorraum $\mathbb{H}$ ist dann durch

$$(h_1, h_2) = \frac{1}{2}(h_1 \overline{h}_2 + h_2 \overline{h}_1)$$

ein Skalarprodukt mit Orthonormalbasis $1, i, j, k$ definiert.

Beobachtungen

1. $(h, h) = h\overline{h}$
2. $h \in \mathbb{H}' \Longleftrightarrow h \perp 1 \Longleftrightarrow \overline{h} = -h \Longrightarrow h^2 = -h\overline{h} \in \mathbb{R}_{\leq 0}$
3. $(h_1, h_2 h_3) = (h_1 \overline{h}_3, h_2)$
4. $h_1, h_2 \in \mathbb{H}' \Longrightarrow [\, h_1 \perp h_2 \Longleftrightarrow h_1 h_2 \in \mathbb{H}' \Longleftrightarrow h_1 h_2 = -h_2 h_1 \,]$.
5. sind $h_1, h_2 \in \mathbb{H}'$, so gilt für ihr Vektorprodukt $\quad h_1 \times h_2 = \frac{1}{2}(h_1 h_2 - h_2 h_1)$

Sei nun f ein orthogonaler Endomorphismus von $\mathbb{H}'$. Nenne $f(i) = I, f(j) = J, f(k) = K$. Es gilt

$$I^2 = J^2 = K^2 = -1,\; IJ = -JI,\; IK = -KI,\; JK = -KJ \quad \text{sowie}$$

$$IJ = \begin{cases} K & \text{falls } \det f = +1 \\ -K & \text{falls } \det f = -1 \;. \end{cases}$$

Beachte hier, daß $\det f = \det(I, J, K)$ und $\det(I, J, IJ) \geq 0$, weil $IJ = I \times J$ ist.

Lemma 12.a. *Es gibt ein* $a \in \mathbb{H}$ *mit* $u = a - Iai - Jaj - Kak \neq 0$.

Ansonsten erreiche einen Widerspruch durch Ersetzen von a in

$$a = Iai + Jaj + Kak$$

durch ai und durch Multiplikation obiger Gleichung von rechts mit i; entsprechend mit j und k anstelle von i.

Satz 12.B. $f(x) = \begin{cases} uxu^{-1} & \textit{falls } \det f = 1 \\ -uxu^{-1} & \textit{falls } \det f = -1 \;. \end{cases}$

Hier ist $x \in \mathbb{H}'$ und $u \in \mathbb{H}$ wie in Lemma 12.a.

Zum Beweis überlegt man sich, daß $uxu^{-1} \in \mathbb{H}'$ für $x \in \mathbb{H}'$, daß $x \mapsto uxu^{-1}$ linear ist und mit f auf der Basis i, j, k übereinstimmt.

Im Satz 12.B darf man natürlich ohne Einschränkung u als von der Länge 1 annehmen.

Die allgemeine längen- und winkeltreue Abbildung f des 3-dimensionalen affinen Euklidischen Raumes $A^3 = (P, \mathbb{H}')$ sieht folglich so aus

$$f(x) = \pm uxu^{-1} + c\,, \quad x, c \in \mathbb{H}',\; u \in \mathbb{H}\,, |u| = 1 \;;$$

das c steht für die Translation mit c.

Geometrische Interpretation im Falle des Plus-Zeichens [1]

Wir berechnen Fixpunkte von $f(x) = uxu^{-1} + c$ und schreiben dazu

$$\begin{aligned} u &= \alpha + u' \,,\ \alpha \in \mathbb{R} \,,\ u' \in \mathbb{H}',\ \alpha^2 - (u')^2 = 1 \,,\ \alpha = \cos\varphi \\ x &= y + \mu u' \text{ mit einem } y \in \mathbb{H}' \text{ so daß } y \perp u' \\ c &= d + \lambda u' \,,\ d \perp u' \,. \end{aligned}$$

1. $u(y + \mu u')u^{-1} + d + \lambda u' = y + \mu u' \iff \lambda = 0 \ \&\ [\, y = (1-u^2)^{-1} d \text{ oder } u = \pm 1 \,]$.

   Beweis: $uyu^{-1} = u^2 y$, weil $y \perp u'$; $uu' = u'u$. Also $(1-u^2)y = d + \lambda u'$. Wende darauf $(\ \ ,u')$ an: $(\,(1-u^2)y, u'\,) = (d + \lambda u', u') = \lambda(u', u')$; $(1-u^2)y = (1-\alpha^2-(u')^2-2\alpha u')y \overset{\text{def}}{=} (\beta - 2\alpha u')y$ mit $\beta \in \mathbb{R}$; $(\,(1-u^2)y, u'\,) = (\beta y - 2\alpha u'y, u') = -2\alpha(u'y, u') = 2\alpha(yu', u')$; $(yu', u') = \frac{1}{2}(-yu'u' + u'\overline{yu'}) = \frac{1}{2}(-y(u')^2 + (u')^2 y) = 0$. Damit ist $\lambda = 0$ und $y = (1-u^2)^{-1}d$, sofern $y \neq \pm 1$.

2. Falls $u = \pm 1$ ist, ist f die Translation zu c.

3. Falls $u \neq \pm 1$ und $c \perp u'$, also $\lambda = 0$, ist $\{(1-u^2)^{-1}d + \nu u' : \nu \in \mathbb{R}\}$ die Drehachse von f und f die Rotation um 2φ in der Ebene senkrecht zur Drehachse.

   Zu beweisen ist nur die Winkelaussage. Sei dazu $x \perp u'$. Wegen

   $$uxu^{-1} = u^2 x \,,\ |x| = \sqrt{x\overline{x}} \,,\ |u^2 x| = \sqrt{u^2 x \overline{u^2 x}} = \sqrt{u^2 x \overline{x}\, \overline{u}^2} = \sqrt{x\overline{x}}(u\overline{u})$$

   $$\text{und} \quad (u^2 x, x) = \frac{1}{2}(u^2 x\overline{x} + x\overline{u^2 x}) = \frac{1}{2} x\overline{x}(u^2 + \overline{u}^2)$$

   folgt

   $$\frac{(u^2x, x)}{|u^2x||x|} = \frac{1}{2}\frac{u^2 + \overline{u}^2}{u\overline{u}} = \frac{\alpha^2 + (u')^2}{1} = 2\alpha^2 - 1 = 2\cos^2\varphi - (\cos^2\varphi + \sin^2\varphi) = \cos(2\varphi) \,.$$

4. Ist die Voraussetzung $c \perp u'$ nicht erfüllt, so hänge dem f eine geeignete Translation in Richtung u' an, um dies zu erreichen. Es handelt sich dann bei f selbst um eine “Schraubung”.

In diesem Zusammenhang machen wir einen kurzen Abstecher in die *dreidimensionale sphärische Geometrie*; dazu brauchen wir das Vektorprodukt. Ein sphärisches Dreieck im $\mathbb{R}^3$ ist durch drei linear unabhängige Vektoren $a, b, c \in \mathbb{R}^3$ gegeben. Wir orientieren es durch die Forderung $\det(a, b, c) > 0$ und definieren dann die Winkel zwischen den Seiten durch

$$\cos(C) = (a, b) \,,\ \cos(B) = (a, c) \,,\ \cos(A) = (b, c) \,,$$

sofern alle drei Vektoren die Länge 1 haben, und den Winkel zwischen den von je zwei dieser Vektoren aufgespannten Ebenen durch

$$\cos\alpha = \frac{(a \times c, a \times b)}{|a \times c||a \times b|} \,,\ \text{etc.}$$

[1] Die für den Fall des Minus-Zeichens anzugeben, ist eine Übungsaufgabe.

Beispiel Wir berechnen die Entfernung vom Ort A zum Ort B, die durch ihre geographischen Breiten $-\pi/2 \le \beta_A, \beta_B \le \pi/2$ und Längen $-\pi \le \lambda_A, \lambda_B \le \pi$ gegeben seien [2]. Dazu arbeiten wir im sphärischen Dreieck ABN, wobei N der Nordpol ist. Der Winkel des Dreiecks bei N zwischen den Ebenen durch O, A, N und O, B, N heiße α; hier bezeichne O den Erdmittelpunkt. Also: $\alpha = \lambda_B - \lambda_A$. Die Längen der Seiten des sphärischen Dreicks ABN sind (im Bogenmaß gemessen)

$$AN = \frac{\pi}{2} - \beta_A \ , \ BN = \frac{\pi}{2} - \beta_B \ .$$

Gesucht ist $\cos AB$. Dazu nützen wir den sogenannten Kosinussatz aus

$$\cos AB = \cos AN \cos BN + \sin AN \sin BN \cos\alpha \ ,$$

der wiederum aus $\cos AB = (a, b)$ mit $a = OA,\ b = OB$ etc. und $\cos\alpha = (a\times n, b\times n)/|a\times n||b\times n|$ bei Normierung des Erdradius auf 1 resultiert.

13. Kapitel : Bilinearformen

Das Skalarprodukt auf dem reellen Vektorraum V_n ist ein Beispiel einer Bilinearform auf V_n. In diesem Kapitel werden einige grundsätzliche Eigenschaften von Bilinearformen auf endlich dimensionalen Vektorräumen über kommutativen Körpern K vorgestellt. Die einzige Bedingung an K ist $2 \ne 0$. Das Ziel unseres Studiums ist, mit Methoden der Linearen Algebra Informationen über quadratische Gleichungen in mehreren Variablen zu gewinnen.

Definition 13.1. *V sei ein n-dimensionaler Vektorraum über K. Eine Bilinearform auf V ist eine Abbildung*

$$(\cdot,\cdot) : V \times V \to K,$$

die additiv in beiden Argumenten ist und die $(\alpha v, w) = (v, \alpha w) = \alpha(v, w)$ für $\alpha \in K,\ v, w \in V$ erfüllt.

1. *Gilt stets $(v, w) = (w, v)$, so heißt die Bilinearform symmetrisch und V ein quadratischer Raum; gilt stets $(v, w) = -(w, v)$, so heißt die Bilinearform schiefsymmetrisch oder alternierend und V ein symplektischer Raum.*
2. *Eine symmetrische oder alternierende Bilinearform heißt nicht-ausgeartet, falls*

$$[\,(v, w) = 0\ (\forall\, w \in V) \Longrightarrow v = 0\,]\ ^3 .$$

Im folgenden ist V stets ein mit einer Bilinearform versehener n-dimensionaler Vektorraum über K.

[2] der Äquator hat also die Breite $\beta = 0$, Greenwich die Länge $\lambda = 0$

[3] Wenn wir künftig von einem nicht-ausgearteten Raum sprechen, meinen wir damit ein quadratisches oder symplektisches V mit nicht-ausgearteter Bilinearform.

Beobachtungen

a) Jede Bilinearform ist die Summe aus einer symmetrischen und einer schiefsymmetrischen Bilinearform. In der Tat, $(v,w) = \langle v,w\rangle + [v,w]$ mit der symmetrischen Bilinearform $\langle v,w\rangle = \frac{1}{2}((v,w)+(w,v))$ und der schiefsymmetrischen Bilinearform $[v,w] = \frac{1}{2}((v,w)-(w,v))$.

b) V ist genau dann symplektisch, wenn $(v,v)=0$ für alle $v\in V$ gilt.

c) Es sei $v_1,\dots,v_n$ eine Basis von V. Zu ihr gehört die Begleitmatrix

$$S_{\{v_i\}} \overset{\text{def}}{=} \Big((v_i,v_j)\Big)_{1\le i,j\le n}.$$

Für $v = \sum\limits_{i=1}^{n}\alpha_i v_i$ und $w = \sum\limits_{j=1}^{n}\beta_j v_j$ gilt $(v,w) = v^T S_{\{v_i\}} w$.

Die Begleitmatrix $S_{\{Av_i\}}$ zur Basis Av_i mit $A\in GL_n(K)$ berechnet sich aus $S_{\{v_i\}}$ über die Formel $S_{\{Av_i\}} = A^T S_{\{v_i\}} A$. Insbesondere: $\det S_{\{Av_i\}} = \det A^2\cdot \det S_{\{v_i\}}$.

V ist quadratisch, genau wenn $S_{\{v_i\}}$ symmetrisch ist, und symplektisch, genau wenn $S_{\{v_i\}}$ schiefsymmetrisch ist, also $S^T_{\{v_i\}} = -S_{\{v_i\}}$ gilt.

Definition 13.2. $D_V = \{\det S_{\{v_i\}}\cdot\alpha^2 : 0\neq\alpha\in K\}\subset K$ *ist die Diskriminante von* V *(wenn* v_i *wieder Basis von* V *ist).*

Satz 13.A. *a)* V *sei ein quadratischer Raum. Dann gilt*

$$(v,w) = \frac{1}{2}((v+w,v+w)-(v,v)-(w,w));$$

die "Quadrate" bestimmen also schon die Bilinearform.

b) Quadratische V *besitzen Orthogonalbasen, also Basen* $e_1,\dots,e_n$ *mit* $(e_i,e_j)=0$ *für* $i\neq j$.

c) V *ist nicht-ausgeartet genau wenn* $D_V\neq 0$. *In diesem Fall ist*

$$v\mapsto\varphi_v : V\to\hat V,\ \varphi_v(w) = (v,w)$$

ein Isomorphismus.

d) Sei V *nicht-ausgeartet. Ist* U *Teilraum von* V, *so hat* $U^\perp$ *die Dimension* $\dim V - \dim U$.

13 a. Quadratische Räume

In diesem Teilkapitel ist V stets quadratisch. Die Bilinearform ist also durch die (v,v), $v\in V$ schon vollständig bestimmt. Wir haben $(v,v) = v^T S v$, wenn S die Begleitmatrix einer Basis bezeichnet, bezüglich der die $v\in V$ als Spaltenvektoren geschrieben sind. Jede Gleichung $v^T S v = \alpha$ stellt eine quadratische Gleichung in den Koordinaten von v dar, die um so zugänglicher ist, je einfacher die Matrix S gebaut ist. S kann zu $A^T S A$ mit $A\in GL_n(K)$ verändert werden; wir versuchen, möglichst einfache Formen unter den $A^T S A$ zu finden (Satz 13.B).

Definition 13.3. *1. $v \in V$ heißt isotrop, falls $(v,v)=0$. Ein Teilraum $U \le V$ heißt isotrop, wenn $(u_1,u_2)=0$ für alle $u_1,u_2 \in U$ gilt.*

2. Jeder 2-dimensionale Teilraum E von V mit Basis

$$u,h\ :\ (u,u)=(h,h)=0,\ (u,h)=1$$

heißt eine hyperbolische Ebene; u,h heißt eine hyperbolische Basis von E.

3. Die Summe U_1+U_2 zweier Teilräume $U_1,U_2 \le V$ wird mit $U_1 \perp U_2$ notiert, wenn

$$U_1 \cap U_2 = 0 \quad \& \quad (u_1,u_2)=0\ (\forall\, u_1 \in U_1\,,\ \forall\, u_2 \in U_2)\,.$$

Satz 13.B. *V sei nicht-ausgeartet.*

1. Zu jedem isotropem $0 \neq u \in V$ existiert ein isotropes $h \in V$ mit $(u,h)=1$. Es gilt: $V = \langle u,h\rangle \perp W$ mit dem nicht-ausgearteten Teilraum $W=\{v \in V : (v,u)=0=(v,h)\}$.

2. $V = E_1 \perp E_2 \perp \ldots \perp E_r \perp U$ mit hyperbolischen Ebenen E_i $(1 \le i \le r \ge 0)$ und einem Teilraum U ohne isotrope Vektoren $\neq 0$. Die Begleitmatrix zu einer aus hyperbolischen Basen u_i,h_i der E_i und einer Orthogonalbasis von U zusammengesetzten Basis von V hat also die Gestalt

$$\begin{pmatrix}
0 & 1 & & & & & & & & \\
1 & 0 & & & & & & & & \\
 & & 0 & 1 & & & & & & \\
 & & 1 & 0 & & & & & & \\
 & & & & \ddots & & & & & \\
 & & & & & 0 & 1 & & & \\
 & & & & & 1 & 0 & & & \\
 & & & & & & & \alpha_{r+1} & & \\
 & & & & & & & & \ddots & \\
 & & & & & & & & & \alpha_n
\end{pmatrix}.$$

Die $\alpha_{r+1},\ldots,\alpha_n$ sind so, daß die quadratische Gleichung

$$\sum_{j=r+1}^{n} \alpha_j x_j^2 = 0 \quad \textit{nur die triviale Lösung hat.}$$

Bei Wahl der Basis $e_i = u_i + \frac{1}{2}h_i$, $f_i = u_i - \frac{1}{2}h_i$ von E_i werden aus den $2{\times}2$-Kästchen $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ die neuen Kästchen $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

$\langle h_1,\ldots,h_r\rangle$ ist ein maximaler isotroper Teilraum von V.

3. Umgekehrt, für einen isotropen Teilraum $U \le V$ mit größtmöglicher Dimension gilt

$$2\dim U \le \dim V\,,$$

und V hat eine Basis u_i, h_i, v_j $(1 \leq i \leq r = \dim U\,,\ 2r+1 \leq j \leq n)$ mit

$$(u_i, u_i) = (h_i, h_i) = 0\,,\ (u_i, h_i) = 1\,,\ (u_i, h_k) = 0 \text{ für } i \neq k\,,$$

$$(u_i, v_j) = 0\,,\ (h_i, v_j) = 0\,,\ (v_j, v_l) = 0 \text{ für } j \neq l\,.$$

Die quadratische Gleichung $\sum\limits_{j=2r+1}^{n} (v_j, v_j) x_j^2 = 0$ *hat nur die triviale Lösung in K.*

4. *(Sylvester) Sei $K = \mathbb{R}$. Dann existiert eine Basis $e_1, \ldots, e_r, \ldots, e_n$ von V mit $(e_i, e_j) = 0$ für $i \neq j$, $(e_i, e_i) = 1$ für $i \leq r$ und $(e_i, e_i) = -1$ für $i \geq r+1$; dabei kann $r = 0$ oder $r = n$ sein. Die Begleitmatrix zu dieser Basis ist also diagonal.*

Ein Fazit aus dem Satz 13.B ist, daß jede quadratische Gleichung $\sum\limits_{1 \leq i \leq j \leq n} \alpha_{ij} x_i x_j = \alpha$ in Diagonalgestalt $\sum\limits_{i=1}^{n} \alpha_i x_i^2 = \alpha$ mit recht guter Information über die Koeffizienten α_i transformiert werden kann [4].

Die im Satz vorkommende Zahl r ist eine Invariante von V; mit anderen Worten: alle maximalen isotropen Teilräume von V haben dieselbe Dimension. Insbesondere ist im Falle $K = \mathbb{R}$ die sogenannte Signatur von V erklärt, das ist die Differenz $s_{\{+1\}} - s_{\{-1\}}$ aus den Anzahlen der $+1$ bzw. -1 in der Diagonalen der durch Sylvester definierten Begleitmatrix

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \end{pmatrix}.$$

Definition 13.4. *Im Falle $s_{\{+1\}} = 0$ heißt die Bilinearform negativ definit, im Fall $s_{\{+1\}} = n$ positiv definit.*

Offenbar gilt

$$s_{\{+1\}} = 0 \Longleftrightarrow (v,v) < 0 \quad \text{für alle } 0 \neq v \in V$$

$$s_{\{+1\}} = n \Longleftrightarrow (v,v) > 0 \quad \text{für alle } 0 \neq v \in V.$$

Satz 13.C. *Die symmetrische, nicht-ausgeartete Bilinearform des reellen Vektorraums V ist genau dann positiv definit, wenn $\det S_k > 0$ für $1 \leq k \leq n$, und negativ definit, wenn $(-1)^k \det S_k > 0$ für $1 \leq k \leq n$. Dabei ist $S_k = ((v_i, v_j))_{1 \leq i,j \leq k}$ und $v_1, \ldots, v_n$ eine Basis von V.*

Der Fall des Skalarproduktes auf V aus Kapitel 9 ist also der einer positiv definiten symmetrischen Bilinearform.

Für $K = \mathbb{R}$ wissen wir tatsächlich mehr.

[4] Die α_i sind aber i.A. nicht die Eigenwerte von (α_{ij})!

Satz 13.D. *Jede Matrix $B \in GL_n(\mathbb{R})$ läßt sich eindeutig als Produkt $B = D \cdot U \cdot O$ schreiben, wobei D eine Diagonalmatrix mit lauter positiven Einträgen auf der Diagonalen, U eine obere Dreiecksmatrix mit lauter Einsen auf der Diagonalen ("unipotent") und $O \in \mathfrak{O}_n(\mathbb{R})$ ist.*

Satz 13.E. *a. Ist die symmetrische Matrix $S \in GL_n(\mathbb{R})$ positiv definit, so existiert ein $A \in GL_n(\mathbb{R})$ mit $S = A \cdot A^T$. Umgekehrt, solche Produkte $A \cdot A^T$ sind symmetrisch und positiv definit.*

b. Ist $S \in GL_n(\mathbb{R})$ symmetrisch und positiv definit, so besitzt S genau eine symmetrische, positiv definite Wurzel $S' \in GL_n(\mathbb{R})$, $(S')^2 = S$.

c. Jedes $B \in GL_n(\mathbb{R})$ ist eindeutig als Produkt $B = PO$ mit symmetrischem, positiv definitem P und orthogonalem O darstellbar. (P heißt die Pfaffsche Matrix zu B.)

S', oder vielleicht besser $\sqrt{S}$, erhält man durch Rückkonjugation mit der orthogonalen Matrix O aus $O^T S O = \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix}$, wobei $\mu_i > 0$ und die μ_i^2 die Eigenwerte von S sind. Die Pfaffsche Matrix P ist $= \sqrt{B \cdot B^T}$.

Zurück zu allgemeinerem K.

Definition 13.5. *Eine lineare Abbildung $f : V \to W$ zwischen den quadratischen Räumen V, W über K heißt isometrisch, falls $(f(v), f(v')) = (v, v')$ für alle $v, v' \in V$ gilt. Ist f surjektiv, so heißt f eine Isometrie.*

Beachte, daß isometrische Abbildungen zwischen nicht-ausgearteten Räumen stets injektiv sind.

Die Menge aller Isometrien von V wird mit $\mathfrak{O}(V)$ notiert. Mit der Wahl einer Basis $v_1, \ldots, v_n$ von V wird sie eine Untergruppe der linearen Gruppe $GL_n(K)$ und besteht aus allen invertierbaren Matrizen A, die $A^T S A = S$ erfüllen, wenn $S = ((v_i, v_j))$ die Begleitmatrix von V zu $v_1, \ldots, v_n$ bezeichnet. $\mathfrak{O}(V)$ heißt die orthogonale Gruppe von V; sie stimmt mit der im 9. Kapitel definierten orthogonalen Gruppe überein, wenn $K = \mathbb{R}$ und die Bilinearform positiv definit ist. Wir interessieren uns für Erzeugende von $\mathfrak{O}(V)$.

1. Sei H eine nicht-ausgeartete Hyperebene in V, also $H \leq V$, $\dim H = n - 1$, $H \cap H^\perp = 0$. Wähle $0 \neq y \in H^\perp$. Dann ist
$$\sigma_H(v) = v - 2\frac{(v, y)}{(y, y)} y$$
ein Element in $\mathfrak{O}(V)$ mit $\sigma_H^2 = \mathrm{id}$, $\det \sigma_H = -1$. σ_H hängt nicht von der Wahl von y ab. Es handelt sich bei σ_H um die Spiegelung an H:
$$\sigma_H(v) = v \text{ für } v \in H\,,\ \sigma_H(y) = -y\ .$$
Man bemerke, daß die nicht-ausgearteten Hyperebenen genau die $\langle y \rangle^\perp$ zu den anisotropen Vektoren $y \in V$ sind.

2. Sind v, w anisotrope Vektoren gleicher Länge, so existiert ein $\sigma \in \mathfrak{O}(V)$ mit $\sigma(v) = w$.

Ein direkte Konsequenz ist

SATZ 13.F. *Jedes $\sigma \in \mathfrak{O}(V)$ ist ein Produkt $\sigma_{H_1} \cdot \ldots \cdot \sigma_{H_r}$ von $r \leq n$ Spiegelungen an nichtausgearteten Hyperebenen H_i $(1 \leq i \leq r)$.*

FOLGERUNG. *1. Gilt $\sigma = \sigma_{H_1} \cdot \ldots \cdot \sigma_{H_r}$ mit $r < n$, so existiert ein $0 \neq v$ mit $\sigma(v) = v$.*

2. n gerade, $\det \sigma = -1 \Rightarrow \exists\, 0 \neq v : \sigma(v) = v$
n ungerade, $\det \sigma = 1 \Rightarrow \exists\, 0 \neq v : \sigma(v) = v$

3. Ist σ Produkt von nicht weniger als n Spiegelungen, so läßt sich $\sigma = \sigma_H \sigma_{H_2} \cdot \ldots \cdot \sigma_{H_n}$ für jede Spiegelung σ_H lösen.

13 b. Symplektische Räume

SATZ 13.G. *V sei ein nicht-ausgearteter symplektischer Vektorraum über K. Es gibt eine Basis $v_1, w_1, v_2, w_2, \ldots, v_s, w_s$ mit*

$$(v_i, v_j) = (w_i, w_j) = 0 \text{ für } 1 \leq i, j \leq n$$

$$(v_i, w_i) = 1 \ , \ (v_i, w_j) = 0 \text{ für } 1 \leq i \neq j \leq n \ .$$

Die Begleitmatrix hat also die Form

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

Insbesondere ist $\dim V = 2s$ *gerade.*

FOLGERUNG. *Ist $S \in K_{n \times n}$ eine schiefsymmetrische Matrix, so ist* $\det S$ *ein Quadrat.*

14. Kapitel: Jordanzerlegungen von Vektorräumen

In diesem Kapitel, das die Diskussionen aus Kapitel 8 aufgreift, ist V ein n-dimensionaler Vektorraum über einem beliebigen kommutativen Körper und f ein Endomorphismus von V mit Matrix A bezüglich einer vorgegebenen Basis $v_1, \ldots, v_n$. Wie üblich sei $m = m_A(x)$ das Minimalpolynom und $\chi_A(x)$ das charakteristische Polynom von A (oder von f). Wir erklären eine Abbildung

$$K[x] \times V \to V \quad \text{durch} \quad (h, v) \mapsto hv \stackrel{\text{def}}{=} h(A) \cdot v \, .$$

Sie ist additiv in beiden Argumenten und erfüllt

$$[\,(h,v) = 0\ (\forall\, v \in V) \iff m|h\,]$$ [5].

Folgende Eigenschaften von $K[x]$ sind für uns wichtig.

1. (Euklidischer Divisionsalgorithmus) Zu $g, 0 \neq h \in K[x]$ existieren Polynome $s, r \in K[x]$ mit $g = sh + r$ und $[\,r = 0$ oder $\mathrm{grad}(r) < \mathrm{grad}(h)\,]$.
2. Ist $d \in K[x]$ der größte gemeinsame Teiler von $g \neq 0$ und $h \neq 0$, also
$$d|g\,,\ d|h\,; \quad d_1|g\,,\ d_1|h \Longrightarrow d_1|d\,; \quad d \text{ ist normiert}\,,$$
so existieren Polynome $s, t \in K[x]$ mit $sg + th = d$.

Beobachtung Zerfällt $m = g_1 g_2$ wie gezeigt und mit $\mathrm{ggT}(g_1, g_2) = 1$, so ist

$$V = V_1 \oplus V_2 \quad \text{mit} \quad V_1 = \ker(g_1(A))\,,\ V_2 = \ker(g_2(A))\,.$$

Des weiteren induziert f durch Einschränkung Endomorphismen f_1, f_2 von V_1 bzw. V_2. Sind A_1 bzw. A_2 zugehörige Matrizen, so ist $A \sim \begin{pmatrix} A_1 & \\ & A_2 \end{pmatrix}$. Es gilt: $m_{A_1}(x) = g_1(x)\,,\ m_{A_2}(x) = g_2(x)\,.$

Definition 14.1. *Für $0 \neq v \in V$ bezeichne $[v]$ den Teilraum $K[x] \cdot v$, also den Aufspann*

$$[v] = \langle v, Av, \ldots, A^{k-1}v \rangle \quad \text{mit} \quad k = \mathrm{grad}(m)\,,$$

und $m_v \in K[x]$ das normierte Polynom kleinsten Grades mit $m_v v = m_v(A)v = 0$

Beachte hier, daß wegen $mv = 0\ (\forall\, v \in V)$ die Vektoren $A^\nu v$ für $\nu \geq k$ linear durch $v, Av, \ldots, A^{k-1}v$ ausdrückbar sind. Insbesondere ist $[v]$ ein A-stabiler Unterraum von V. Die gleiche Begründung, $mv = 0$, liefert die Teilbarkeit $m_v|m$. Die Eindeutigkeit von m_v resultiert aus den weiter oben genannten Eigenschaften von $K[x]$.

Aufgrund der obigen Beobachtung können und wollen wir fürs erste die Bezeichnungen vereinfachen und nehmen nun an, daß $m_A(x) = m = p^e$ mit einem irreduziblen Polynom $p = p(x) \in K[x]$ und einer natürlichen Zahl e gilt.

Lemma 14.a. *Es sei* $m_v(x) = \alpha_{v,0} + \alpha_{v,1}x + \ldots + \alpha_{v,k_v-1}x^{k_v-1} + x^{k_v}$. *Dann folgt:*

1. *$v, Av, \ldots, A^{k_v-1}v$ ist eine Basis von $[v]$.*
2. *Der von A auf $[v]$ vermittelte Endomorphismus wird bezüglich obiger Basis durch die Matrix*
$$\begin{pmatrix} 0 & 0 & \ldots & \ldots & 0 & -\alpha_{v,0} \\ 1 & 0 & \ldots & \ldots & 0 & -\alpha_{v,1} \\ 0 & 1 & \ldots & \ldots & 0 & -\alpha_{v,2} \\ \vdots & \vdots & & & \vdots & \vdots \\ 0 & 0 & \ldots & \ldots & 0 & -\alpha_{v,k_v-2} \\ 0 & 0 & \ldots & \ldots & 1 & -\alpha_{v,k_v-1} \end{pmatrix}$$
beschrieben.

[5] Für Polynome $b, c \in K[x]$ bedeutet $b|c$ (lies b teilt c): $\exists t \in K[x]$ mit $tb = c$.

Lemma 14.b. *Sind für $0 \neq v \in V$ die Vektoren $A^\nu v$, $0 \leq \nu \leq k_v - 1$ linear unabhängig und gilt $A^{k_v} v = -\alpha_0 v - \alpha_1 A v - \ldots - \alpha_{k-1} A^{k_v - 1} v$, so ist $m_v(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_{k-1} x^{k-1} + x^k$. Des weiteren, ist $v_1, \ldots, v_n$ irgendeine Basis von V, so gilt $m_A(x) = \mathrm{kgV}(\, m_{v_1}(x), \ldots, m_{v_n}(x)\,)$.*

Satz 14.A. *Hat $f \in \mathrm{End}_K(V)$ das Minimalpolynom p^e mit irreduziblem $p \in K[x]$ und natürlichem Exponenten $e \geq 1$, so existieren l Vektoren $v_1, \ldots, v_l$ in V mit*

$$V = [v_1] \oplus \ldots \oplus [v_l] \quad \textit{und} \quad m_{v_j} = p^{a_j}\,,\ 1 \leq a_1 \leq \ldots \leq a_l = e\,.$$

Die Zahlen l und $a_1, \ldots, a_l$ sind durch A eindeutig bestimmt (die a_j heißen die Elementarteiler von A).

Ein Beweis ist weiter unten angehangen. Zunächst aber lösen wir uns von der Voraussetzung, das Minimalpolynom $m_A(x)$ von A sei eine Potenz eines irrreduziblen Polynoms $p(x)$, und konstatieren:

Besitzt das Minimalpolynom $m_A(x)$ der Matrix $A \in K_{n\times n}$ die Zerlegung

$$m_A(x) = \prod_{i=1}^{z} p_i^{e_i}(x)$$

in Potenzen von paarweise verschiedenen normierten irreduziblen Polynomen $p_i(x)$, so ist A zu einer Matrix in Kästchenform

$$\begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_z \end{pmatrix}$$

konjugiert. Es gilt $m_{A_i} = p_i^{e_i}$. Die Kästchen A_i ihrerseits haben die Form

$$\begin{pmatrix} B_{i1} & & \\ & \ddots & \\ & & B_{il_i} \end{pmatrix} \quad \text{mit} \quad B_{ij} = \begin{pmatrix} 0 & 0 & \ldots & \ldots & 0 & -\alpha_{ij,0} \\ 1 & 0 & \ldots & \ldots & 0 & -\alpha_{ij,1} \\ 0 & 1 & \ldots & \ldots & 0 & -\alpha_{ij,2} \\ \vdots & \vdots & & & \vdots & \vdots \\ 0 & 0 & \ldots & \ldots & 0 & -\alpha_{ij,k_{ij}-2} \\ 0 & 0 & \ldots & \ldots & 1 & -\alpha_{ij,k_{ij}-1} \end{pmatrix},$$

und es gilt $m_{B_{ij}}(x) = \alpha_{ij,0} + \alpha_{ij,1} x + \ldots + \alpha_{ij,k_{ij}-1} x^{k_{ij}-1} + x^{k_{ij}} = p_i^{a_{ij}}$. Des weiteren $a_{i1} \leq a_{i2} \leq \ldots \leq a_{il_i} = e_i$.

Durch geschickte Umgruppierung erhält man auch

A kann in die Kästchenform

$$\begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_s \end{pmatrix} \quad \text{mit} \quad B_i = \begin{pmatrix} 0 & 0 & \ldots & \ldots & 0 & -\alpha_{i,0} \\ 1 & 0 & \ldots & \ldots & 0 & -\alpha_{i,1} \\ 0 & 1 & \ldots & \ldots & 0 & -\alpha_{i,2} \\ \vdots & \vdots & & & \vdots & \vdots \\ 0 & 0 & \ldots & \ldots & 0 & -\alpha_{i,k_i-2} \\ 0 & 0 & \ldots & \ldots & 1 & -\alpha_{i,k_i-1} \end{pmatrix}$$

transformiert werden und es gilt

$$m_{B_1}(x) \mid m_{B_2}(x) \mid \cdots \mid m_{B_s}(x)$$
$$m_{B_i}(x) = \alpha_{i,0} + \alpha_{i,1}x + \ldots + \alpha_{i,k_i-1}x^{k_i-1} + x^{k_i}\ .$$

Eine solche Zerlegung von A ist eindeutig. In ihr sind die $m_{B_i}(x)$ die Elementarteiler.

FOLGERUNG. *Der Satz von Cayley-Hamilton gilt für beliebiges kommutatives K.*

Tatsächlich ist $\chi_A(x) = \prod\limits_{i=1}^{z} \chi_{A_i}(x)$ und $\chi_{A_i}(x) = p_i^{\sum_j a_{ij}}$ (die Summe läuft von 1 bis l_i – wir beschränken unsere Argumentation hier wieder auf den Fall $m_A(x) = p(x)^e$).

SATZ 14.B. *Hat $A_i \in K_{n_i\times n_i}$ das Minimalpolynom $m_i(x) = (x-\lambda_i)^{e_i}$ mit $\lambda_i \in K$, so gilt*

$$A_i \sim \begin{pmatrix} B_{i1} & & \\ & \ddots & \\ & & B_{il_i} \end{pmatrix} \text{ mit } B_{ij} \sim \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix}^T .$$

Der Eigenraum $\{v : A_i v = \lambda_i v\}$ von A_i zum Eigenwert λ_i hat die Dimension l_i.

Die Matrix B_{ij} ist die von Lemma 14.a zur Basis $v_{ij}, B_{ij}v_{ij}, \ldots, B_{ij}^{k_{ij}-1}v_{ij}$ von $[v_{ij}]$; für Satz 14.B ersetze diese Basis durch

$$v_{ij}, (B_{ij}-\lambda_i)v_{ij}, \ldots, (B_{ij}-\lambda_i)^{k_{ij}-1}v_{ij}\ .$$

BEWEIS VON SATZ 14.A.: Da $p^e v = 0$ ($\forall\, v \in V$), gilt $m_v | p^e$, also $m_v = p^{a_v}$ mit $a_v \le e$. Wähle nun in V eine Basis $y_1, \ldots, y_n$. Offenbar ist $e = \max\{a_{y_1}, \ldots, a_{y_n}\}$, also nach eventueller Umnumerierung, $a_{y_1} = e$. Setze $y_1 = v_l$, $a_l = e$ und $\overline{V} = V/[v_l]$. Auf $\overline{V}$ induziert A den Endomorphismus $\overline{w} \mapsto \overline{Aw}$; wir können deshalb in Analogie zum zu $w \in V$ gebildeten $m_w \in K[x]$ für $\overline{w} \in \overline{V}$ das Polynom $m_{\overline{w}} \in K[x]$ definieren.

Die grundlegende Beobachtung ist nun die

$$(*) \qquad \text{Zu } \overline{w} \in \overline{V} \text{ existiert ein } w_1 \in V \text{ mit } \overline{w}_1 = \overline{w}\,,\ m_{w_1} = m_{\overline{w}}\,.$$

Weiß man $(*)$, so nützt man Induktion nach $\dim V$ für die Existenz von Vektoren $\overline{v}_1, \ldots, \overline{v}_{l-1} \in \overline{V}$ mit

$$\overline{V} = [\overline{v}_1] \oplus \ldots \oplus [\overline{v}_{l-1}]\,, \quad m_{\overline{v}_j} = p^{a_j}\,,\ a_1 \le \ldots \le a_{l-1}$$

aus. Sodann suchen wir Urbilder $v_j \in V$ mit $m_{v_j} = p^{a_j}$; $(*)$ garantiert solche v_j. Offenbar gilt

$$V = [v_1] + \ldots + [v_l]\,.$$

Die Summe ist tatsächlich direkt. Denn sind $g_j, g_l \in K[x]$ so, daß

$$\sum_{j=1}^{l-1} g_j v_j + g_l v_l = 0\,,$$

so ist $\sum\limits_{j=1}^{l-1} g_j\overline{v}_j = 0$, also $m_{\overline{v}_j}|g_j$, i.e. $g_j\overline{v}_j = 0\,,\ 1 \le j \le l-1$. Wegen $m_{\overline{v}_j} = m_{v_j}$ ist auch $g_jv_j = 0$ und es bleibt $g_lv_l = 0$.

Wir müssen noch $(*)$ nachweisen.

Wir starten mit irgendeinem Urbild $w \in V$ von $\overline{w}$. Setze

$$m_{v_l} = p^{a_l}\,,\ m_{\overline{w}} = p^{\overline{a}}\,,\ m_w = p^a\,.$$

Sicher gilt $\overline{a} \le a$ sowie $p^{\overline{a}}w \in [v_l]$, also $p^{\overline{a}}w = gv_l$ mit einem $g \in K[x]$.
Sollte $gv_l = 0$ sein, ist nichts mehr zu beweisen. Im anderen Fall schreibe $g \ne 0$ als Produkt

$$g = p^bg_1 \quad \text{mit} \quad g_1 \in K[x]\,,\ p \nmid g_1\,,\ b \ge 0\,.$$

Nun existieren $s, t \in K[x]$ mit $sp^{a_l} + tg_1 = 1$. Das zeigt $v_l = t(g_1v_l)$ und weiter

$$m_{v_l} = m_{g_1v_l} = p^{a_l}\,,\ m_{gv_l} = p^{a_l-b}\,.$$

Diese Information hat sofort $m_w = p^{\overline{a}+a_l-b}$ zur Folge. Die Maximalität von a_l erzwingt $a_l \ge \overline{a} + a_l - b$, mithin $b \ge \overline{a}$ und

$$p^{\overline{a}}w = gv_l = p^bg_1v_l = p^{\overline{a}}(p^{b-\overline{a}}g_1v_l)\,.$$

Das gesuchte w_1 ist nun $w_1 = w - p^{b-\overline{a}}g_1v_l\,.$

Die Eindeutigkeitsaussage folgt aus dem Vergleich von zwei Zerlegungen von V mit den davon induzierten Zerlegungen auf $V/\{v \in V : p(A)v = 0\}$ und auf $V/p(A)V$.

Die Jordan-Chevalley-Zerlegung einer zerfallenden Matrix:

DEFINITION 14.2. *$A \in K_{n\times n}$ heißt zerfallend, wenn alle Eigenwerte von A zu K gehören, wenn also das charakteristische Polynom $\chi_A(x)$ (oder, gleichbedeutend, das Minimalpolynoms $\mu_A(x)$) über K in Linearfaktoren zerfällt.*

Ist zum Beispiel $K = \mathbb{C}$, so zerfällt jedes A. Oder: Ist A nilpotent oder idempotent, so zerfällt A. Erinnerung: *Nilpotent* bedeutet die Existenz eines $k \in \mathbb{N}$ mit $A^k = 0$, *idempotent* heißt $A^2 = A$.

SATZ 14.D. *Zerfällt A, so existieren Matrizen D und N mit*

$$A = D + N\,, \quad DN = ND\,, \quad D \text{ ist diagonalisierbar}\,, \quad N \text{ ist nilpotent}\,.$$

Diese Zerlegung ist eindeutig.

Im Beweis nützt man zweierlei aus – erstens daß, falls $D \in K_{n\times n}$ vertauschbar mit der nilpotenten Matrix $N \in K_{n\times n}$ ist, also $DN = ND$ erfüllt, $\chi_D(x) = \chi_{D+N}(x)$ gilt, und zweitens, daß für $B \in GL_n(K)$ die inverse Matrix B^{-1} ein Polynom in B ist.

FOLGERUNG. *1. A und D haben dieselben Eigenwerte.*

2. Zerfällt A und ist A symmetrisch, so sind D und N symmetrisch.

3. *Ist $A \in \mathbb{R}_{n\times n}$ beliebig, so gilt die Behauptung des Satzes unabhängig davon, ob A zerfällt oder nicht, wenn man sich damit zufrieden gibt, daß das $D \in \mathbb{R}_{n\times n}$ über $\mathbb{C}$ diagonalisierbar ist (solche D heißen* halbeinfach*).*

15. Kapitel: Tensorprodukt von Vektorräumen

In diesem Kapitel, das eine überall in der Mathematik und Physik Anwendung findende Konstruktion von neuen Vektorräumen aus bekannten gegebenen vorstellt, sind V und W (nicht notwendig endlich dimensionale) Vektorräume über dem kommutativen Körper K.

DEFINITION 15.1. *Eine Abbildung $a : V \times W \to A$ der Produktmenge $V \times W$ in eine abelsche Gruppe A heißt ausgeglichen, falls*

1. *a linear in jedem Argument ist, also*
   $a(v_1 + v_2, w) = a(v_1, w) + a(v_2, w)$ *und* $a(v, w_1 + w_2) = a(v, w_1) + a(v, w_2)$
   für alle $v, v_1, v_2 \in V$ und alle $w, w_1, w_2 \in W$ gilt,
2. *a der Regel $a(\lambda v, w) = a(v, \lambda w)$ $(\forall\, v \in V, w \in W, \lambda \in K)$ genügt.*

DEFINITION 15.2. *Die abelsche Gruppe $V \otimes_K W$ heißt ein Tensorprodukt von V und W (über K), falls*

1. *es eine ausgeglichene Abbildung*
$$\otimes : V \times W \to V \otimes_K W\,,\ (v, w) \mapsto v \otimes w$$
   gibt
2. *$V \otimes W$ von den $v \otimes w$, $v \in V, w \in W$ erzeugt ist,*
3. *jede ausgeglichene Abbildung $a : V \times W \to A$ einen Gruppenhomomorphismus $V \otimes W \to A$, $v \otimes w \mapsto a(v, w)$ definiert.*

Aus der Definition 15.2 folgt sofort, daß zwei Tensorprodukte $V \otimes_K W$, $V \otimes'_K W$ derselben Vektorräume V, W kanonisch isomorph sind: $v \otimes w \leftrightarrow v \otimes' w$.

SATZ 15.A. *Zu V und W existiert ein Tensorprodukt.*

Wegen obiger Bemerkung sprechen wir ab jetzt von *dem* Tensorprodukt $V \otimes_K W$.

SATZ 15.B.

1. *$V \otimes_K W$ trägt auf natürliche Weise eine Vektorraumstruktur:*
$$\lambda(v \otimes w) \overset{\text{def}}{=} \lambda v \otimes w = v \otimes \lambda w\ .$$
2. *$V \otimes_K W \simeq W \otimes_K V$, $v \otimes w \leftrightarrow w \otimes v$*
3. *Ist Z ein dritter K-Vektorraum, so gilt*
$$V \otimes_K (W \otimes_K Z) \simeq (V \otimes_K W) \otimes Z\ ,\ v \otimes (w \otimes z) \leftrightarrow (v \otimes w) \otimes z$$

Satz 15.C. *1. Sind $v_1, \ldots, v_n$ linear unabhängige Vektoren in V und ist $\sum\limits_{i=1}^{n} v_i \otimes w_i = 0$ in $V \otimes_K W$, so sind alle $w_i = 0$ $(1 \leq i \leq n)$.*

2. Ist V endlich dimensional und $v_1, \ldots, v_n$ eine Basis, so läßt sich jedes Element des Vektorraums $V \otimes_K W$ eindeutig als $\sum\limits_{i=1}^{n} v_i \otimes w_i$ mit geeigneten Vektoren $w_i \in W$ schreiben.

3. Ist auch W endlich dimensional mit Basis $w_1, \ldots, w_m$, so ist $V \otimes_K W$ ein Vektorraum der Dimension $n \cdot m$ $(= \dim V \cdot \dim W)$ und $v_i \otimes w_j$ $(1 \leq i \leq n, 1 \leq j \leq m)$ eine Basis von $V \otimes_K W$.

Beispiele :

1. W sei ein K-Vektorraum und L Erweiterungskörper von K (etwa $K = \mathbb{Q}$, $L = \mathbb{R}$ oder $\mathbb{C}$; $K = \mathbb{R}$, $L = \mathbb{C}$). Dann ist L auch ein K-Vektorraum und $L \otimes_K W$ ist definiert. Dies wird durch $\lambda(\mu \otimes w) = \lambda\mu \otimes w$ ein L-Vektorraum. Es gilt $\dim_L(L \otimes W) = \dim_K W$ [6]. Des weiteren ist W über die Identifizierung $w \leftrightarrow 1 \otimes w$ ein K-Teilvektorraum von $L \otimes_K W$.

2. $\mathrm{Hom}_K(V, W) \simeq \hat{V} \otimes_K M$ (kanonisch), wobei $\hat{V} = \mathrm{Hom}_K(V, K)$ der Dualraum von V ist.

Satz 15.D. *1. $f : V \to V_1$, $g : W \to W_1$ seien K-lineare Abbildungen von Vektorräumen. Durch*

$$f \otimes g : V \otimes_K W \to V_1 \otimes_K W_1 \ , \ v \otimes w \to f(v) \otimes g(w)$$

ist eine wohldefinierte K-lineare Abbildung von $V \otimes_K W$ nach $V_1 \otimes_K W_1$ definiert.

2. Es seien $V = V_1$, $W = W_1$ endlich dimensional mit Basen $v_1, \ldots, v_n$ bzw. $w_1, \ldots, w_m$. Zu f gehöre bezüglich $\{v_i\}$ die Matrix $A = (\alpha_{ik})$, zu g bezüglich $\{w_j\}$ die Matrix $B = (\beta_{jl})$. Dann gehört zu $f \otimes g$ bezüglich $\{v_i \otimes w_j\}$ die Matrix $A \otimes B \in K_{nm \times nm}$ mit

$$\begin{pmatrix}
\alpha_{11}\beta_{11} & \ldots & \alpha_{11}\beta_{1m} & \alpha_{12}\beta_{11} & \ldots & \alpha_{12}\beta_{1m} & \ldots & \alpha_{1n}\beta_{11} & \ldots & \alpha_{1n}\beta_{1m} \\
\alpha_{11}\beta_{21} & \ldots & \alpha_{11}\beta_{2m} & \alpha_{12}\beta_{11} & \ldots & \alpha_{12}\beta_{1m} & \ldots & \alpha_{1n}\beta_{21} & \ldots & \alpha_{1n}\beta_{2m} \\
\vdots & \vdots & & & \vdots & & \vdots & & \vdots & \\
\alpha_{11}\beta_{m1} & \ldots & \alpha_{11}\beta_{mm} & \alpha_{12}\beta_{m1} & \ldots & \alpha_{12}\beta_{mm} & \ldots & \alpha_{1n}\beta_{m1} & \ldots & \alpha_{1n}\beta_{mm} \\
\vdots & \vdots & & & \vdots & & \vdots & & \vdots & \\
\vdots & \vdots & & & \vdots & & \vdots & & \vdots & \\
\alpha_{n1}\beta_{11} & \ldots & \alpha_{n1}\beta_{1m} & \alpha_{n2}\beta_{11} & \ldots & \alpha_{n2}\beta_{1m} & \ldots & \alpha_{nn}\beta_{11} & \ldots & \alpha_{nn}\beta_{1m} \\
\vdots & \vdots & & & \vdots & & \vdots & & \vdots & \\
\alpha_{n1}\beta_{m1} & \ldots & \alpha_{n1}\beta_{mm} & \alpha_{n2}\beta_{11} & \ldots & \alpha_{n2}\beta_{mm} & \ldots & \alpha_{nn}\beta_{m1} & \ldots & \alpha_{nn}\beta_{mm}
\end{pmatrix} .$$

Beobachtung : Spur $(A \otimes B) =$ Spur $A \cdot$ Spur B

Das alternierende Produkt V sei ein endlich dimensionaler K-Vektorraum und $k \in \mathbb{N}$.

Definition 15.3. *1.* $V^{\otimes k} = \underbrace{V \otimes \ldots \otimes V}_{k\,\mathrm{mal}}$

[6] Für $\dim_K L = \infty$ siehe den Anhang zu diesem Kapitel.

2. $\Lambda^k(V) = V^{\otimes k}/U_k$ *mit*

$$U_k = \langle v_1 \otimes \ldots \otimes v_k \ , \ \exists 1 \leq i \leq k-1 \text{ mit } v_i = v_{i+1} \rangle \ .$$

Das Bild von $v_1 \otimes \ldots \otimes v_k \in V \otimes \ldots \otimes V$ *in* $V^{\otimes k}/U_k$ *bezeichnen wir mit* $v_1 \wedge \ldots \wedge v_k$.

Beobachtungen :

1. Die Vertauschung zweier Vektoren in $v_1 \wedge \ldots \wedge v_k$ bewirkt ein Vorzeichen.
2. Sind zwei Vektoren in $v_1 \wedge \ldots \wedge v_k$ gleich, so ist $v_1 \wedge \ldots \wedge v_k = 0$.
3. Ist $e_1, \ldots, e_n$ Basis von V, so ist $e_{i_1} \wedge \ldots \wedge e_{i_k}$ mit $i_1 < \ldots < i_k$ Basis von $\Lambda^{(k)}(V)$. Insbesondere ist $\Lambda^1(V) = V$, $\Lambda^k(V) = 0$ für $k > n$ und $\dim \Lambda^n(V) = 1$.
4. Ist $f \in \mathrm{End}_K(V)$, so induziert f einen Endomorphismus von $\Lambda^{(k)}(V)$ und insbesondere einen Skalar für $k = n : \mathrm{End}(\Lambda^{(n)}(V)) = K$. Dieser Skalar ist $\det f$.

Satz 15.E. *1. Es sei* $V_1 \leq V$ *und* $V_2 = V/V_1$. *Des weiteren sei* $n_1 = \dim V_1$, $n_2 = \dim V_2$, $n = n_1 + n_2 = \dim V$. *Es gilt*

$$\Lambda^{n_1} V_1 \otimes_K \Lambda^{n_2} V_2 \simeq \Lambda^n V \ , \ (v_1 \wedge \ldots \wedge v_{n_1} \, , \, \overline{w}_1 \wedge \ldots \wedge \overline{w}_{n_2}) \mapsto v_1 \wedge \ldots \wedge v_{n_1} \wedge w_1 \wedge \ldots \wedge w_{n_2} \, .$$

2. Es sei $V = V_1 \oplus V_2$. *Dann gilt* $\bigoplus\limits_{r+s=m} \Lambda^r V_1 \otimes_K \Lambda^s V_2 = \Lambda^m V \ (\forall\, m \geq 0)$.

Bemerkung : $\Lambda V = \bigoplus\limits_{r \geq 0} \Lambda^r V$, mit $\Lambda^0 V \overset{\text{def}}{=} K$, ist eine K-Algebra, also ein K-Vektorraum mit einem zusätzlichem (assoziativen und distributiven) Produkt, für das die Elemente aus K selbst zentral sind. Das Produkt ist durch $\wedge$ definiert; es gilt $a \wedge b = (-1)^{rs} b \wedge a$ für $a \in \Lambda^r V$, $b \in \Lambda^s V$. Diese Algebra wird die Graßmann-Algebra zu V genannt.

Man kann, analog zur Graßmann-Algebra, die symmetrische Algebra $\mathfrak{S}(V)$ zu V definieren

$$\mathfrak{S}(V) = \bigoplus_{r \geq 0} \mathfrak{S}^r(V) \, , \ \mathfrak{S}^0(V) = K \, , \ \mathfrak{S}^r(V) = V^{\otimes r}/U_r \text{ für } r > 0 \, ,$$

$$U_r \overset{\text{def}}{=} \langle v_1 \otimes \ldots \otimes v_r - v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(r)} : \sigma \in S_r \rangle \, .$$

$\mathfrak{S}(V)$ ist eine kommutative K-Algebra mit dem durch $\otimes$ vermittelten Produkt. In der Tat ist $\mathfrak{S}(V)$ isomorph zum Polynomring $K[x_1, \ldots, x_n]$ in $n \, (= \dim V)$ kommutierenden unabhängigen Variablen $x_1, \ldots, x_n$ über K : Ist $e_1, \ldots, e_n$ eine Basis von V, so identifiziere

$$e_{i_1} \otimes \ldots \otimes e_{i_r} \bmod U_r \quad \text{mit} \quad x_{i_1} \cdots x_{i_r} \, .$$

Satz 15.F. *Sind* A *und* B *Algebren über* K, *so wird der Vektorraum* $A \otimes_K B$ *durch*

$$(a_1 \otimes b_1)(a_2 \otimes b_2) = a_1 a_2 \otimes b_1 b_2 \ , \ a_1, a_2 \in A \, , \, b_1, b_2 \in B$$

ebenfalls eine K*-Algebra.*

Anhang: Unendlich dimensionale Vektorräume

Es sei V ein unendlich dimensionaler Vektorraum über K (wie etwa $V = \mathbb{R}$ über $K = \mathbb{Q}$). Wir zeigen, daß V eine K-Basis besitzt, daß also Vektoren $e_\nu \in V$ mit ν aus einer (unendlichen) Indexmenge I existieren, die

1. *linear unabhängig* sind, d.h. per definitionem, daß je endlich viele der e_ν linear unabhängig sind

2. ganz V erzeugen, d.h.

$$\forall\, v \in V \exists\, \nu_1, \ldots, \nu_{n(v)} \in I \;\&\; \alpha_1, \ldots, \alpha_{n(v)} \in K : v = \sum_{i=1}^{n(v)} \alpha_i e_{\nu_i} \;.$$

Dazu benutzen wir ein Axiom, das sogenannte LEMMA VON ZORN, dessen Zusammenhang mit anderen Axiomen aus der Mengenlehre, wie etwa dem Auswahlaxiom, hier nicht diskutiert werden soll (vgl. dazu etwa H. Kneser, Math. Z. **53** (1950), 110).

Es sei M eine nichtleere Menge. M heißt teilweise geordnet, wenn es eine Relation $\leq$ zwischen gewissen (nicht notwendig allen) Elementen von M gibt, die folgendes erfüllt:

$$\begin{array}{l} m_1 \leq m_2 \;\&\; m_2 \leq m_1 \Longrightarrow m_1 = m_2 \\ m_1 \leq m_2 \;\&\; m_2 \leq m_3 \Longrightarrow m_1 \leq m_3 \end{array} .$$

(Ein Beispiel ist $M = \mathbb{N} \setminus \{3^j : j \geq 2\}$ mit $m_1 \leq m_2 \iff m_1 \mid m_2$.) Sie heißt vollständig geordnet, wenn noch für je zwei Elemente $m_1, m_2 \in M$ gilt: $m_1 \leq m_2$ oder $m_2 \leq m_1$. Ein maximales Element in der teilweise geordneten Menge M ist ein Element $m_0 \in M$ mit $[\,m_0 \leq m_1 \in M \Longrightarrow m_0 = m_1\,]$ (im Beispiel etwa $m_0 = 3$). Eine Kette in M ist eine in der Teilordnung von M vollständig geordnete nichtleere Teilmenge N von M (wie im obigen Beispiel die Teilmenge $N_2 = \{2^i : i = 0, 1, 2, \ldots\}$). Das Zornsche Lemma besagt nun: *Ist $M \neq \emptyset$ teilweise geordnet und besitzt jede Kette N in M eine obere Schranke in M (d.h. $\exists m \in M : n \leq m \; (\forall n \in N)$), so gibt es maximale Elemente in M.* (Hätten wir im Beispiel $\mathbb{N}$ statt $\mathbb{N} \setminus \{3^j : j \geq 2\}$ genommen, so gäbe es da kein einziges maximales Element. Allerdings genügt auch $\mathbb{N} \setminus \{3^j : j \geq 2\}$ nicht der Zornschen Bedingung: N_2 besitzt keine obere Schranke.)

Warnung: mit zu sorglosem Umgang mit dem Zornschen Lemma kann aller möglicher Unsinn erreicht werden. Wichtig ist, daß, erstens, M eine wohldefinierte Menge und, zweitens, M nichtleer ist. Zum Beispiel gibt es nicht die Menge aller Mengen, noch die Menge aller Erweiterungskörper von K, etc.

Die Standardanwendung betrifft eine Menge M, deren Elemente Teilmengen A einer vorgebenen Menge $\mathfrak{M}$ sind und deren Teilordnung durch $A_1 \leq A_2 \iff A_1 \subset A_2$ gegeben ist. Als obere Schranke einer Kette $N = \{A\}$ versucht man das Element $\bigcup_{A \in N} A$, das natürlich eine Teilmenge von $\mathfrak{M}$ ist, aber vielleicht nicht unbedingt in M liegt.

Für die Existenz der Basis von V betrachten wir nun die Menge M, deren Elemente genau die Teilmengen von V sind, die (im Sinne von 1. oben) aus linear unabhängigen Vektoren bestehen. Weil $V \neq 0$, existiert ein $v \neq 0$ in V und also ist $\{v\} \in M$, somit M nichtleer. Die Ordnung in M sei durch '$\subset$' gegeben. Man sieht sofort, daß Zorns Lemma anwendbar ist und maximale Elemente liefert. Jedes solche seinerseits besteht dann aus einer Basis von V.

Startet man von vorgebenen linear unabhängigen Vektoren $\{v_\nu\} \subset V$, so kann man für M auch die Teilmenge des obigen M nehmen, deren Elemente $\{v_\nu\}$ enthalten. Man folgert daus, daß jeder Satz linear unabhängiger Vektoren von V zu einer Basis von V ergänzt werden kann, und daraus wiederum, daß jeder Teilraum $U \leq V$ in V ein Komplement besitzt: $\exists\, U_1 \leq V : U \oplus U_1 = V$.

16. Kapitel: Lineare Ungleichungen

Hier ist die Aufgabenstellung des Kapitels.

Zu vorgegebenen reellen Parametern

$$\alpha_1, \ldots, \alpha_m\,;\ \beta_1, \ldots, \beta_n\,;\ \alpha_{ik}\,,\ 1 \leq i \leq n\,,\ 1 \leq k \leq m\,,$$

sind $x_k \in \mathbb{R}\,,\ 1 \leq k \leq m\,,$ mit

$$x_k \geq 0\,,\ \sum_{k=1}^{m} \alpha_{ik} x_k \leq \beta_i\ (1 \leq i \leq n)\,,\ \sum_{k=1}^{m} x_k \alpha_k \text{ maximal}$$

zu finden.

Solche Probleme tauchen etwa so auf. Eine Fabrik kauft n Rohstoffe $R_1, \ldots, R_n$ und fertigt daraus die m Güter $G_1, \ldots, G_m$. Zur Herstellung von G_k wird von R_i der Teil α_{ik} verbraucht. Stellt also die Fabrik das Endprodukt G_k genau x_k mal her, so wird von R_i insgesamt $\sum_{k=1}^{m} \alpha_{ik} x_k$ verbraucht. Die Gesamtmenge des eingekauften Rohstoffes R_i sei β_i; also $\sum_{k=1}^{m} \alpha_{ik} x_k \leq \beta_i$ für $1 \leq i \leq n$. Natürlich sind alle $x_k \geq 0$. Beim Verkauf von G_k mache die Fabrik den Gewinn α_k. Der Gesamtgewinn ist dann $\sum_{k=1}^{m} x_k \alpha_k$ – und der soll so groß wie möglich sein.

Übersetzung obigen Problems in die metrische Sprache

V sei der reelle m-dimensionale Vektorraum mit Orthonormalbasis $v_1, \ldots, v_m$ und W der reelle n-dimensionale Vektorraum mit Orthonormalbasis $w_1, \ldots, w_n$. Des weiteren seien

$$v_0 = \sum_{k=1}^{m} \alpha_k v_k \in V\ ,\ w_0 = \sum_{i=1}^{n} \beta_i w_i \in W$$

und f der bezüglich obiger Basen zur Matrix $(\alpha_{ik}) = A$ gehörige Homomorphismus von V nach W. Wir vereinbaren noch, daß i stets zwischen 1 und n, k zwischen 1 und m laufe.

Mit diesen Daten liest die Aufgabe sich nun so: finde ein $v \in V$ mit

$$(v, v_k) \geq 0\ ,\ (f(v), w_i) \leq (w_0, w_i)\ ,\ (v, v_0) \text{ maximal}.$$

Wir erklären eine dazu korrespondierende "duale" Aufgabe.

$$\text{finde } w \in W \quad \text{mit } (w, w_i) \geq 0 \ , \ (v_k, f^a(w)) \geq (v_k, v_0) \ , \ (w, w_0) \text{ minimal } ^7 .$$

Dabei ist $f^a : W \to V$ die zu f adjungierte Abbildung, $(f(v), w) = (v, f^a(w))$, gehört also zur Matrix A^T bezüglich der gewählten Basen.
Beachte $(\ker f)^\perp = \operatorname{im} f^a$.

DEFINITION 16.1. $$\begin{cases} L_V &=& \{v \in V : (v, v_k) \geq 0 \,, \ (f(v), w_i) \leq (w_0, w_i) \ (\forall\, i, k)\} \\ L_V^{\text{opt}} &=& \{v \in L_V : (v, v_0) \ \textit{maximal}\} \\ L_W &=& \{w \in W : (w, w_i) \geq 0 \,, \ (v_k, f^a(w)) \geq (v_k, v_0) \ (\forall\, i, k)\} \\ L_W^{\text{opt}} &=& \{w \in L_W : (w, w_0) \ \textit{minimal}\} \end{cases}$$

LEMMA 16.a. *$v \in L_V$, $w \in L_W \Longrightarrow (v, v_0) \leq (w, w_0)$. Gilt sogar Gleichheit, ist $v \in L_V^{\text{opt}}$ und $w \in L_W^{\text{opt}}$.*

Denn
$$\begin{aligned} (v, v_0) &= \textstyle\sum_k (v, v_k)(v_0, v_k) \leq \sum\limits_k (v, v_k)(v_k, f^a(w)) = (v, f^a(w)) \\ &= (f(v), w) = \sum\limits_i (f(v), w_i)(w, w_i) \leq \sum\limits_i (w_0, w_i)(w, w_i) = (w_0, w) \ . \end{aligned}$$

SATZ 16.A. *i) $L_V \neq \emptyset \neq L_W \Longrightarrow L_V^{\text{opt}} \neq \emptyset \neq L_W^{\text{opt}}$*
ii) $L_V = \emptyset \Longrightarrow L_W^{\text{opt}} = \emptyset \,, \quad L_W = \emptyset \Longrightarrow L_V^{\text{opt}} = \emptyset$.

Der Beweis sei für den Moment zurückgestellt.

DEFINITION 16.2. *Sind $u_1, \ldots, u_r$ Vektoren in W, so heißt die Menge*

$$\mathfrak{K}(u_1, \ldots, u_r) = \{\sum_{j=1}^{r} \lambda_j u_j \ , \ \lambda_j \geq 0\}$$

der von den $u_1, \ldots, u_r$ in W aufgespannte Kegel.

LEMMA 16.b. *1. Ist $y \in W$, aber $y \notin \mathfrak{K}(u_1, \ldots, u_r)$, so existiert eine Hyperebene $H_w = \langle w \rangle^\perp$, die y und $\mathfrak{K}(u_1, \ldots, u_r)$ trennt, d.h.*

$$(w, u) \geq 0 \text{ für } u \in \mathfrak{K}(u_1, \ldots, u_r) \,, \quad (w, y) = -1 \,.$$

2. Entweder hat $[\,(v, v_k) \geq 0 \,, (f(v), w_i) \leq (w_0, w_i)\,]$ eine Lösung in V, oder es ist $[\,(w, w_i) \geq 0 \ , \ (v_k, f^a(w)) \geq 0 \ , \ (w, w_0) < 0\,]$ lösbar in W.

[7] Solch ein Minimierungsproblem ist z.B. dieses: Die R_i sind Silos von Düngemittel. Über Rohre der Dicke α_{ik} sind sie mit Behältern G_k verbunden; allen aus dem R_i herausführenden Rohren zusammen ist ein regulierbares Ventil y_i vorgesetzt. Die Behälter G_k tropfen auf speziele Kulturen, jedoch nur dann in ausreichendem Maße, wenn sie bis zu einer Mindesthöhe α_k gefüllt sind. Der Fluß von R_i in G_k beträgt also $\sum_{i=1}^n \alpha_{ik} y_i \geq \alpha_k$. Der Dünger in R_i kostet β_i; natürlich sollen die Gesamtkosten $\sum_{i=1}^n \beta_i y_i$ minimal sein.

In der Tat, falls $y \notin U \stackrel{\text{def}}{=} \langle u_1, \ldots, u_r \rangle$ ist, existiert ein $w \in U^\perp$ mit $(w, y) = -1$. Ist allerdings $y \in U = \langle u_1, \ldots, u_r \rangle$, so schließen wir mit Induktion nach r wie folgt.

1. $r = 1:$ $\quad y = \mu u_1 \Longrightarrow \mu < 0$. Setze $w = -\frac{1}{\mu \cdot |u_1|^2} u_1$.

2. $r - 1 \to r:$ Nach Voraussetzung ist $y \notin \mathfrak{K}(u_2, \ldots, u_r) \subset \mathfrak{K}(u_1, \ldots, u_r) = \mathfrak{K}$. Somit existiert ein c mit $(c, y) = -1$, $(c, u_j) \geq 0$ $(2 \leq j \leq r)$ und wir dürfen $(c, u_1) < 0$ annehmen. Setze

   $$\tilde{u}_j = (c, u_j) u_1 - (c, u_1) u_j \text{ für } 2 \leq j \leq r \ , \ \tilde{y} = (c, y) u_1 - (c, u_1) y \, .$$

   Dann gilt: $\tilde{u}_j \in \mathfrak{K}$, aber $\tilde{y} \notin \tilde{\mathfrak{K}} \stackrel{\text{def}}{=} \mathfrak{K}(\tilde{u}_2, \ldots, \tilde{u}_r)$. Andernfalls wäre $\tilde{y} = \sum_{j \geq 2} \lambda_j \tilde{u}_j = (c, y) u_1 - (c, u_1) y$ mit $\lambda_j \geq 0$, also $y = \frac{-1}{(c, u_1)} \sum_j \lambda_j \tilde{u}_j + \frac{(c, y)}{(c, u_1)} u_1 \in \mathfrak{K}$, ein Widerspruch. Folglich gibt es ein d mit $(d, \tilde{u}_j) \geq 0$ für $2 \leq j \leq r$, $(d, \tilde{y}) = -1$, und $w = (d, u_1) c - (c, u_1) d$ leistet das Gewünschte.

Nun zu 2.: Beide Aussagen zugleich widersprechen sich, weil

$$(f(v), w) = \sum_i (f(v), w_i)(w, w_i) \leq \sum_i (w_0, w_i)(w, w_i) = (w_0, w) < 0 \ ,$$

$$(f(v), w) = (v, f^a(w)) = \sum_k (v, v_k)(f^a(w), v_k) \geq 0 \ .$$

Nun gelte die erste Aussage in 2. nicht. Wäre

$$w_0 = \sum_k \lambda_k f(v_k) + \sum_i \mu_i w_i \ , \ \lambda_k \geq 0 \ , \ \mu_i \geq 0 \ ,$$

so lieferte

$$(w_0, w_{i'}) = \sum_k \lambda_k (f(v_k), w_{i'}) + \underbrace{\sum_i \mu_i (w_i, w_{i'})}_{= \mu_{i'} \geq 0} \geq \sum_k \lambda_k (f(v_k), w_{i'}) = (f(\sum_k \lambda_k v_k), w_{i'})$$

ein $v = \sum\limits_k \lambda_k v_k \in L_V$ im Widerspruch zur Annahme $L_V = \emptyset$. Mithin

$$w_0 \notin \mathfrak{K}(f(v_1), \ldots, f(v_m), w_1, \ldots, w_n)$$

und es gibt eine trennende Hyperebene H_w

$$(w, w_0) = -1 \ , \ (v_k, f^a(w)) = (w, f(v_k)) \geq 0 \ , \ (w, w_i) \geq 0 \ .$$

Lemma 16.b ist bewiesen.

Wir kommen nun zum Beweis von Satz 16.A .

Wir vergrößern V, W zu $\tilde{V} := V \perp W$ bzw. $\tilde{W} := W \perp V \perp \mathbb{R}$. Als Orthonormalbasen von $\tilde{V}$ und $\tilde{W}$ wählen wir [8]

$$\begin{aligned} e_\lambda &= [v_1, 0], \ldots, [v_m, 0] \, , \, [0, w_1], \ldots, [0, w_n] \\ g_\mu &= [w_1, 0, 0], \ldots, [w_n, 0, 0] \, , \, [0, v_1, 0], \ldots, [0, v_m, 0] \, , \, [0, 0, 1] \, . \end{aligned}$$

[8] [,] statt (,), um Verwechslungen mit dem Skalarprodukt zu vermeiden

Schließlich sei $\tilde{f} : \tilde{V} \to \tilde{W}$ durch

$$[v, w] \mapsto [f(v), -f^a(w), -(v_0, v) + (w_0, w)]$$

definiert, so daß also $\tilde{f}^a$ den Effekt

$$[w, v, \alpha] \mapsto [f^a(w) - \alpha v_0, -f(v) + \alpha w_0]$$

hat. Setze noch $\tilde{w_0} \overset{\text{def}}{=} [w_0, -v_0, 0] \in \tilde{W}$.

Mit diesen Daten formulieren wir die Aufgabe

$$\text{suche } [v, w] \in \tilde{V} \text{ mit } ([v, w], e_\lambda) \geq 0 \ , \ (f([v, w]), g_\mu) \leq (\tilde{w_0}, g_\mu) \, .$$

Also

$$(v, v_k) \geq 0 \, , \, (w, w_i) \geq 0 \, , \, (f(v), w_i) \leq (w_0, w_i) \, , \, (f^a(w), v_k) \geq (v_0, v_k) \, , \, (w_0, w) \leq (v_0, v) \, .$$

Finden wir eine Lösung $[v, w]$, so ist nach Lemma 16.a automatisch $v \in L_V^{\text{opt}}$ und $w \in L_W^{\text{opt}}$. Finden wir aber keine Lösung, so gibt es nach Lemma 16.b.2 jedenfalls ein $[w, v, \alpha] \in \tilde{W}$ mit

$$([w, v, \alpha], g_\mu) \geq 0 \, , \, (\tilde{f}^a[w, v, \alpha], e_\lambda) \geq 0 \, , \, ([w, v, \alpha], \tilde{w_0}) < 0 \, ,$$

oder, ausgeschrieben,

$$(w, w_i) \geq 0 \ , \ (v, v_k) \geq 0 \ , \ \alpha \geq 0$$

$$(f^a(w), v_k) \geq \alpha(v_0, v_k) \ , \ (f(v), w_i) \leq \alpha(w_0, w_i)$$

$$(w, w_0) < (v, v_0) \ .$$

Wäre $\alpha > 0$, so wären $\frac{1}{\alpha} v \in L_V$ und $\frac{1}{\alpha} w \in L_W$ im Widerspruch zu $(w, w_0) < (v, v_0)$ und Lemma 16.a . Also ist $\alpha = 0$. Wähle nun, sofern möglich, $u \in L_V$, $z \in L_W$. Dann gilt

$$\begin{aligned} (w_0, w) &= \sum_i (w_0, w_i)(w, w_i) \geq \sum_i (f(u), w_i)(w, w_i) \\ &= (f(u), w) = (u, f^a(w)) = \sum_k (u, v_k)(f^a(w), v_k) \geq 0 \end{aligned}$$

und genauso

$$\begin{aligned} (v, v_0) &= \sum_k (v, v_k)(v_0, v_k) \leq \sum_k (v, v_k)(f^a(z), v_k) = (f^a(z), v) \\ &= (z, f^a(v)) = \sum_i (z, w_i)(w_i, f^a(v)) \leq 0 \ ; \end{aligned}$$

zusammengefaßt $(w_0, w) - (v, v_0) \geq 0$, ein Widerspruch zu $(w, w_0) < (v, v_0)$.

Was ist, wenn $L_V = \emptyset$? Aus Lemma 16.b.2 folgt dann die Existenz eines $w \in W$ mit

$$(w, w_i) \geq 0 \ , \ (v_k, f^a(w)) \geq 0 \ , \ (w, w_0) < 0 \ .$$

Wähle, falls möglich, $z \in L_W$. Also

$$(z, w_i) \geq 0 \ , \ (v_k, f^a(z)) \geq (v_k, v_0) \ .$$

Setze $y = z + \lambda w$ mit $\lambda \geq 0$. Offenbar ist $y \in L_W$. Berechne (w_0, y):

$$(w_0, y) = (w_0, z) + \lambda(w_0, w) \ .$$

Da $(w_0, w) < 0$, wird (w_0, y) beliebig klein, sofern nur λ groß genug gewählt wird. Also ist $L_W^{\text{opt}} = \emptyset$ und Satz 16.A bewiesen.

L_V geometrisch gesehen

Wir beginnen mit der Definition dreier geometrischer Begriffe.

Definition 16.3. *Eine Teilmenge $L \subset P$ des m-dimensionalen Euklidischen Raumes $A^m = (P, V)$ heißt konvex, falls mit je zwei Punkten $l_1, l_2 \in L$ die ganze Verbindungsstrecke zwischen l_1 und l_2 zu L gehört. Durch die Wahl eines Grundpunktes $o \in P$ identifizieren wir P mit V: $p \leftrightarrow op = v$. Dadurch wird L Teilmenge von V und die Konvexität bedeutet*

$$[\, x = \lambda_1 l_1 + \lambda_2 l_2 \ , \ \lambda_1, \lambda_2 \geq 0 \ , \ \lambda_1 + \lambda_2 = 1 \,] \Longrightarrow x \in L \ .$$

Ein verallgemeinertes Polyeder ist eine konvexe Menge L, die keine Halbgrade

$$\{p + \gamma v : \gamma \geq 0 \quad (p \text{ Punkt}, \ v \text{ Richtung})\}$$

enthält.

Ein Extremalpunkt der konvexen Menge L ist ein $l \in L$, das auf keiner echt in L verlaufenden Strecke liegt:

$$[\, l = \lambda_1 l_1 + \lambda_2 l_2 \ , \ \lambda_1, \lambda_2 > 0 \ , \ \lambda_1 + \lambda_2 = 1 \ , \ l_1, l_2 \in L \,] \Longrightarrow l = l_1 = l_2 \ .$$

Unser Ziel ist, L_V un L_V^{opt} zu bestimmen. Der folgende Trick [9] erleichtert dies.
Setze $V' = V \perp W$ und $f' : V' \to W$, $f'([v, w]) = f(v) + w$. Wir stellen uns die Aufgabe

$$\text{suche } [v, w] \in V' \text{ mit } f'([v, w]) = w_0 \, , \, ([v, w], e_\lambda) \geq 0 \, ;$$

hier ist $e_1 = [v_1, 0], \ldots, e_m = [v_m, 0], e_{m+1} = [0, w_1], \ldots, e_{m+n} = [0, w_n]$. Die Lösungsmenge heiße L'.

Beobachtungen

(1) $v \in L_V \Longrightarrow [v, w_0 - f(v)] \in L'$

(2) $[v, w] \in L' \Longrightarrow v \in L_V$

(3) $v \in L_V^{\text{opt}} \iff [v, w_0 - f(v)] \in L'$ & $(\, [v, w_0 - f(v)], [v_0, 0]\,)$ maximal

(4) L_V und L' sind konvex

(5) $[v, w]$ Extremalpunkt von $L' \Longrightarrow v$ Extremalpunkt von $L_V \Longrightarrow [v, w_0 - f(v)]$ Extremalpunkt von L'.

[9] in der Praxis ist es allerdings nicht gerade von Vorteil, die Dimension von V zu erhöhen

Im folgenden nennen wir V' einfach wieder V und f' wieder f; wir haben so obengenannte Aufgabe in die angenehmere Gestalt

$$\text{suche } v = \sum_k x_k v_k \in V \text{ mit } f(v) = w_0 \text{ und } x_k \geq 0$$

gebracht [10]. Die zugehörige Lösungsmenge heiße jetzt L; L^{opt} bezieht sich auf die Zusatzbedingung "(v, v_0)" ist maximal mit einem gegebenen Vektor $v_0 \in V$. Beachte, daß nun $m > n$ ist.

Ist $w_0 \notin \operatorname{im} f$, so ist $L = \emptyset$; ist $w_0 \in \operatorname{im} f$, so können wir ohne Schaden für unsere Fragestellung W durch $\operatorname{im} f$ ersetzen, also f als surjektiv annehmen (f' ist surjektiv). Dies tun wir im folgenden.

Weiter unten zeigen wir der Reihe nach

(1) die Extremalpunkte von L sind genau die $l \in L$ mit folgender Eigenschaft: Es gibt eine (von l abhängige) n-elementige Indexmenge $I = \{i_1, \ldots, i_n\} \subset \{1, \ldots, m\}$ mit

- 1a) $f(v_{i_1}), \ldots, f(v_{i_n})$ ist Basis von W
- 1b) $(l, v_{i_j}) = \{w_0, f(v_{i_j})\}$ für $j = 1, \ldots, n$ – mit der Notation $w_0 = \sum_{j=1}^{n} \{w_0, f(v_{i_j})\} f(v_{i_j})$
- 1c) $(l, v_k) = 0$ für $k \notin I$.

(2) $L \neq \emptyset \implies$ es gibt Extremalpunkte
$L^{\text{opt}} \neq \emptyset \implies$ es gibt einen Extremalpunkt von L, der in L^{opt} liegt.

(3) $L \neq \emptyset$, L verallgemeinerter Polyeder $\implies$ es gibt einen Extremalpunkt, der in L^{opt} liegt.

Wie ist das anzuwenden? Zunächst gibt es nach (1) nur endlich viele Extremalpunkte und diese lassen sich alle so finden: Suche die möglichen Mengen I mit 1a), setze sodann l durch 1b) und 1c) fest und überprüfe schließlich, ob wirklich $l \in L$, also, da sicher $f(l) = w_0$, ob $\{w_0, f(v_{i_j})\} \geq 0$ ist. Nachdem alle Extremalpunkte gefunden sind, wähle man den mit dem "größten Winkel" zu v_0: dieser ist eine optimale Lösung – jedenfalls wenn L ein verallgemeinerter Polyeder ist.

Beweis von (1): Tatsächlich zeigen wir, daß die Extremalpunkte von L genau die $l \in L$ sind, für die

$$\Big\{ f(v_j) : j \in I_l \overset{\text{def}}{=} \{1 \leq k \leq m : (l, v_k) \neq 0\} \Big\}$$

nur aus linear unabhängigen Vektoren besteht.

Es erfülle $l \in L$ die Bedingungen 1a), 1b) und 1c) und liege auf der Strecke zwischen $l_1, l_2 \in L$:

$$l = \lambda_1 l_1 + \lambda_2 l_2 \ , \ \lambda_1, \lambda_2 > 0 \ , \ \lambda_1 + \lambda_2 = 1 \ .$$

Aus $(l, v_k) = 0$ für $k \notin I$ folgt $\lambda_1(l_1, v_k) + \lambda_2(l_2, v_k) = 0$, und deshalb $(l_1, v_k) = 0 = (l_2, v_k)$. Des weiteren

$$f(l_1) = w_0 = \sum_{j=1}^{n} \{w_0, f(v_{i_j})\} f(v_{i_j}), .$$

[10] Denn mit $p : V' \to V$, $[v, w] \mapsto v$, und $i : V \to V'$, $v \mapsto [v, w_0 - f(v)]$, gilt: $pi = \text{id}$, $p(L') = L_V$, $i(L_V) \subset L'$, p und i sind mit den Begriffen *extremal* und *optimal* verträglich.

Mithin $(l_1, v_{i_j}) = \{w_0, f(v_{i_j})\} = (l_2, v_{i_j})$. Daher ist $l_1 = l_2$ und also auch $= l$, i.e. l ist Extremalpunkt.

Umgekehrt, sei nun l ein Punkt aus L. Wir definieren wie oben die Indexmenge („Träger von l")

$$I_l = \{j : 1 \leq j \leq m \ , \ (l, v_j) > 0\}$$

und unterscheiden zwei Fälle:

a) Die $f(v_j)$ sind linear unabhängig für $j \in I_l$.
Weil f surjektiv ist, kann I_l zu einer Menge I mit 1a) vergrößert werden; automatisch ist dann 1c) richtig, weil immer $(l, v_k) \geq 0$. Wegen

$$f(l) = w_0 \ , \ \text{also} \ \sum_{j=1}^{n} (l, v_{i_j}) f(v_{i_j}) = \sum_{j=1}^{n} \{w_0, f(v_{i_j})\} f(v_{i_j})$$

(denn l hat Träger in I) ist auch 1b) richtig; l ist also Extremalpunkt.

b) Die $f(v_j)$ sind linear abhängig.
Wir wählen eine nichttriviale Relation $\sum_{j \in I_l} \lambda_j f(v_j) = 0$. Der Vektor $z = \sum_{j \in I_l} \lambda_j v_j$ liegt also im Kern von f und ist $\neq 0$. Es sei $\mu = \min \{|\lambda_j|^{-1}(l, v_j) : \lambda_j \neq 0\}$. Also

$$\mu > 0 \text{ und } (l, v_j) \geq \mu |\lambda_j| \text{ falls } j \in I_l \ .$$

Damit sind mit $l - \mu z$ und $l + \mu z$ zwei Punkte in L gefunden, die l als Mittelpunkt auf ihrer Verbindungsstrecke haben. Ist l aber Extremalpunkt, so ist das ein Widerspruch. Der Fall b) tritt also nicht ein, wenn l ein Extremalpunkt ist.

Beweis von (2): Sei l eine Lösung. Ist $l = 0$, so ist wegen $x_k \geq 0$ für $v = \sum_k x_k v_k \in L$ offenbar l ein Extremalpunkt. Es sei deshalb $l \neq 0$ und I_l die Trägermenge von l. Gilt a) aus dem vorigen Beweis, so ist l schon Extremalpunkt, wie dort gezeigt wurde. Es gelte also b), so daß wir zu den beiden zusätzlichen Lösungen $l - \mu z$ und $l + \mu z$ kommen (beachte: l ist dann kein Extremalpunkt). Sollte sogar $l \in L^{\text{opt}}$ sein, so gehören diese beiden neuen Lösungen auch zu L^{opt}, weil $(z, v_0) = 0$, da

$$(z, v_0) < 0 \implies (l - \mu z, v_0) = (l, v_0) - \mu(z, v_0) > (l, v_0)$$

$$(z, v_0) > 0 \implies (l + \mu z, v_0) = (l, v_0) + \mu(z, v_0) > (l, v_0) \ .$$

Das Minimum $\mu = \min \{|\lambda_j|^{-1}(l, v_j) : \lambda_j \neq 0\}$ werde beim Index j_0 angenommen, so daß also

$$(l \pm \mu z, v_{j_0}) = (l, v_{j_0}) \pm \mu(z, v_{j_0}) = |\lambda_{j_0}|\mu \pm \mu\lambda_{j_0}$$

für eines der beiden Vorzeichen $= 0$ ist. Damit haben wir eine neue Lösung $l_1 \in L$ gefunden (bzw. $\in L^{\text{opt}}$), deren Träger kleiner als der von l ist: der Index j_0 fehlt. Wir wiederholen das Verfahren.

Beweis von (3): Dies ist Geometrie: Ein Polyeder [11] ist die konvexe Hülle seiner Extremalpunkte; d.h. in unserem Fall

[11] d.i. ein verallgemeinerter Polyeder mit endlich vielen Extremalpunkten

$$(*) \qquad L = \{\textstyle\sum_{\nu=1}^t \lambda_\nu l_\nu : \lambda_\nu \geq 0 \ , \ \sum_{\nu=1}^t \lambda_\nu = 1\} \ ,$$

wenn $l_1, \ldots, l_t$ die sämtlichen Extremalpunkte von L sind. Also

$$(\sum_{\nu=1}^t \lambda_\nu l_\nu, v_0) = \sum_{\nu=1}^t \lambda_\nu (l_\nu, v_0) \leq \sum_{\nu=1}^t \lambda_\nu (l_{\nu_0}, v_0) = (l_{\nu_0}, v_0) \ ;$$

dabei ist ν_0 so ausgewählt, daß

$$(l_{\nu_0}, v_0) = \max\{(l_\nu, v_0) \ : \ 1 \leq \nu \leq t\} \text{ ist } .$$

Beweis von $(*)$: Falls $w_0 = 0$, ist $\{0\} = L$, da mit $l \neq 0$, auch $\gamma l \in L$ für $\gamma \geq 0$. Es sei also $w_0 \neq 0$, $l \in L$ und, wie früher schon, I_l der Träger von l. Der Beweis von $(*)$ geschieht durch Induktion nach $|I_l|$. Wegen $l \neq 0$, ist im Falle $|I_l| = 1$ der Punkt l automatisch ein Extremalpunkt, vgl. den Beweis von (1), Fall a). Im Induktionsschritt dürfen wir wegen des genannten Beweisschrittes jetzt annehmen, daß die $f(v_j)$ für $j \in I_l$ linear abhängig sind. Wir arbeiten wieder mit dem Vektor z von früher. Seine Koordinaten λ_j können nicht alle ≤ 0 sein, weil sonst die Halbgerade $\{l - \gamma z : \gamma \geq 0\}$ in L läge. Also existiert

$$\overline{\mu} = \min \ \{\lambda_j^{-1}(l, v_j) \ : \ \lambda_j > 0\} \ ,$$

und es ist $\overline{l} := l - \overline{\mu} z \in L$. Offenbar gilt noch

$$I_{\overline{l}} \subsetneq I_l \ , \ \text{also } \overline{l} \in \{\sum_{\nu=1}^t \lambda_\nu l_\nu\} \ .$$

Wir suchen einen zweiten Punkt $\overline{\overline{l}} \in \{\sum_{\nu=1}^t \lambda_\nu l_\nu\}$ mit $l = \overline{\lambda} \ \overline{l} + \overline{\overline{\lambda}} \ \overline{\overline{l}}$, $\overline{\lambda}, \overline{\overline{\lambda}} \geq 0$, $\overline{\lambda} + \overline{\overline{\lambda}} = 1$.

Setze dazu $\mu_1 = \max\{(\overline{l}, v_j)(l, v_j)^{-1} \ : \ j \in I_l\}$.

Es ist $\mu_1 > 1$, denn sonst wäre $(\overline{l}, v_j) \leq (l, v_j)$ und es müßten noch wegen $(\overline{l}, v_j) = (l, v_j) - \overline{\mu}\lambda_j$ alle $\lambda_j \geq 0$ sein. Folglich wäre die Halbgerade $\{l + \gamma z : \gamma \geq 0\}$ in L. Daher

$$\overline{\overline{l}} = (\mu_1 - 1)^{-1}(\mu_1 l - \overline{l}) \in L$$

und $I_{\overline{\overline{l}}} \subsetneq I_l$ (es fehlt der Index, bei dem μ_1 erreicht wird). Außerdem ist $l = \mu_1^{-1}\overline{l} + (1 - \mu_1^{-1})\overline{\overline{l}}$.

Wir haben $L \subset \{\sum \lambda_\nu l_\nu\}$ gezeigt; die umgekehrte Inklusion ist trivial. $(*)$ ist damit bewiesen.

Bemerkung: Algorithmen zur Bestimmung von L^{opt} werden in den Vorlesungen zur Optimierungstheorie vorgestellt. Hier ist ein "theoretisches" Verfahren.

Da $f : V \to W$ surjektiv ist, gilt $W = \langle f(v_1), \ldots, f(v_m)\rangle$ (mit $m = \dim V$). Wähle aus diesen W aufspannenden Vektoren eine Basis von W, etwa $f(v_{i_1}), \ldots, f(v_{i_n})$, und schreibe w_0 als Linearkombination $w_0 = \sum_{j=1}^n x_j f(v_{i_j})$. Sind alle $x_j \geq 0$, so ist $l = \sum_{j=1}^n x_j v_j$ ein Extremalpunkt in L. Ist einmal ein $x_j < 0$, so vergiß die Basis $f(v_{i_1}), \ldots, f(v_{i_n})$ und wähle eine neue. Sofern $L \neq \emptyset$ und L keine Halbgerade enthält, gibt es aber Basen mit nichtnegativen zugehörigen x-Koordinaten von w_0. Liste alle diese auf und erhalte so alle Extremalpunkte in L. Der mit dem größten "Winkel" (l, v_0) ist dann optimal.

17. Kapitel: Ein wenig aus der projektiven Geometrie zum Schluß

V sei ein endlich dimensionaler Vektorraum über dem kommutativen Körper K; $\dim V = n \geq 1$.

Definition. *1. $p(V)$ sei die Menge der eindimensionalen (linearen) Unterräme von V.*

2. Der projektive Raum $P(V)$ über V sei die Menge aller (linearen) Unterräume von V. Sind $U_1, U_2 \leq V$, so definiere $U_1 \preceq U_2$ in $P(V)$ durch $U_1 \subset U_2$.

3. Für $U \leq V$ setze $\dim_p(U) = \dim U - 1$. "Punkte" von $P(V)$ sind die U der Dimension 1, "Geraden" die U der Dimension 2.

Beobachtung: Über $\preceq$ wird $P(V)$ zu einer teilgeorneten Menge, in der zu je zwei Elementen U_1, U_2 das Maximum $U_1 + U_2 \leq V$ und das Minimum $U_1 \cap U_2 \leq V$ existiert. Diese Teilordnung fassen wir auch als eine Inzidenzstruktur auf $P(V)$ auf. Der Teilraum $p(V)$ von $P(V)$ erzeugt $P(V)$ über die Maximums- und Minimumsbildung.

Wir führen Koordinaten auf $p(V)$ ein und wählen dazu eine Basis $v_1, \ldots, v_n$ in V. Ein eindimensionaler Unterraum U von V ist dann durch ein n-Tupel $\underline{\alpha} \overset{\text{def}}{=} (\alpha_1, \ldots, \alpha_n) \neq (0, \ldots, 0)$ angebbar: $U = \langle \sum_{i=1}^n \alpha_i v_i \rangle$. Zwei solche Tupel $\underline{\alpha}$ und $\underline{\beta}$ gehören genau dann zum selben U, wenn es ein $0 \neq \lambda \in K$ mit $\lambda \alpha_i = \beta_i$ $(1 \leq i \leq n)$ gibt. Dem "Punkt" U in $p(V)$ ordnen wir nun das Tupel $[\alpha_1, \ldots, \alpha_n]$ zu und identifizieren $[\alpha_1, \ldots, \alpha_n]$ und $[\beta_1, \ldots, \beta_n]$, wenn $\lambda \alpha_i = \beta_i$ mit einem (von i unabhängigen) $\lambda \in K^\times \overset{\text{def}}{=} K \setminus \{0\}$ gilt. Die α_i heißen die homogenen Koordinaten vom Punkt U; man beachte, daß sie nur bis auf einen gemeinsamen Faktor $\neq 0$ festgelegt sind und daß $[0, \ldots, 0]$ nicht erlaubt ist.

So wie wir den n-dimensionalen Vektorraum V zum affinen Raum $\mathbb{A}^n$ ausbauen, können wir aus $p(V)$ einen projektiven Raum $\mathbb{P}^{n-1}$ gewinnen. Diesen identifiziere am besten mit $P(V)$, indem einer Punktmenge $\{[x_{j1}, \ldots, x_{jn}]\}$ die Summe (das Maximum) der Unterräume $\langle (x_{j1}, \ldots, x_{jn}) \rangle$ zugeordnet wird (hier ist j ein laufender Index).

Es sei zum Beispiel $n = 3$; wir sprechen also über die projektive Ebene $\mathbb{P}^2 = P(V)$ und insbesondere hier über (projektive) Punkte und (projektive) Geraden. Offenbar gilt

1. durch zwei verschiedene Punkte geht genau eine Gerade: $[\alpha_1, \alpha_2, \alpha_3]$ und $[\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3]$ liefern nämlich die linear unabhängigen Vektoren $v = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3$, $\tilde{v} = \tilde{\alpha}_1 v_1 + \tilde{\alpha}_2 v_2 + \tilde{\alpha}_3 v_3$, die zusammen einen 2-dimensionalen Teilraum von V aufspannen, also eine Gerade in $\mathbb{P}^2 = P(V)$, und dies ist die Gerade durch die beiden Punkte im Sinne unserer Inzidenzstruktur

2. zwei verschiedene Geraden in $\mathbb{P}^2 = P(V)$ entsprechen zwei verschiedenen 2-dimensionalen Unterräumen U_1, U_2 in V. Ihr Schnitt $U_1 \cap U_2$ hat die Dimension 1 (weil $\dim V = 3$), liefert also einen Punkt in $\mathbb{P}^2$.

Vielleicht ist folgendes Bild für eine geometrische Vorstellung des $\mathbb{P}^{n-1} = P(V)$ hilfreich: Betrachte den $(n-1)$-dimensionalen affinen Raum $\mathbb{A}^{n-1} \overset{\text{def}}{=} \mathbb{A}^{n-1}(K)$ als Teilraum des $\mathbb{A}^n$. Wähle dazu Grundpunkte $o_{n-1} \in \mathbb{A}^{n-1}$ und $o_n \in \mathbb{A}^n \setminus \mathbb{A}^{n-1}$ und setze $H = \{v \in V : o_{n-1}v \in \mathbb{A}^{n-1}\}$ [12], so daß also H eine Hyperebene in V ist.
Die *eigentlichen* Punkte p des $\mathbb{P}^{n-1}$ sind nun genau die Punkte des $\mathbb{A}^{n-1}$ und wir interpretieren sie als die Geraden durch p und o_n, somit als die linearen Unterräumen $\langle v \rangle$ von V mit $v \notin H$.

[12] einer einfachen Notation wegen identifizieren wir hier die affinen Räume mit ihren Punktmengen und schreiben ov für die Anwendung des Vektors v auf den Punkt o

Die anderen, sogenannten *uneigentlichen*, Punkte des $\mathbb{P}^{n-1}$ sind die zu $\mathbb{A}^{n-1}$ parallelen Geraden des $\mathbb{A}^n$ durch o_n, also die Geraden $\{o_n(\lambda h_0) : \lambda \in K\}$, welche wir mit den linearen Unterräumen $\langle h_0 \rangle$ von $H \leq V$ identifizieren (h_0 durchläuft $H \setminus \{0\}$).
In Koordinatenschreibweise bekommen dann die eigentlichen Punkte die Koordinaten

$$[\alpha_1, \ldots, \alpha_{n-1}, 1] \leftrightarrow (\alpha_1, \ldots, \alpha_{n-1})$$

und die uneigentlichen die Koordinaten

$$[\alpha_1, \ldots, \alpha_{n-1}, 0] \quad (\text{mit } H = \langle v_1, \ldots, v_{n-1} \rangle\,,\ v_1, \ldots, v_n \text{ Basis von } V)\,.$$

Um Unabhängigkeit von Basiswahlen zu erreichen, hatten wir früher lineare Abbildungen zwischen Vektorräumen studieren müssen. Wir verallgemeinern ins Projektive.

Definition. *V_1, V_2 seien K-Vektorräume der Dimensionen n_1 bzw. n_2. Die Gruppe aller Körperautomorphismen von K werde mit $\mathrm{Aut}(K)$ bezeichnet* [13].

1. *Zu $\sigma \in \mathrm{Aut}(K)$ werde eine σ-lineare Abbildung $f : V_1 \to V_2$ durch die Regeln*

$$f(v + \tilde{v}) = f(v) + f(\tilde{v})\,,\ f(\lambda v) = \sigma(\lambda) f(v) \quad (\forall\, v, \tilde{v} \in V_1\,,\ \lambda \in K)$$

definiert.

2. *Für $n_1 = n_2$ heiße $\gamma : P(V_1) \to P(V_2)$ eine Kollineation, wenn γ eine mit $\preceq$ verträgliche Bijektion ist.*

Auch die σ-linearen Abbildungen entsprechen Matrizen $A \in K_{n_2 \times n_1}$; sind nämlich $e_1, \ldots, e_{n_1}$ und $w_1, \ldots, w_{n_2}$ Basen von V_1 bzw. V_2, so erhalte A wie früher aus $f(e_i) = \sum_{j=1}^{n2} \alpha_{ji} w_j$ $(1 \leq i \leq n_1)$, i.e. $A = (\alpha_{ji})$, aber $v = \sum_i \beta_i e_i$ hat nun das Bild $f(v) = \sum_j \alpha_{ji} \sigma(\beta_i) w_j$, der Spaltenvektor $(\beta_1, \ldots, \beta_{n_1})^T$ geht also in $A \cdot (\sigma(\beta_1), \ldots, \sigma(\beta_{n_1}))^T$ über.

Bezüglich Kollinetionen beobachtet man:

1. σ-lineare bijektive Abbildungen $f : V_1 \to V_2$ induzieren über

$$\gamma(U_1) \overset{\text{def}}{=} f(U_1) \quad \text{für} \quad U_1 \leq V_1$$

   Kollineationen $\gamma : P(V_1) \to P(V_2)$

2. $\dim U_1 = \dim \gamma(U_1)$ für $U_1 \leq V_1$

3. mit γ ist auch $\gamma^{-1} : P(V_2) \to P(V_1)$ eine Kollineation

4. $\gamma(U \cap \tilde{U}) = \gamma(U) \cap \gamma(\tilde{U})\,,\ \gamma(U + \tilde{U}) = \gamma(U) + \gamma(\tilde{U})$ für $U, \tilde{U} \leq V_1$

5. $U \subset U' + \tilde{U} \Longrightarrow \gamma(U) \subset \gamma(U') + \gamma(\tilde{U})$ für eindimensionale Unterräume $U, U', \tilde{U} \leq V_1$

[13] z.B. $\mathrm{Aut}(\mathbb{Q}) = 1$; $\mathrm{Aut}(\mathbb{R}) = 1$ (weil $\sigma \in \mathrm{Aut}(\mathbb{R})$ wegen $[\alpha > 0 \iff \alpha = \beta^2]$ mit $\geq$ verträglich ist); $1 \neq \sigma = [x + iy \mapsto x - iy] \in \mathrm{Aut}(\mathbb{C})$

Im Fall $V_1 = V_2 = V$ ist $GL(V)$ (oder $GL_n(K)$) die Gruppe aller K-linearen Bijektionen $f : V \to V$. Wenn hingegen σ nicht länger trivial ist und die Gruppe $\mathrm{Aut}(K)$ durchläuft, erhält man die größere Gruppe $A(V)$ aller σ-linearen Bijektionen $V \to V$ (mit der Hintereinanderausführung solcher als Produkt). Es bezeichne noch $PA(V)$ die Gruppe aller Kollineationen $P(V) \to P(V)$. Wir haben folgendes Diagramm

$$\begin{array}{ccccc} K^\times & \rightarrowtail & GL(V) & \twoheadrightarrow & PGL(V) \\ & & \downarrow & & \\ & & A(V) & \to & PA(V) \\ & & \downarrow & & \\ & & \mathrm{Aut}(K) & & \end{array} \quad ,$$

in dem "$\rightarrowtail$" injektiv, "$\twoheadrightarrow$" surjektiv bedeutet und jede Zeile oder Spalte des Typs $G' \overset{\varphi'}{\rightarrowtail} G \overset{\varphi''}{\twoheadrightarrow} G''$ erfüllt: $\mathrm{im}\,\varphi' = \ker\varphi'' = \{x \in G : \varphi''(x) = 1\}$. Solche kurzen Gruppensequenzen $G' \rightarrowtail G \twoheadrightarrow G''$ nennen wir fortan *exakt.*

$PGL(V)$ ist definiert als $GL(V)/K^\times \simeq GL_n(K)\Big/\left\{\begin{pmatrix} \lambda & & \\ & \ddots & \\ & & \lambda \end{pmatrix} : \lambda \in K^\times\right\}$ [14]; sie ist die *projektive lineare Gruppe* von V.

SATZ 17.1. *1. Im Falle $n_1 = n_2 \geq 3$ ist eine bijektive Abbildung $\hat\gamma : p(V_1) \to p(V_2)$ mit*

$$U \subset U' + \tilde U \Longrightarrow \hat\gamma(U) \subset \hat\gamma(U') + \hat\gamma(\tilde U)$$

für eindimensionale Unterräume $U, U', \tilde U \leq V_1$ zu einer Kollineation $\gamma : P(V_1) \to P(V_2)$ fortsetzbar; genauer

$$\exists^!\, \sigma \in \mathrm{Aut}(K)\, \exists\, f : V_1 \overset{\sigma-\text{linear}}{\longrightarrow} V_2 : \quad \hat\gamma(U_1) = \gamma(U_1) = f(U_1) \quad \text{für } U_1 \leq V_1\,,\ \dim U_1 = 1\,.$$

f ist bis auf Multiplikation mit einem Skalar $\lambda \in K^\times$ eindeutig bestimmt.

2. Ist $n \geq 3$, so ist die Abbildung $A(V) \to PA(V)$ surjektiv und liefert das kommutative Diagramm [15] mit exakten Zeilen un Spalten

$$\begin{array}{ccccc} K^\times & \rightarrowtail & GL(V) & \twoheadrightarrow & PGL(V) \\ \| & & \downarrow & & \downarrow \\ K^\times & \rightarrowtail & A(V) & \twoheadrightarrow & PA(V) \\ & & \downarrow & & \downarrow \\ & & \mathrm{Aut}(K) & = & \mathrm{Aut}(K) \end{array} \quad .$$

[14] die Notation ist weiter unten erklärt

[15] "kommutativ" bedeutet, daß der gewählte Weg für die Hintereinanderausführung zweier Abbildungen des Diagramms keinen Einfluß auf das Ergebnis hat

Bemerkungen:

1. *Doppelverhältnis*

   Auf $\mathbb{P}^1(K) \subset P(V)$ wählen wir vier verschiedene Punkte $P_i = [\alpha_{i1}, \alpha_{i2}]$, $1 \leq i \leq 4$. Deren Doppelverhältnis ist durch

   $$\frac{(\alpha_{11}\alpha_{42} - \alpha_{41}\alpha_{12})(\alpha_{31}\alpha_{22} - \alpha_{21}\alpha_{32})}{(\alpha_{31}\alpha_{42} - \alpha_{41}\alpha_{32})(\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12})}$$

   definiert; das ist eine wohlbestimmte Zahl aus K. Man zeigt, daß es invariant bei Anwendung von projektiven linearen Abbildungen $\gamma \in PGL(V)$ ist; eine Kollineation $\gamma \in PA(V)$ verschiebt es um den durch γ induzierten Automorphismus $\sigma \in \mathrm{Aut}(K)$. – Dies ist, so man will, eine Verallgemeinerung des sogenannten Strahlensatzes im $\mathbb{A}^n$: zeichne eine Gerade durch einen Punkt o und wähle auf ihr zwei weitere von o verschiedene Punkte $p_1 \neq p_2$. Eine bijektive lineare Abbildung $f : V^n \to V^n$ (oder eine längen- und winkeltreue Abbildung $f : \mathbb{A}^n \to \mathbb{A}^n$ mit Fixpunkt o) liefert über $f(op_i) = oq_i$ die neue Gerade durch o, q_1, q_2, und der Strahlensatz betrifft dann die jeweilig entstehenden Verhältnisse, wie etwa $\frac{|op_1|}{|op_2|} = \frac{|oq_1|}{|oq_2|}$.

2. *Topologie*

   Interpretiere die Punkte des $\mathbb{P}^{n-1}(K)$, so wie früher beschrieben, als eigentliche Punkte $[\alpha_1, \ldots, \alpha_{n-1}, 1]$ und als uneigentliche Punkte $[\alpha_1, \ldots, \alpha_{n-1}, 0]$. Das liefert n Einbettungen $\mathbb{A}^{n-1} \rightarrowtail \mathbb{P}^{n-1}$, indem die "1 nach vorn wandert". Sei $K = \mathbb{R}$. Der Euklidische Raum $\mathbb{A}^{n-1}(\mathbb{R})$ trägt eine natürliche Topologie, die über die n genannten Einbettungen widerspruchsfrei auf den $\mathbb{P}^{n-1}(\mathbb{R})$ übertragen werden kann. Man zeigt, daß damit der $\mathbb{P}^{n-1}(\mathbb{R})$ ein kompakter Raum wird. Im Gegensatz zu der für analytische Zwecke günstigen Einpunktkompaktifizierung des lokalkompakten $\mathbb{A}^{n-1}(\mathbb{R})$ (Hinzunahme des Punktes ∞) erhalten wir mit dem $\mathbb{P}^{n-1}(\mathbb{R})$ eine mehr geometrische Kompaktifizierung.

3. *Gruppen*

   Ein generelles mathematisches Prinzip ist, Objekte über ihre Automorphismen zu studieren, so z.B. Vektorräume V über die lineare Gruppe $GL(V)$ (das ist klassische lineare Algebra). In diesem Sinn ist der projektive Raum $P(V)$ über die Gruppe $PA(V)$ zu studieren und das Diagramm in Satz 17.1 gibt über $PA(V)$ gute Auskunft. (Andereseits ist es vorteilhaft, explizite Beispiele von Gruppen zur Verfügung zu haben, und die Geometrie trägt wesentlich dazu bei, an solche Beispiele heranzukommen – siehe $GL(V), PGL(V), \ldots$.)

4. *Hilfsmittel aus der Gruppentheorie*

   G sei eine Gruppe und U eine Untergruppe von G. Definiere mithilfe von U eine neue Gleichheit auf G über

   $$x_1 \equiv x_2 \bmod U \iff x_1 x_2^{-1} \in U\,.$$

   Dies liefert eine reflexive, symmetrische und transitive Äquivalenzrelation auf G; die Menge der Äquivalenzklassen bezeichne mit $\overline{G}$, also

   $$\overline{G} = \{\overline{x} : x \in G\,;\ \overline{x}_1 = \overline{x}_2 \iff x_1 x_2^{-1} \in U\}\,.$$

Gilt $gU = Ug$, i.e., $gug^{-1} \in U$ ($\forall g \in G, \forall u \in U$) [16], so heißt die Untergruppe U ein Normalteiler von G, und statt $\overline{G}$ schreiben wir dann auch G/U.

G/U wird über $\overline{x} \cdot \overline{y} = \overline{xy}$ eine Gruppe (die Faktorgruppe von G modulo U), und $\varphi_U : G \to G/U$, $x \mapsto \overline{x}$, ist dann ein Epimorphismus von Gruppen, den wir die kanonische Restklassenabbildung modulo U nennen.

Beispiele:

- $G = GL(V)$; $U = K^\times$ ist darin ein Normalteiler mit Faktorgruppe $PGL(V)$
- in $PA(V)$ ist $PGL(V)$ ein Normalteiler mit Faktorgruppe $\mathrm{Aut}(K)$
- $G = S_3$; hier ist $U = \langle(123)\rangle$ ein Normalteiler (mit Faktorgruppe $\{\pm 1\}$); $U = \langle(12)\rangle$ ist kein Normalteiler
- ist $\varphi : G \to G'$ ein Gruppenhomomorphismus, so ist $B = \mathrm{im}\,\varphi$ eine Untergruppe von G' und $U = \ker\varphi = \{x \in G : \varphi(x) = 1\}$ ein Normalteiler in G. Obiges φ_U hat ebenfalls $\ker\varphi_U = U$. Mit $\varphi : G \to G'$ und φ_U erhalten wir: $\exists^1 \varphi' : G/U \to G'$, ein Gruppenhomomorphismus, mit

$$\begin{tikzcd} G \arrow[rr, "\varphi"] \arrow[dr, "\varphi_U"'] & & G' \\ & G/U \arrow[ur, "\varphi'"'] & \end{tikzcd}$$ kommutiert.

Geometrische Deutung; Perspektivitäten:

Es seien V_1, V_2 Unterräume gleicher Dimension $n \geq 2$ eines K-Vektorraums W. Dann existiert ein gemeinsames Komplement C in W und wir wählen eines:

$$W = V_1 \oplus C = V_2 \oplus C \ .$$

Für $U_1 \leq V_1$ und $C \leq \tilde{C} \leq W$ resultiert

$$U_1 = \tilde{C} \cap V_1 \iff \tilde{C} = U_1 + C \ ,$$

und folglich erhalten wir eine von C abhängige Korrespondenz zwischen den $U_1 \leq V_1$ und den $U_2 \leq V_2$ (nämlich durch $U_1 + C = U_2 + C$) und zugleich eine Kollineation

$$\gamma : P(V_1) \to P(V_2)\,, \quad U_1 \mapsto U_2\,,$$

die *die Projektion - oder Perspktivität - von* $P(V_1)$ *auf* $P(V_2)$ *mit Zentrum* $P(C)$ heißt.

Satz 17.2. *1. Diese Perspektivitäten sind von K-linearen Bijektionen $f : V_1 \to V_2$ mit $f(v) = v$ für $v \in V_1 \cap V_2$ induziert; umgekehrt, jedes solche f bestimmt eine Perspektivität (mit von f abhängigem Zentrum).*

2. $P(V_1)$ und $P(V_2)$ seien echte Unterräume eines projektiven Raumes $P(W)$; es gelte des weiteren $\dim V_1 = \dim V_2 \geq 2$. *Ist dann $\gamma : P(V_1) \to P(V_2)$ eine Kollineation, die alle Punkte in $P(V_1) \cap P(V_2)$ festläßt und gilt darüber hinaus* $\dim_p((P(V_1) \cap P(V_2)) \geq 1$, *so ist γ eine Perspektivität.*

[16] das ist gewiß der Fall, wenn G abelsch (kommutativ) ist