Lineare Algebra II

Prof. Dr. Uwe Jannsen   Sommersemester 2006

§1 Transformation auf Dreiecksgestalt

Sei K ein Körper.

Definition 1.1 Zwei Matrizen A und $A' \in M_n(K)$ heißen **ähnlich** (oder **konjugiert**), wenn es eine invertierbare Matrix $B \in M_n(K)$ gibt (also $B \in GL_n(K)$) mit

$$A' = B^{-1}AB\,.$$

Bemerkungen 1.2 (a) Nach I 15.1 und I 17.15 haben ähnliche Matrizen dieselbe Determinante und dasselbe charakteristische Polynom, also auch dieselbe Spur.

(b) Stellt man einen Endomorphismus durch zwei verschiedene Basen dar, so erhält man konjugierte Matrizen, nach I 10.27 und I 10.29. Weiter kann man in 1.1 B als die Transformationsmatrix M_b^e auffassen, wobei e die Standardbasis des K^n ist und b aus den Spalten von B besteht. Dann ist A' die lineare Abbildung, die $x \mapsto Ax$ bezüglich der Basis b darstellt.

Wir wollen untersuchen, wann eine Matrix **diagonalisierbar** ist, d.h., ähnlich zu einer Diagonalmatrix

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

ist. Nach Hauptachsentransformation/Spektralsatz ist dies der Fall für reelle symmetrische oder komplexe hermitesche Matrizen. Aber nicht jede Matrix ist diagonalisierbar – z.B. nicht die Matrix

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in M_2(\mathbb{R})$$

(siehe Beispiel 1.6) oder die Matrix

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in M_2(\mathbb{C})$$

(siehe Beispiel 3.7 (b)).

Wir untersuchen in diesem Abschnitt zuerst, welche Matrizen sich auf obere Dreieicksgestalt (vergl. I 13.15)

$$\begin{pmatrix} * & & & \\ & \ddots & * & \\ & 0 & \ddots & \\ & & & * \end{pmatrix}$$

oder untere Dreiecksgestalt

$$\begin{pmatrix} * & & & \\ & \ddots & 0 & \\ & * & \ddots & \\ & & & * \end{pmatrix}$$

transformieren lassen.

Satz 1.3 (a) Ein Endomorphismus

$$\varphi : V \to V$$

eines endlich-dimensionalen K-Vektorraums V läßt sich genau dann durch eine obere Dreiecksmatrix darstellen (d.h., es gibt eine Basis b von V, so dass $M_b^b(\varphi)$ obere Dreiecksgestalt hat), wenn das charakteristische Polynom **über** K zerfällt, d.h., Produkt von Linearfaktoren in $K[x]$ ist.

(b) Eine Matrix $A \in M_n(K)$ ist genau dann ähnlich zu einer oberen Dreiecksmatrix, wenn $\chi_A(x)$ in $K[x]$ zerfällt.

Dasselbe gilt für untere Dreiecksmatrizen.

Beweis Es genügt, (a) zu zeigen: Zwei Matrizen sind genau dann ähnlich, wenn sie durch einen Basiswechsel, d.h., Übergang zu einer anderen Basis auseinander hervorgehen (siehe Bemerkung 1.2(b)).

(a): Die eine Richtung folgt aus:

Lemma 1.4 Sei $A \in M_n(K)$ ein obere oder untere Dreiecksmatrix mit Diagonalelementen $a_{ii} \quad (i = 1, \dots, n)$. Dann ist

$$\chi_A(x) = \prod_{i=1}^{n} (x - a_{ii}) \, .$$

Insbesondere sind die a_{ii} die Eigenwerte von A.

Beweis $xE - A$ ist eine obere oder untere Dreiecksmatrix mit Diagonalelementen $x - a_{ii} \quad (i = 1, \dots, n)$. Daher folgt die Aussage aus Satz I 13.15, bzw. analog dazu nach Bemerkung I 17.9.

Die andere Richtung von 1.3 (a) folgt mit vollständiger Induktion nach $n = \dim V$: Der Fall $n = 1$ ist trivial, also sei $n > 1$. Zerfällt $\chi_\varphi(x)$ in $K[x]$, so hat φ einen Eigenwert $\lambda \in K$. Sei $v \neq 0$ ein Eigenvektor zu λ und sei

$$b = (b_1 = v, b_2, \dots, b_n)$$

eine Basis von V. Die Matrix-Darstellung von φ bezüglich b ist dann von der Gestalt

$$A = M_b^b(\varphi) = \left(\begin{array}{c|ccc} \lambda & a_{12} & \dots & a_{1n} \\ \hline 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{array}\right)$$

mit einer gewissen Matrix $A' \in M_{n-1}(K)$. Durch Entwicklung nach der ersten Spalte folgt:

$$\chi_\varphi(x) = \det(xE_n - A) = (x-\lambda)\cdot\det(xE_{n-1} - A') = (x-\lambda)\chi_{A'}(x)$$

($E_m \in M_m(K)$ die $(m \times m)$-Einheitsmatrix).

Andererseits ist nach I 17.11 (λ ist Nullstelle von $\chi_A(x)$) und I 16.10 (dann ist $\chi_A(x)$ durch $(x-\lambda)$ teilbar)

$$\chi_A(x) = (x-\lambda)\cdot g(x)\,,$$

mit einem Polynom $g(x) \in K[x]$, und $g(x)$ zerfällt nach Voraussetzung in Linearfaktoren. Wegen der Eindeutigkeit der Polynomdivision folgt

$$\chi_{A'}(x) = g(x)\,;$$

dies zerfällt also in Linearfaktoren in $K[x]$. Wir können A' als Endomorphismus von $V_2 =< b_2, \ldots, b_n >_K$ auffassen, mit $A'(b_j) = \sum\limits_{i=2}^{n} a_{ij}b_i \quad (j = 2, \ldots, n)$. Nach Induktionsvoraussetzung gibt es dann eine neue Basis $b'_2, \ldots, b'_n$ von V_2 bezüglich derer A' obere Dreiecksgestalt hat. Dann hat A in der Basis $b' = (b_1, b'_2, \ldots, b'_n)$ obere Dreiecksgestalt

$$\left(\begin{array}{c|ccccc}
\lambda & a_{12} & \ldots & \ldots & \ldots & a_{1n} \\
\hline
0 & * & & & & \\
\vdots & & \ddots & & * & \\
\vdots & & & \ddots & & \\
\vdots & & 0 & & \ddots & \\
0 & & & & & *
\end{array}\right)$$

Der Fall unterer Dreiecksmatrizen ergibt sich durch Betrachtung der transponierten Matrix.

Da über $\mathbb{C}$ jedes Polynom in Linearfaktoren zerfällt (I 16.15), erhalten wir:

Corollar 1.5 Über $\mathbb{C}$ ist jede Matrix trigonalisierbar, d.h., ähnlich zu einer (oberen) Dreiecksmatrix (d.h., läßt sich durch Basistransformation auf solche Gestalt bringen).

Beispiel 1.6 Für $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ ist das chrakteristische Polynom $\chi_A(x) = x^2 + 1$. Dies hat keine reelle Nullstelle; daher ist A nicht über $\mathbb{R}$ trigonalisierbar. Nach 1.5 ist A aber über $\mathbb{C}$ trigonalisierbar. Tatsächlich ist A über $\mathbb{C}$ sogar diagonalisierbar (siehe 3.7 (a)).

§2 Eigenräume

Sei K ein Körper.

Erinnerung 2.1 (siehe I.17.4) Für einen K-Vektorraum V, einen Endomorphismus $\varphi : V \to V$ und ein $\lambda \in K$ heißt

$$V(\lambda) := \ker(\varphi - \lambda \cdot \ \mathrm{id}) \subseteq V$$

der Eigenraum von φ zu λ. Um die Abhängigkeit von φ auszudrücken, schreiben wir auch $V(\varphi, \lambda)$.

Bemerkung 2.2 (a) Manche Bücher verwenden die Bezeichnung V_λ.

(b) Es gilt also:
$V(\lambda) \neq 0 \quad \Leftrightarrow \quad \lambda$ ist Eigenwert von φ
$V(\lambda) \smallsetminus \{0\}$ ist die Menge der Eigenvektoren zu λ.

(c) Für eine Matrix $A \in M_n(K)$ und $\lambda \in K$ ist entsprechend

$$V(\lambda) \; := \; V(A, \lambda) \; := \; \ker(A - \lambda \cdot E)\,.$$

der Eigenraum der linearen Abbildung $A : K^n \to K^n$ (Spezialfall $V = K^n$).

Definition 2.3 Sei $\varphi : V \to V$ ein Endomorphismus eines endlich-dimensionalen K-Vektorraums und sei $\lambda \in K$.

(a) Die Dimension $\dim_K V(\lambda)$ des Eigenraums von λ heißt die **geometrische Vielfachheit** von λ als Eigenwert von φ.

(b) Die **algebraische Vielfachheit** von λ als Eigenwert von φ ist die Vielfachheit von λ als Nullstelle des charakteristischen Polynoms $\chi_\varphi(x)$ von φ.

Hierbei definieren wir

Definition 2.4 Sei $f(x) \in K[x]$ ein Polynom über K. Die Vielfachheit von $\lambda \in K$ als Nullstelle von $f(x)$ ist gleich $m \quad (m \in \mathbb{N}_0)$, wenn gilt

$$f(x) = (x - \lambda)^m \cdot g(x)$$

mit $g(\lambda) \neq 0$. Wir nennen λ dann auch eine m-fache (einfache, zweifache...) Nullstelle von $f(x)$.

Bemerkung 2.5 (a) Die Nullstellen-Vielfachheit von $\lambda \in K$ für $f \in K[x]$ ist $\leq \deg(f)$, und 0 genau dann, wenn λ keine Nullstelle von f ist.

(b) $\lambda \in K$ ist kein Eigenwert von $\varphi : V \to V$
$\Leftrightarrow$ die geometrische Vielfachheit von λ ist 0
$\Leftrightarrow$ die algebraische Vielfachheit von λ ist 0.

Beispiel 2.6 Betrachte die Matrix

$$A = \begin{pmatrix} -1 & 0 & 0 & 6 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Die Eigenwerte von A sind 1 und -1; das chrakteristische Polynom ist

$$\chi_A(x) = (x+1)(x-1)^3\,.$$

Die algebraische Vielfachheit ist also 1 für $\lambda = -1$ und 3 für 1.

Was sind die geometrischen Vielfachheiten?

$\boldsymbol{\lambda = -1}$: Betrachte

$$V(-1) = \ker(A - (-1) \cdot E) = \ker \begin{pmatrix} 0 & 0 & 0 & 6 \\ 0 & 2 & 0 & -3 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

Der Rang dieser Matrix ist 3: die letzten 3 Spaltenvektoren sind linear unabhängig, denn es sind schon die 3 Vektoren $\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} \begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix}$ linear unabhängig (die Matrix mit diesen Spalten hat Det $= 8 \neq 0$). Nach der Rangformel ist also $\dim V(-1) = 4 - 3 = 1$.

$\boldsymbol{\lambda = 1}$:

$$V(1) = \ker(A - E) = \ker \begin{pmatrix} -2 & 0 & 0 & 6 \\ 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Diese Matrix hat Rang 2: der 1. und 4. Spaltenvektor sind linear unabhängig; die mittleren Spaltenvektoren sind null. Also ist $\dim V(1) = 2$, die geometrische Vielfachheit von 1 ist also $2 < 3 =$ algebraische Vielfachheit. Diese Vielfachheiten können also verschieden sein.

Allgemein gilt aber:

Proposition 2.7 Die geometrische Vielfachheit eines Eigenwertes ist kleiner oder gleich der algebraischen Vielfachheit.

Beweis Sei $\dim V(\lambda) = m$ und $v_1, \ldots, v_m$ eine Basis von $V(\lambda)$. Ergänze dies zu einer Basis $b = (v_1, \ldots, v_m, v_{m+1}, \ldots, v_n)$ von V. Die Matrixdarstellung von φ bezüglich dieser Basis hat dann wegen $\varphi(v_i) = \lambda \cdot v_i$ für $i = 1, \ldots, m$ die Gestalt

$$A = M_b^b(\varphi) = \begin{array}{c} \\ m\left\{ \begin{array}{c} \\ \\ \\ \end{array} \right. \\ \\ \\ \end{array} \left(\begin{array}{c|c} \overbrace{\begin{matrix} \lambda & & \\ & \ddots & 0 \\ 0 & \ddots & \\ & & \lambda \end{matrix}}^{m} & * \\ \hline 0 & A' \end{array} \right)$$

Durch iterierte Entwicklung nach den ersten m Spalten folgt

$$\chi_A(x) = (x - \lambda)^m \cdot \chi_{A'}(x)$$

(exakter Beweis durch vollständige Induktion nach m). Damit ist die algebraische Vielfachheit m' von λ größer oder gleich m: ist $\chi_{A'}(x) = (x-\lambda)^r \cdot g(x)$ mit $r \geq 0$ und $g(\lambda) \neq 0$, so ist $m' = m + r \geq m=$ geometrische Vielfachheit von λ.

Lemma 2.8 (vergleiche auch I 19.22) Sind $\lambda_1, \ldots, \lambda_s$ verschiedene Eigenwerte von φ, so bilden $V(\lambda_1), \ldots, V(\lambda_s)$ eine direkte Summe in V, d.h., es gilt:

(2.8.1) Ist $v_1 + \ldots + v_s = 0$ mit $v_i \in V(\lambda_i) \quad (i = 1, \ldots, s)$, so folgt $v_1 = \ldots = v_s = 0$.

Insbesondere gilt $\sum\limits_{i=1}^{s} \dim V(\lambda_i) \leq \dim V$, und es ist $V(\lambda_i) \cap V(\lambda_j) = 0$ für $i \neq j$.

Beweis durch Induktion über s. Für $s = 1$ ist nicht zu zeigen. Sei $s > 1$ und

$$(*) \qquad v_1 + v_2 + \ldots + v_s = 0 \text{ mit } v_i \in V(\lambda_i)\,.$$

Durch Anwenden von φ erhalten wir die Gleichung

$$\lambda_1 v_1 + \ldots + \lambda_s v_s = 0$$

Durch Multiplizieren von $(*)$ mit λ_s und Subtraktion folgt

$$(\lambda_1 - \lambda_s)v_1 + (\lambda_2 - \lambda_s)v_2 + \ldots + (\lambda_{s-1} - \lambda_s)v_{s-1} = 0$$

Wegen $(\lambda_i - \lambda_s)v_i \in V(\lambda_i)$ und (2.8.1) für $s - 1$ (Induktionsvoraussetzung) schließen wir

$$(\lambda_i - \lambda_s)v_i = 0 \quad \forall\, i = 1, \ldots, s-1$$

Wegen $\lambda_i \neq \lambda_s$ für $i \neq s$ folgt hieraus

$$v_i = 0 \quad \forall\, i = 1, \ldots, s-1\,.$$

Wegen $(*)$ gilt dann auch $v_s = 0$.

Die anderen Aussagen folgen sofort.

§3 Diagonalisierbarkeit

Die Gleichheit von geometrischer und algebraischer Vielfachheit gilt gerade bei den **diagonalisierbaren** Matrizen:

Definition 3.1 Ein Endomorphismus $\varphi : V \to V$ eines endlich-dimensionalen K-Vektorraums heißt **diagonalisierbar**, wenn es eine Basis b von V gibt, bezüglich derer die Matrixdarstellung von φ Diagonalgestalt hat, d.h., für die

$$M_b^b(\varphi) = \begin{pmatrix} \lambda_1 & & & & \\ & \lambda_2 & & 0 & \\ & & \ddots & & \\ & 0 & & \ddots & \\ & & & & \lambda_n \end{pmatrix}$$

mit $\lambda_1, \dots, \lambda_n \in K$.

Bemerkung 3.2 Damit ist auch definiert, wann eine $n \times n$-Matrix $A \in M_n(K)$ diagonalisierbar ist: wenn sie **ähnlich** zu einer Diagonalmatrix

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

ist.

Denn die Beschreibung des Endomorphismus $A : K^n \to K^n, v \mapsto Av$, in einer neuen Basis $b = (b_1, \dots, b_n)$ bedeutet grade Übergang zu $A' = B^{-1}AB$, wobei $B = M_b^e (:= M_b^e(\mathrm{id}))$ die Matrix ist, bei der in der j-ten Spalte gerade der j-te Basisvektor b_j steht, also $B = (b_1 \mid \dots \mid b_n) =$ Matrix mit Spaltenvektoren $b_1, \dots, b_n$ (vergleiche 1.2(b)).

Satz 3.3 Für einen Endomorphismus $\varphi : V \to V$ eines endlich-dimensionalen K-Vektorraums sind die folgenden Aussagen äquivalent:

(a) φ ist diagonalisierbar.

(b) φ ist trigonalisierbar, und für jeden Eigenwert λ von φ ist die geometrische Vielfachheit gleich der algebraischen Vielfachheit.

(c) Das charakteristische Polynom $\chi_\varphi(x)$ von φ zerfällt in $K[x]$, und für jeden Eigenwert λ von φ ist die geometrische gleich der algebraischen Vielfachheit.

(d) V besitzt eine Basis aus Eigenvektoren von φ

(e) Sind $\lambda_1, \dots, \lambda_r \in K$ die verschiedenen Eigenwerte von φ, so ist $\dim V = \sum\limits_{i=1}^{r} \dim V(\lambda_i)$.

(f) Sind $\lambda_1, \dots, \lambda_v \in K$ die verschiedenen Eigenwerte von φ, so ist $V = \bigoplus\limits_{i=1}^{r} V(\lambda_i)$.

Wir zeigen zuerst:

Lemma 3.4 Sei $b = (b_1, \dots, b_n)$ eine Basis von V und $\varphi : V \to V$ ein Endomorphismus. Dann gilt: Die Matrixdarstellung $M_b^b(\varphi)$ von φ bezüglich b ist genau dann eine Diagonalmatrix, wenn alle b_i Eigenvektoren von φ sind.

Beweis: Dies ist klar, denn es gilt:

$$\varphi(b_i) = \lambda_i \cdot b_i \quad \forall\, i = 1, \dots, n \quad \Leftrightarrow \quad M_b^b(\varphi) = (\delta_{ij} \cdot \lambda_i) = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & 0 & \\ & 0 & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

Beweis von 3.3 (a) $\Leftrightarrow$ (d) folgt aus Lemma 3.4 und (b) $\Leftrightarrow$ (c) folgt aus Satz 1.3.

Seien nun $\lambda_1, \ldots, \lambda_r$ die verschiedenen Eigenwerte von φ in K. Nach Lemma 2.8 bilden die $V(\lambda_i)$ eine direkte Summe

$$V' = \bigoplus_{i=1}^{r} V(\lambda_i) \subseteq V ,$$

und es ist $\dim V' = \sum_{i=1}^{r} \dim V(\lambda_i) \leq \dim(V)$. Dies zeigt, dass (e) $\Leftrightarrow$ (f), denn es ist $V' = V$ genau dann wenn $\dim V' = \dim V$.

(c) $\Rightarrow$ (e): Zerfällt $\chi_\varphi(x)$, so ist

$$(*) \qquad \chi_\varphi(x) = \prod_{i=1}^{r} (x - \lambda_i)^{m_i} \text{ mit } \lambda_i \in K ,$$

wobei $m_i =$ algebraische Vielfachheit von λ_i. Ist $m_i = \dim V(\lambda_i)$, so folgt

$$\sum_{i=1}^{r} \dim V(\lambda_i) = \sum_{i=1}^{r} m_i \stackrel{(*)}{=} \deg \chi_\varphi(x) = \dim V$$

(f) $\Rightarrow$ (d): Wähle für jedes $i = 1, \ldots, r$ eine Basis $b^i = (b_1^i, \ldots, b_{n_i}^i), n_i = \dim V(\lambda_i)$. Dann ist $b = (b_1^i, \ldots, b_{n_1}^i, \ldots, b_1^r, \ldots, b_{n_r}^r)$ eine Basis von V'. Ist also $V' = V$, so gilt (d).

(d) $\Rightarrow$ (b): Sei b eine Basis aus Eigenvektoren. Wir können b so anordnen, dass

$$b = (b_1^1, \ldots, b_{n_1}^1, b_1^2, \ldots, b_{n_2}^2, \ldots, b_1^t, \ldots, b_{n_t}^t)$$

wobei $b_1^i, \ldots, b_{n_i}^i$ gerade Eigenvektoren zum Eigenwert λ_i sind, und $\lambda_i \neq \lambda_j$ für $i \neq j$. Dann ist die Matrix von φ bezüglich b gleich

$$A = \begin{array}{c} n_1\left\{\begin{array}{c} \\ \\ \\ \end{array}\right. \\ n_2\left\{\begin{array}{c} \\ \\ \\ \end{array}\right. \\ \\ n_t\left\{\begin{array}{c} \\ \\ \\ \end{array}\right. \end{array} \begin{pmatrix} \lambda_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & \lambda_1 & & & & & & \\ & & & \lambda_2 & & & & & \\ & & & & \ddots & & & & \\ & & & & & \lambda_2 & & & \\ & & & & & & \ddots & & \\ & & & & & & & \lambda_t & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & \lambda_t \end{pmatrix} ,$$

also

$$\chi_\varphi(x) = \chi_A(x) = \prod_{i=1}^{n} (x - \lambda_i)^{n_i} .$$

$\lambda_1, \ldots, \lambda_t$ sind also alle Eigenwerte von φ, und n_i ist die algebraische Vielfachheit von λ_i. Andererseits ist für alle i

$$< b_1^i, \ldots, b_{n_i}^i >_K \subseteq V(\lambda_i)$$

und damit $n_i \leq \dim V(\lambda_i)$. Mit 2.7 folgt $n_i = \dim V(\lambda_i)$ für $i = 1, \ldots, t$ und wir erhalten (b). Damit ist die Äquivalenz aller Aussagen in 3.1 gezeigt.

Corollar 3.5 Sei V ein K-Vektorraum der Dimension $n \in \mathbb{N}$, und sei $\varphi : V \to V$ ein Endomorphismus. Hat φ n verschiedene Eigenwerte in K, so ist φ diagonalisierbar.

Beweis Seien $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von φ. Dann ist $1 \leq \dim V(\lambda_i)$ für alle $i = 1, \ldots, n$ und

$$\sum_{i=1}^{n} \dim V(\lambda_i) \leq n$$

wegen Lemma 2.8. Also gilt Gleichheit überall und damit $\bigoplus\limits_{i=1}^{n} V(\lambda_i) = V$, d.h., V besitzt eine Basis aus Eigenvektoren für φ.

Bemerkung 3.6 Sei $A \in M_n(K)$. Ist A diagonalisierbar, so sei $v_1, \ldots, v_n$ eine Basis von K^n zu den Eigenvektoren $\lambda_1, \ldots, \lambda_n$ (die nicht notwendig verschieden sind). Sei $T = (v_1 \mid \ldots \mid v_n)$ die Matrix, deren Spalten die Eigenvektoren $v_1, \ldots, v_n$ sind. Dann ist

$$T^{-1}AT = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & 0 & \\ & 0 & \ddots & \\ & & & \lambda_n \end{pmatrix} \tag{3.6.1}$$

eine Diagonalisierung von A. Denn es ist

$$\begin{aligned} T^{-1}AT &= T^{-1}A(v_1 \mid \ldots \mid v_n) = T^{-1}(Av_1 \mid \ldots \mid Av_n) \\ &= T^{-1}(\lambda_1 v_1 \mid \ldots \mid \lambda_n v_n) \quad (v_i \text{ Eigenvektor zum Eigenwert } \lambda_i) \\ &= (\lambda_1 e_1 \mid \ldots \mid \lambda_n e_n) = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & 0 & \\ & 0 & \ddots & \\ & & & \lambda_n \end{pmatrix}, \end{aligned}$$

denn wegen $T^{-1}T = E$ muss gelten $(T^{-1}v_1 \mid \ldots \mid T^{-1}v_n) = E = (e_1 \mid \ldots \mid e_n)$, also $T^{-1}v_i = e_i$, damit $T^{-1}\lambda_i v_i = \lambda_i T^{-1} v_i = \lambda_i e_i$.

Die Matrix aus den Eigenvektoren liefert also eine Transformationsmatrix zur Diagonalisierung.

Ein anderer Beweis von (3.6.1) ergibt sich daraus, dass $T^{-1}AT$ die Darstellungsmatrix von A (d.h., der linearen Abbildung $A : K^n \to K^n$) bezüglich der Basis $v = (v_1, \ldots, v_n)$ ist (da $T = M_v^e$, für $e = (e_1, \ldots, e_n)$ die Standardbasis von K^n. Nun verwende (den Beweis von) Lemma 2.3.

Beispiele 3.7 (a) Wir betrachten die Matrix

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

aus Beispiel 1.6. Das charakteristische Polynom ist $\chi_A(x) = x^2 + 1$ und hat die komplexen Nullstellen i und $-i$ $\quad (i = \sqrt{-1})$. Nach 3.5 ist A also über $\mathbb{C}$ diagonalisierbar.

Diagonalisierung: Eigenvektor zu i ist $\begin{pmatrix}1\\i\end{pmatrix}$, und Eigenvektor zu $-i$ ist $\begin{pmatrix}i\\1\end{pmatrix}$. Die Transformationsmatrix T besteht aus den Eigenvektoren

$$T = \begin{pmatrix}1 & i\\ i & 1\end{pmatrix}.$$

Es ist $T^{-1} = \frac{1}{2}\begin{pmatrix}1 & -i\\ -i & 1\end{pmatrix}$ und

$$\begin{aligned} T^{-1}AT &= \tfrac{1}{2}\begin{pmatrix}1 & -i\\ -i & 1\end{pmatrix}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}1 & i\\ i & 1\end{pmatrix}\\ &= \tfrac{1}{2}\begin{pmatrix}1 & -i\\ -i & 1\end{pmatrix}\begin{pmatrix}i & 1\\ -1 & -i\end{pmatrix} = \begin{pmatrix}i & 0\\ 0 & -i\end{pmatrix}\end{aligned}$$

eine Diagonalisierbarkeit über $\mathbb{C}$.

(b) Betrachte die Matrix

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

Eigenwert ist 1, mit algebraischer Multiplizität 2. Der Eigenraum zum Eigenwert 1 ist

$$V(1) = \ker\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} = \left\{\begin{pmatrix}x\\y\end{pmatrix}\middle|\, y=0\right\} = K\begin{pmatrix}1\\0\end{pmatrix}.$$

Also ist die geometrische Vielfachheit 1 und A ist nicht diagonalisierbar. Das gilt über jedem Körper K, also $K = \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{F}_2, \ldots$ (vergleiche auch die späteren Paragraphen über die Jordansche Normalform).

(c) Nach Hauptachsentransformation/Spektralsatz (LA I) ist über $\mathbb{R}$ jede symmetrische Matrix diagonalisierbar und über $\mathbb{C}$ jede hermitesche Matrix.

Satz 3.3 motiviert die folgende Definition:

Definition 3.8 Ein Endomorphismus $\varphi : V \to V$ eines beliebigen (nicht notwendig endlich-dimensionalen) K-Vektorraums V heißt diagonalisierbar, wenn V eine Basis aus Eigenvektoren für φ besitzt.

Wegen 3.3. gibt dies für endlich-dimensionales V die alte Definition!

Definition 3.9 Sei V ein K-Vektorraum und sei $(V_i)_{i\in I}$ eine Familie von Untervektorräumen $V_i \subseteq V$, wobei die Indexmenge I nicht notwendig endlich sei.

(a) Die Summe $\sum\limits_{i\in I} V_i$ ist der von den V_i erzeugte Unterraum, d.h., der kleinste Unterraum der alle V_i enthält.

(b) Man sagt, dass die V_i eine direkte Summe bilden (Bez. $\sum\limits_{i\in I} V_i = \bigoplus\limits_{i\in I} V_i$), wenn für jede Familie $i_1, \ldots, i_r \in I$ und jede Familie $v_{i_1}, \ldots, v_{i_r}$ mit $v_{i_j} \in V_{i_j}$ gilt

$$v_{i_1} + \ldots + v_{i_r} = 0 \quad \Rightarrow \quad v_{i_j} = 0 \qquad \forall\, j = 1, \ldots, r\,.$$

(c) Man sagt, dass V die direkte Summe der V_i ist (Bez. $V = \bigoplus\limits_{i \in I} V_i$), wenn $\sum\limits_{i \in I} V_i = \bigoplus\limits_{i \in I} V_i$ und $\sum\limits_{i \in I} V_i = V$.

Lemma 3.10 Ein Endomorphismus $\varphi : V \to V$ ist genau dann diagonalisierbar, wenn gilt

$$V = \bigoplus_{\lambda \in K} V(\lambda)$$

(wobei $V(\lambda) := V(\varphi, \lambda)$).

Beweis Nach dem ersten Teil von Lemma 2.8 bilden die $V(\lambda)$ eine direkte Summe. Besitzt nun V eine Basis aus Eigenvektoren für φ, so wird V von den Eigenräumen $V(\lambda)$ erzeugt, d.h., es ist $\bigoplus\limits_{\lambda \in K} V(\lambda)$ gleich V. Ist umgekehrt $\bigoplus\limits_{\lambda \in K} V(\lambda) = V$, und wählen wir für jedes λ ein Basis von $V(\lambda)$, so erhalten wir insgesamt eine Basis von V, die nach Konstruktion aus Eigenvektoren für φ besteht.
(Hier haben wir benutzt, dass jeder Vektorraum eine Basis besitzt – für unendlich-dimensionale Vektorräume ist dies äquivalent zum Auswahlaxiom)

Der folgende Satz ist wichtig in der Physik (verallgemeinert auf Hilberträume):

Satz 3.11 Sei V ein K-Vektorraum, und seien $\varphi, \psi : V \to V$ zwei Endomorphismen, die **miteinander vertauschen**, d.h., für die gilt

$$(*) \qquad \varphi\psi = \psi\varphi\,.$$

(a) Sei $\lambda \in K$. Ist $v \in V$ ein Eigenvektor von φ zum Eigenwert λ, so ist $\psi(v)$ wieder ein Eigenvektor von φ zum Eigenwert λ, oder $\psi(v) = 0$; d.h., es gilt $\psi(V(\varphi, \lambda)) \subseteq V(\varphi, \lambda)$.

(b) Sind φ und ψ beide diagonalisierbar, so sind sie **simultan diagonalisierbar**, d.h., es gibt eine Basis von V, deren Elemente Eigenvektoren von φ und von ψ sind.

Beweis (a): $\varphi(v) = \lambda \cdot v \Rightarrow \varphi(\psi(v)) = (\varphi\psi)(v) \underset{(*)}{=} (\psi\varphi)(v) = \psi(\varphi(v)) = \psi(\lambda \cdot v) = \lambda \cdot \psi(v)$.
Damit folgt die Behauptung (für die zweite Formulierung beachte, dass immer $\psi(0) = 0$ gilt).

(b): Wir zeigen zuerst

Lemma 3.12 Sei $\psi : V \to V$ ein Endomorphismus, und sei $W \subseteq V$ ein Untervektorraum mit

$$\psi(W) \subseteq W\,.$$

Ist ψ diagonalisierbar, so auch die Einschränkung

$$\psi_{|W} : W \to W\,.$$

Beweis Nach Lemma 3.10 gilt

$$(*) \qquad V = \bigoplus_{\lambda \in K} V(\lambda) \qquad (V(\lambda) = V(\psi, \lambda))\,,$$

und wir haben zu zeigen

Behauptung: $W = \bigoplus\limits_{\lambda \in K} W(\lambda)$

Beweis: Die Summe ist direkt (wegen Lemma 2.8, oder wegen $(*)$ und $W(\lambda) = W \cap V(\lambda)$), also ist zu zeigen, dass W Summe der $W(\lambda)$ ist. Sei $w \in W \setminus \{0\}$. Dann gibt es nach $(*)$ verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_r$ von $\psi : V \to V$ und Eigenvektoren $v_i \in V(\psi, \lambda_i) \setminus \{0\}$ $\quad (i = 1, \ldots, r)$ mit

$$w = \sum_{i=1}^{r} v_i \,.$$

Es genügt zu zeigen, dass $v_i \in W$ ist für alle $i = 1, \ldots, r$; denn $W(\lambda_i) = W \cap V(\lambda_i)$. Für $r = 1$ ist nichts zu zeigen, also sei $r > 1$. Für $j \in \{1, \ldots, r\}$ gilt dann

$$\prod_{\substack{i=1 \\ i \neq j}}^{r} (\psi - \lambda_i) w \in W \quad \text{(wegen } \psi(W) \subseteq W\text{)}$$

und andererseits

$$\prod_{\substack{i=1 \\ i \neq j}}^{r} (\psi - \lambda_i) w = \prod_{\substack{i=1 \\ i \neq j}}^{r} (\psi - \lambda_i) v_j = \prod_{\substack{i=1 \\ i \neq j}}^{r} (\lambda_j - \lambda_i) v_j$$

(wegen $(\psi - \lambda_i) v_i = 0$). Zusammen folgt

$$v_j \in W \,,$$

wegen $\prod\limits_{\substack{i=1 \\ i \neq j}}^{r} (\lambda_j - \lambda_i) \neq 0$, da $\lambda_i \neq \lambda_j \ \forall \ i \neq j$.

Beweis von 3.9 (b): Wegen $V = \bigoplus\limits_{\lambda \in K} V(\varphi, \lambda)$ genügt es zu zeigen, dass jedes $V(\varphi, \lambda) \neq \{0\}$ eine Basis aus Eigenvektoren für ψ hat. Dies gilt aber wegen $\psi(V(\varphi, \lambda)) \subseteq V(\varphi, \lambda)$ (Teil (a)) und Lemma 3.12.

Bemerkung 3.13 (a) Es gilt auch die folgende Umkehrung: Sind die Endomorphismen φ und ψ simultan diagonalisierbar, so vertauschen sie (Beweis?).

§4 Das Minimalpolynom eines Endomorphismus

Sei K ein Körper und sei $\varphi : V \to V$ ein Endomorphismus eines K-Vektorraums.

Definition 4.1 Für ein Polynom

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0$$

$\in K[x]$ definiere

$$f(\varphi) = a_n \varphi^n + a_{n-1} \varphi^{n-1} + \ldots + a_1 \varphi + a_0 \in \ \mathrm{End}(V)$$

wobei $\varphi^i = \varphi \circ \ldots \circ \varphi$ (i-mal) und $a_0 = a_0 \cdot$ id).

Dies liefert eine wohldefinierte Abbildung

$$\begin{array}{rcl} K[x] & \to & \mathrm{End}(V) \\ f(x) & \mapsto & f(\varphi)\,. \end{array}$$

Offenbar gilt für diese

(i) $f + g \mapsto f(\varphi) + g(\varphi)$

(ii) $f \cdot g \mapsto f(\varphi) \cdot g(\varphi)$

d.h., die Abbildung ist ein Ringhomomorphismus. Weiter gilt

(iii) $\lambda f \mapsto \lambda f(\varphi)$

d.h., die Abbildung ist ein **K-Algebren-Homomorphismus**. Hierzu:

Definition 4.2 Eine K-**Algebra** ist ein Ring R, der zusätzlich ein K-Vektorraum ist, so dass gilt:

$$\begin{array}{rcl} \lambda(f \cdot g) & = & (\lambda f) \cdot g \\ & = & f \cdot (\lambda g) \end{array}$$

für $\lambda \in K$ und $f, g \in R$.

Ein **Homomorphismus** $\phi : R_1 \to R_2$ **von** K-**Algebren** ist ein Ringhomomorphismus, der gleichzeitig K-linear ist (d.h., es gilt noch zusätzlich

$$\phi(\lambda f) = \lambda \phi(f)$$

für $\lambda \in K$ und $f \in R_1$).

Lemma 4.3 Ist V endlich-dimensional, so gibt es ein $f \in K[x] \setminus \{0\}$ mit $f(\varphi) = 0$ (der Nullhomomorphismus).

Beweis Ist $\dim V = n$, so ist $\dim \mathrm{End}\ (V) = n^2$. Daher sind

$$1, \varphi, \varphi^2, \varphi^3, \ldots, \varphi^{n^2}$$

linear abhängig; es gibt also eine nicht-triviale Linearkombination zu 0.

Lemma 4.4 Sei V endlich-dimensional. Es gibt ein eindeutig bestimmtes normiertes Polynom kleinsten Grades $p(x) \in K[x]$ mit $p(\varphi) = 0$. Dieses heißt das **Minimalpolynom von** φ. Ist $f(x) \in K[x]$ ein anderes nicht-triviales Polynom mit $f(\varphi) = 0$, so wird f von p geteilt.

Beweis der Behauptung: Sei $p(x)$ ein normiertes Polynom minimalen Grades mit $p(\varphi) = 0$; dies existiert nach 4.3. Sei $f(x) \in K[x] \setminus \{0\}$ ein anderes Polynom mit $f(\varphi) = 0$. Division mit Rest gibt

$$f(x) = q(x) \cdot p(x) + r(x)$$

mit $q, r \in K[x]$, $\deg r(x) < \deg p(x)$. Dann gilt

$$0 = f(\varphi) = q(\varphi) \cdot p(\varphi) + r(\varphi) = r(\varphi)\,.$$

Wegen der Minimalität von $\deg p(x)$ muss also $r(x) = 0$ sein, d.h., p teilt f. Hieraus folgt auch die Eindeutigkeit von p: Ist $\deg f = \deg p$ und ist f ebenfalls normiert, so folgt aus $p \mid f$ die Gleichheit $f = p$.

Lemma 4.5 Die Nullstellen des Minimalpolynoms von φ sind gerade die Eigenwerte von φ. (Das Minimalpolynom $p(x)$ und das charakteristische Polynom $\chi_\varphi(x)$ haben also diesselben Nullstellen!)

Beweis 1) Sei $\lambda \in K$ Eigenwert von φ. Sei $v \neq 0$ ein Eigenvektor zum Eigenwert λ:

$$\varphi v = \lambda v\,.$$

Dann ist für das Minimalpolynom $p(x)$ von φ wegen $\varphi^i v = \lambda^i v$

$$0 = p(\varphi)v = p(\lambda)v\,,$$

also $p(\lambda) = 0$, da $v \neq 0$.

2) Sei λ Nullstelle von $p(x)$. Dann ist

$$p(x) = (x - \lambda) \cdot g(x)$$

mit einem $g(x) \in K[x]$. Angenommen, λ ist kein Eigenwert von φ Dann ist $\varphi - \lambda$ id injektiv. Andererseits ist

$$0 = p(\varphi) = (\varphi - \lambda \text{ id }) \cdot g(\varphi)\,.$$

Es folgt $g(\varphi) = 0$, im Widerspruch zur Minimalität von $\deg p(x)$.

Von nun an sei V **endlich-dimensional**.

Definition 4.6 Zwei Polynome $f_1, f_2 \in K[x]$ heißen teilerfremd, wenn sie nur konstante gemeinsame Teiler haben ($g \mid f_1 \wedge g \mid f_2 \Rightarrow g$ konstant).

Satz 4.7 (1. Zerlegungssatz) Sei $f(x) \in K[x]$ und

$$f(x) = g(x) \cdot h(x)$$

mit teilerfremden Polynomen g und h. Dann ist

$$\ker f(\varphi) = \ker g(\varphi) \bigoplus \ker h(\varphi)\,.$$

Beweis 1) $\ker g(\varphi) \cap \ker h(\varphi) = \{0\}$.
Beweis: Der Untervektorraum $U := \ker g(\varphi) \cap \ker h(\varphi)$ ist φ-invariant, d.h., es ist $\varphi(U) \subseteq U$, denn aus $v \in U$ folgt $g(\varphi)v = 0$ und damit $0 = \varphi g(\varphi)v = g(\varphi)\varphi v$, also $\varphi(v) \in \ker g(\varphi)$; entsprechend folgt $\varphi(U) \subseteq \ker h(\varphi)$. Betrachte nun die Einschränkung $\tilde{\varphi} = \varphi_{|U} : U \to U$ von φ. Nach Definition ist $g(\tilde{\varphi}) = 0$ und $h(\tilde{\varphi}) = 0$. Das Minimalpolynom $\tilde{p}$ von $\tilde{\varphi}$ teilt also g und h, ist also konstant, da g und h teilerfremd sind. Dies ist nur möglich, wenn $W = \{0\}$ ist.

2) Sei $W = \ker f(\varphi)$; dann ist $\ker g(\varphi) \subseteq W$, denn mit $g(\varphi)v = 0$ ist auch $f(\varphi)v = g(\varphi)h(\varphi)v = h(\varphi)g(\varphi)v = 0$. Entsprechend folgt $\ker h(\varphi) \subseteq W$.

3) Wie in 1) folgt $\varphi(W) \subseteq W$. Definiere $\overline{\varphi} = \varphi_{|W} : W \to W$. Dann ist $\ker g(\overline{\varphi}) = (\ker g(\varphi)) \cap W \stackrel{2)}{=} \ker g(\varphi)$ und entsprechend $\ker h(\overline{\varphi}) = \ker h(\varphi)$.

4) Auf W gilt $f(\overline{\varphi}) = 0$ und damit

$$\text{im } g(\overline{\varphi}) \subseteq \ker h(\overline{\varphi})$$

denn für $w \in W$ ist $h(\overline{\varphi})g(\overline{\varphi})w = f(\overline{\varphi})w = 0$.

5) Nach der Rangformel gilt

$$\dim \ker g(\overline{\varphi}) + \dim \text{im } g(\overline{\varphi}) = \dim W .$$

6) Es folgt

$$\ker g(\overline{\varphi}) \oplus \ker h(\overline{\varphi}) = W ,$$

denn $\ker g(\overline{\varphi})$ und $\ker h(\overline{\varphi})$ sind nach 3) und 2) Unterräume von W und bilden nach 1) eine direkte Summe; weiter ist die Dimension der linken Seite

$$\begin{aligned} & \dim \ker g(\overline{\varphi}) + \dim \ker h(\overline{\varphi}) \\ \stackrel{4)}{\geq} \; & \dim \ker g(\overline{\varphi}) + \dim \text{ im } G(\overline{\varphi}) \\ \stackrel{5)}{=} \; & \dim W \; . \end{aligned}$$

Mit 6) und 3) folgt die Behauptung des Satzes.

Sei $f(x) \in K[x]$ mit $f(\varphi) = 0$ (z.B. $f(x)$ das Minimalpolynom). Angenommen, f zerfällt über K, d.h.,

$$f(x) = (x - \lambda_1)^{m_1}(x - \lambda_2)^{m_2} \dots (x - \lambda_r)^{m_r}$$

mit paarweise verschiedenen $\lambda_1, \dots, \lambda_r \in K$ und $m_i \in \mathbb{N}$. Dann ist $(x - \lambda_1)^{m_1}$ teilerfremd zu $\prod\limits_{i=2}^{r}(x - \lambda_i)^{m_i}$, und induktiv folgt aus Satz 4.7:

Lemma 4.8 $V = \ker f(\varphi) = \ker(\varphi - \lambda_1)^{m_1} \oplus \dots \oplus \ker(\varphi - \lambda_r)^{m_r}$.

Dies motiviert:

Definition 4.9 Sei $\lambda \in K$.

(a) Für $i \in \mathbb{N}$ heißt

$$V(\lambda)^{(i)} := V(\varphi, \lambda)^{(i)} := \ker(\varphi - \lambda)^i$$

der ***i*-te verallgemeinerte Eigenraum** von φ zum Eigenwert λ.

(b) Sei $i \in \mathbb{N}$. Ein Vektor $v \in V$ mit

$$(\varphi - \lambda)^i v = 0 \text{ und } (\varphi - \lambda)^{i-1} v \neq 0$$

heißt **verallgemeinerter Eigenvektor *i*-ter Stufe** von φ zum Eigenwert λ.

(c) Der Vektorraum

$$V[\lambda] := V[\varphi, \lambda] := \{v \in V \mid \exists\, i \in \mathbb{N} \text{ mit } (\varphi - \lambda)^i v = 0\}$$

heißt der **verallgemeinerte Eigenraum** von φ zum Eigenwert λ. Die Elemente aus $V[\lambda] \setminus \{0\}$ heißen **verallgemeinerte Eigenvektoren** von φ zum Eigenwert λ.

Statt verallgemeinerter Eigenraum/Eigenvektor sagt man auch **Hauptraum/Hauptvektor**.

Lemma 4.10 (a) Es ist $V(\lambda) = V(\lambda)^{(1)} \subseteq V(\lambda)^{(2)} \subseteq V(\lambda)^{(3)} \subseteq \ldots \subseteq V[\lambda] = \bigcup\limits_{i\in\mathbb{N}} V(\lambda)^{(i)}$. Dies zeigt, dass $V[\lambda]$ ein Unterraum von V ist.

(b) Es ist genau dann $V[\lambda] \neq 0$, wenn λ ein Eigenwert von φ ist.

Beweis (a): Nach Definition ist $V(\lambda) = \ker(\varphi - \lambda) = V(\lambda)^{(1)}$. Weiter ist $V(\lambda)^{(i)} \subseteq V(\lambda)^{(i+1)}$, denn mit $(\varphi - \lambda)^i v = 0$ ist auch $(\varphi - \lambda)^{i+1} v = (\varphi - \lambda)(\varphi - \lambda)^i v = 0$. Die Gleichheit $V[\lambda] = \bigcup_{i\in\mathbb{N}} V(\lambda)^{(i)}$ gilt nach Definition, und dies ist ein Unterraum, denn für $v \in V(\lambda)^{(i)}$ und $w \in V(\lambda)^{(j)}$ liegen v und w in $V(\lambda)^{(i+j)}$, also auch $v + w$.

(b) Sei $i_0 \in \mathbb{N}$ minimal mit $(\varphi - \lambda)^{i_0} v = 0$. Dann ist $(\varphi - \lambda)^{i_0 - 1} v \neq 0$ und

$$(\varphi - \lambda)(\varphi - \lambda)^{i_0-1} v = 0\,,$$

also $(\varphi - \lambda)^{i_0-1} v$ ein Eigenvektor von φ zum Eigenwert λ.

Beispiel 4.11 (a) Betrachte den Endomorphismus von $V = K^2$

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in M_2(K)\,.$$

Es ist $\chi_A(x) = (x-1)^2$, A hat also den Eigenwert 1 mit algebraischer Vielfachheit 2. Es ist

$$A - 1E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq 0\,;$$

daher ist die geometrische Vielfachheit gleich 1. Weiter ist

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

Daher ist das Minimalpolynom von A gleich $p(x) = (x-1)^2$ (für alle echten Teiler g von $p(x)$, nämlich $(x-1)$ und 1, ist $g(A) \neq 0$). Es ist

$$V = K^2 = \ker(A - 1E)^2 = V(A,1)^{(2)} = V[1]$$

d.h., der verallgemeinerte Eigenraum ist ganz K^2, während der Eigenraum zu 1,

$$V(1) = \ker(A - 1E) = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in K^2 \,\middle|\, y = 0 \right\} = K \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}\,,$$

eindimensional ist.

(b) Betrachte

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

Es ist $\chi_A(x) = (x-1)^3$, aber das Minimalpolynom ist gleich $p(x) = (x-1)^2$, denn für

$$A - 1E = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq 0$$

ist $(A-1E)^2 = 0$. Es ist

$$\begin{array}{lclcll} V(1) & = & \ker(A-E) & = & Ke_1 \oplus Ke_3 & \text{zweidimensional} \\ V[1] & = & \ker(A-E)^2 & = & K^3 & \text{dreidimensional} \end{array}$$

Sei wieder $\varphi : V \to V$ ein Endomorphismus eines endlich-dimensionalen K-Vektorraums, und $f(x) \in K[x]$ ein Polynom mit $f(\varphi) = 0$ (z. B. das Minimalpolynom), welches über K zerfällt. Wir schreiben

$$f(x) = \prod_{i=1}^{r} (x - \lambda_i)^{m_i}$$

mit paarweise verschiedenen $\lambda_1, \ldots, \lambda_r \in K$ und $m_i \in \mathbb{N}$.

Proposition 4.12 Es ist

$$V = V[\lambda_1] \oplus \ldots \oplus V[\lambda_r]\,.$$

Beweis Nach 4.8 ist

$$(*) \qquad V = \bigoplus_{i=1}^{r} \ker(\varphi - \lambda_i \text{ id})^{m_i}\,,$$

und es genügt zu zeigen

Lemma 4.13 $\ker(\varphi - \lambda_i \text{ id})^{m_i} = V[\lambda_i]$.

Beweis Die Inklusion $\subseteq$ gilt nach Definition. Andererseits gilt für $m'_1, \ldots, m'_r \in \mathbb{N}$ mit $m'_i \geq m_i$ auch $g(\varphi) = 0$ für

$$g(x) = \prod_{i=1}^{r} (x - \lambda_i)^{m'_i}\,,$$

also nach 4.8 auch

$$(**) \qquad V = \bigoplus_{i=1}^{r} \ker(\varphi - \lambda_i \text{ id})^{m'_i}$$

Wegen $\ker(\varphi - \lambda_i \text{ id})^{m_i} \subseteq \ker(\varphi - \lambda_i \text{ id})^{m'_i}$ folgt hieraus

$$\ker(\varphi - \lambda_i \text{ id})^{m_i} = \ker(\varphi - \lambda_i \text{ id})^{m'_i} \quad \forall\, i$$

nach dem folgenden Lemma. Da dies für alle $m'_i \geq m_i$ gilt, folgt die Behauptung (siehe 4.10 (a)).

Lemma 4.14 Sei V ein K-Vektorraum, und sei

$$(*) \qquad V \quad = \quad \bigoplus_{i=1}^{t} V_i$$

$$(**) \qquad V = \bigoplus_{i=1}^{t} V_i'$$

für Unterräume V_i, V_i' mit $V_i \subseteq V_i' \quad (i = 1, \ldots, t)$. Dann ist $V_i = V_i'$ für alle $i = 1, \ldots, t$.

Beweis Sei $v \in V_j'$. Dann ist $v = \sum\limits_{i=1}^{t} v_i$ mit $v_i \in V_i \quad (i = 1, \ldots, t)$. Wegen $v_i \in V_i'$ ist dies auch die Zerlegung von v in Komponenten entsprechend $(**)$. Da diese eindeutig ist und $v \in V_j'$, gilt $v_i = 0$ für $i \neq j$, und damit $v = v_j \in V_j$. (Der Beweis geht genauso für unendliche Summen).

Corollar 4.15 Zerfällt das Minimalpolynom von φ, und sind $\lambda_1, \ldots, \lambda_s$ die verschiedenen Eigenwerte von φ, so ist

$$V = \bigoplus_{i=1}^{s} V[\lambda_i]\,.$$

Beweis Wende 4.12 auf das Minimalpolynom an; dann sind die λ_i in 4.12 gerade die Eigenwerte von φ nach 4.5.

(Dies ist kein Widerspruch zu 4.12: ist λ_i kein Eigenwert von φ, so ist nach 4.10 (b) $V[\lambda_i] = \{0\}$).

Die Zerlegung 4.15 hängt nur von φ ab und ist daher eindeutig. Wir bemerken noch, dass die $V[\lambda_i]$ **φ-invariante Unterräume** sind, d.h., es gilt $\varphi(V[\lambda_i] \subseteq V[\lambda_i]$. Was bedeutet diese Zerlegung für die Matrixdarstellung? Antwort:

Lemma 4.16 Sei $\varphi : V \to V$ ein Endomorphismus eines K-Vektorraums V, und sei

$$V = \bigoplus_{i=1}^{s} V_i$$

eine Zerlegung von V in φ-invariante Unterräume ($\varphi V_i \subseteq V_i$). Sei $b(i) = (b_1^i, \ldots, b_{d_i}^i)$ eine Basis von V_i, und sei

$$A_i = M_{b(i)}^{b(i)}(\varphi_{|V_i}) \in M_{d_i}(K)$$

die Matrixdarstellung der Einschränkung

$$\varphi_{|V_i} : V_i \to V_i$$

bezüglich $b(i) \quad (i = 1, \ldots, s)$. Dann ist

$$b = (b(1), \ldots b(s)) := (b_1^1, \ldots, b_{d_1}^1, b_1^2, \ldots, b_{d_2}^2, \ldots, b_1^s, \ldots, b_{d_s}^s)$$

eine Basis von V, und die Matrixdarstellung von φ bezüglich b hat die Gestalt

$$A = M_b^b(\varphi) = \begin{array}{r} d_1\left\{\vphantom{\boxed{A_1}}\right. \\ d_2\left\{\vphantom{\boxed{A_2}}\right. \\ d_3\left\{\vphantom{\boxed{A_3}}\right. \\ \\ d_n\left\{\vphantom{\boxed{A_s}}\right. \end{array} \begin{pmatrix} \boxed{A_1} & & & & \\ & \boxed{A_2} & & 0 & \\ & & \boxed{A_3} & & \\ & 0 & & \ddots & \\ & & & & \boxed{A_s} \end{pmatrix},$$

(Man nennt eine solche Matrix auch eine **Block-Diagonalmatrix**).

Beweis: Alles folgt sofort aus den Definitionen!

Für die Matrixdarstellung von φ genügt es also, die Einschränkungen

$$\varphi_{|\ V[\lambda_i]} : V[\lambda_i] \to V[\lambda_i]$$

auf die verallgemeinerten Eigenräume zu betrachten.

Beobachtung 4.17 Auf $V[\lambda_i]$ ist $\varphi - \lambda_i$ id nilpotent, d.h., es gibt ein $m \in \mathbb{N}$ mit $\Big((\varphi - \lambda_i \text{ id })_{|\ V[\lambda_i]}\Big)^m = 0$.

Beweis Mit den Bezeichnungen von 4.12 ist $(\varphi - \lambda_i \text{ id})^{m_i}_{|\ V[\lambda_i]} = 0$.

Wir werden im nächsten Paragraphen eine besonders einfache Matrixdarstellung für nilpotente Endomorphismen finden. Wenden wir dies auf $(\varphi - \lambda_i \text{ id})_{|\ V[\lambda_i]}$ an, so erhalten wir dafür eine einfache Matrix B_i, und die Matrix A_i von $\varphi_{|\ V[\lambda_i]} = \lambda_i \text{ id } + (\varphi - \lambda_i \text{ id})$ ist

$$A_i = \lambda E + B_i$$

(die Matrixdarstellung von λ id ist λE, bezüglich *jeder* Basis). Dies wird dann die Jordansche Normalform sein.

§5 Die Jordansche Normalform

Sei K ein Körper und V ein K-Vektorraum.

Definition 5.1 Ein Endomorphismus $\varphi : V \to V$ heißt **nilpotent**, wenn es ein $n \in \mathbb{N}$ gibt mit $\varphi^n = 0$. Ist dann $m \in \mathbb{N}$ minimal $\varphi^m = 0$, so heißt φ **nilpotent von der Stufe** (oder**dem Grad** m.

Proposition 5.2 Sei $\nu : V \to V$ ein nilpotenter Endomorphismus. Für $w \in V$ ist der Unterraum

$$L_\nu(w) := < w, \nu w, \nu^2 w, \nu^3 w, \ldots >_K$$

ein ν-invarianter Unterraum von V. (Wir schreiben einfach $\nu^i w$ für $\nu^i(w)$). Sei $m \in \mathbb{N}_0$ minimal mit

$$\nu^m w = 0\,;$$

dies m heißt die **Periode von** w. Dann ist

$$(w, \nu w, \nu^2 w, \ldots, \nu^{m-1} w)$$

eine Basis von $L_\nu(w)$.

Beweis Die ν-Invarianz von $L_\nu(w)$ ist klar, da $\nu\nu^i w = \nu^{i+1} w \in L_\nu(w) \quad \forall\ i \in \mathbb{N}_0$. Angenommen

$$a_0 w + a_1 \nu w + a_2 \nu^2 w + \ldots + a_{m-1}\nu^{m-1} w = 0$$

mit $a_0, \ldots, a_{m-1} \in K$, nicht alle null. Sei $0 \leq i \leq m$ minimal mit $a_i \neq 0$. Dann folgt durch Anwenden von ν^{m-i-1}

$$a_i \nu^{m-1} w = 0$$

(da $\nu^m w = 0$), im Widerspruch dazu, dass $\nu^{m-1} w \neq 0$.

Beobachtung 5.3 Bezüglich der obigen Basis ist die Matrix von ν auf $L_\nu(w)$ sehr einfach:

(a) Ordnen wir die Basis durch

$$\begin{array}{ccccc} (w, & \nu w, & \nu^2 w, & \ldots & , \nu^{m-1} w)\,, \\ \shortparallel & \shortparallel & \shortparallel & & \shortparallel \\ b_1 & b_2 & b_3 & & b_m \end{array}$$

so ist $\nu b_i = b_{i+1}$ für $i = 1, \ldots, m-1$, und $\nu b_m = 0$, also die Matrix bezüglich dieser Basis

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

(**untere** Dreiecksgestalt, Einsen in der **unteren Nebendiagonalen**).

(b) Ordnen wir die Basis so:

$$\begin{array}{ccccc} (w^{m-1} w, & \nu^{m-2} w, & \ldots & , \nu w, & w)\,, \\ \shortparallel & \shortparallel & & \shortparallel & \shortparallel \\ b_1 & b_2 & & b_{m-1} & b_m \end{array}$$

so gilt $\nu b_i = b_{i-1} \quad \forall\, i = 2, \ldots, m$, $\nu b_1 = 0$; die Matrix ist dann

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

(**obere** Dreiecksgestalt, Einsen in der **oberen Nebendiagonalen**). Diese letzte Basis nennen wir eine Jordan-Basis von $L_\nu(w)$.

Satz 5.4 (2. Zerlegungssatz) Sei V endlich-dimensional und $\nu : V \to V$ ein nilpotenter Endomorphismus. Dann gibt es $w_1, \ldots, w_t \in V$ so, dass

$$V = \bigoplus_{j=1}^{t} L_\nu(w_j)\,.$$

Beweis: durch Induktion über $\dim V$, wobei der Fall $\dim V = 0$ trivial ist. Sei also $\dim V \geq 1$ und die Behauptung für kleinere Dimensionen bewiesen. Sei m der Nilpotenzgrad von ν (also $m \in \mathbb{N}$ minimal mit $\nu^m = 0$) und $w_1 \in V$ ein Vektor mit $\nu^{m-1}w_1 \neq 0$ (also mit Periode von w_1 gleich Nilpotenzgrad von ν). Wir behaupten nun:

Lemma 5.5 Es gibt einen ν-invarianten Unterraum $U \subseteq V$ mit

$$V = L_\nu(w_1) \oplus U\,.$$

Hieraus folgt Satz 5.4: Wegen $L_\nu(w_1) \neq 0$ und $\dim V = \dim L_\nu(w_1) + \dim U$ ist dann nämlich $\dim U < \dim V$, und nach Induktionsannahme gibt es $w_2, \ldots, w_t \in U$ mit $U = L_\nu(w_2) \oplus \ldots \oplus L_\nu(w_t)$. Mit 5.5 folgt $V = L_\nu(w_1) \oplus \ldots \oplus L_\nu(w_t)$.

Beweis von Lemma 5.5 Sei $U \subseteq V$ ein ν-invarianter Unterraum maximaler Dimension, so dass $L_\nu(w_1) \cap U = 0$ ist. Dann bilden U und $L_\nu(w_1)$ eine direkte Summe, und wir haben noch zu zeigen, dass $V = L_\nu(w_1) + U$ ist. Angenommen, dies ist nicht der Fall. Dann gibt es ein $v \in V$ mit $v \notin L_\nu(w_1) + U$. Sei $j \leq m$ minimal mit $\nu^j v \in L_\nu(w_1) + U$ (beachte $\nu^m v = 0 \in L_\nu(w_1) + U$). Für $v' = \nu^{j-1}v$ gilt dann

$$v' \notin L_\nu(w_1) + U, \quad \nu v' \in L_\nu(w_1) + U\,, \tag{5.5.1}$$

insbesondere also

$$\nu v' = v'' + u$$

mit $v'' \in L_\nu(w_1)$ und $u \in U$. Wegen $\nu^m = 0$ ist

$$0 = \nu^m v' = \nu^{m-1}v'' + \nu^{m-1}u\,,$$

also $L_\nu(w_1) \ni \nu^{m-1}v'' = -\nu^{m-1}u \in U$, und somit $\nu^{m-1}v'' = 0$ wegen $L_\nu(w_1) \cap U = 0$. Andererseits ist nach Proposition 5.2

$$\nu'' = \lambda_0 w_1 + \lambda_1 \nu w_1 + \ldots + \lambda_{m-1}\nu^{m-1}w_1\,, \tag{5.5.2}$$

mit $\lambda_0, \ldots, \lambda_{m-1} \in K$ und durch Anwenden von ν^{m-1} folgt

$$0 = \nu^{m-1} v'' = \lambda_0 \nu^{m-1} w_1$$

und damit $\lambda_0 = 0$, weil $\nu^{m-1} w_1 \neq 0$. Aus (5.5.2) folgt nun $v'' = \nu v'''$ mit $v''' = \lambda_1 w_1 + \ldots + \lambda_{m-1} \nu^{m-1} w_1 \in L_\nu(w_1)$, so dass

$$\nu v' = \nu v''' + u\,.$$

Setze $\tilde{v} = v' - v'''$. Dann gilt immer noch $\tilde{v} \notin L_\nu(w_1) + U$ (sonst wäre auch $v' = \tilde{v} + v''' \in L_\nu(w_1) + U$, im Widerspruch zu (5.5.1)) und weiter $\nu(\tilde{v}) = u \in U$. Aus der letzten Beziehung folgt, dass

$$\tilde{U} = U + K\tilde{v}$$

ein ν-invarianter Unterraum ist, und zwar ein größerer als U (wegen $\tilde{v} \notin U$). Weiter ist $L_\nu(w_1) \cap \tilde{U} = 0$: Wäre $0 \neq v_1 \in \tilde{U} \cap L_\nu(w_1)$ etwa $v_1 = u_0 + \lambda\tilde{v}$ mit $u_0 \in U$ und $\lambda \in K$, so wäre $\lambda \neq 0$ wegen $L_\nu(w_1) \cap U = 0$ und daher $\tilde{v} = \lambda^{-1}(v_1 - u_0) \in L_\nu(w_1) + U$ im Widerspruch zur Annahme. Damit erfüllt $\tilde{U}$ diesselben Eigenschaften wie U und es ergibt sich insgesamt ein Widerspruch zur Maximalität von U. Daher kann das v wie oben nicht existieren, und die Behauptung des Lemmas ist bewiesen.

Definition 5.6 Sei $\lambda \in K$ und $m \in \mathbb{N}$. Die Matrix

$$J_m(\lambda) := \begin{pmatrix} \lambda & 1 & & & & \\ & \ddots & \ddots & & 0 & \\ & & \ddots & \ddots & & \\ & 0 & & \ddots & \ddots & \\ & & & & \ddots & 1 \\ & & & & & \lambda \end{pmatrix} \in M_m(K)$$

heißt **Jordankästchen der Größe** m (oder $m \times m$-**Jordanmatrix**) zum Eigenwert λ.

Satz 5.7 (Jordan-Normalform) Sei $\varphi : V \to V$ ein Endomorphismus eines endlichdimensionalen K Vektorraums V. Das Minimalpolynom von φ zerfalle in $K[x]$ (z.B. $K = \mathbb{C}$ oder K algebraisch abgeschlossen). Dann gibt es eine Basis b von V so, dass die Matrixdarstellung bezüglich b Blockdiagonalenform hat

$$A = M_b^b(\varphi) = \begin{pmatrix} \boxed{J_1} & & & & \\ & \boxed{J_2} & & 0 & \\ & & \ddots & & \\ & 0 & & \ddots & \\ & & & & \boxed{J_s} \end{pmatrix}$$

mit Jordankästchen J_α. Die Jordankästchen sind bis auf ihre Reihenfolge eindeutig durch φ bestimmt.

Beweis **Existenz:** Seien $\lambda_1, \dots, \lambda_r$ die verschiedenen Eigenwerte von φ. Wegen der Zerlegung

$$V = \bigoplus_{i=1}^{r} V[\lambda_i]$$

(1. Zerlegungssatz bzw. Corollar 4.15) und der daraus folgenden Blockdiagonalform (Lemma 4.16) genügt es, alle Einschränkungen $\varphi_i = \varphi_{|\,V[\lambda_i]}$ zu betrachten: Finden wir für jeden dieser Endomorphismen eine Basis $b(i)$ von $V[\lambda_i]$, so dass die darstellende Matrix $M^{b(i)}_{b(i)}(\varphi_i) = A_i$ Jordan-Normalform hat, so erhalten wir eine Jordan-Form

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{pmatrix}$$

für die Matrix von φ durch Aneinandersetzen der Basen $b(1), \dots, b(r)$.

Auf $V[\lambda_i]$ ist $\nu_i = \varphi_i - \lambda_i$ nilpotent. Zerlegen wir nun

$$V[\lambda_i] = \bigoplus_{j=1}^{t_i} L_{\nu_i}(w^i_j)$$

gemäß dem 2. Zerlegungssatz 5.4, so erhalten wir auf jedem $L_{\nu_i}(w_j)$ durch eine Jordanbasis ein Jordankästchen, mit Nullen in der Diagonalen (Beobachtung 5.3(b)), und folglich für $\varphi_i = \lambda_i + \nu_i$ gerade ein Jordankästchen

$$J_{m^i_j}(\lambda_i) = \left.\begin{pmatrix} \lambda_i & 1 & & & \\ & \ddots & \ddots & 0 & \\ & & \ddots & \ddots & \\ & 0 & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix}\right\} m^i_j$$

wenn m^i_j = Periode von w^i_j. Durch Aneinanderreihen erhalten wir insgesamt eine Jordan-Normalform A_i auf $V[\lambda_i]$ wie gewünscht.

Eindeutigkeit bis auf Anordnung: Diese folgt aus dem nächsten Satz, denn wir haben nur zu zeigen, dass für jeden Eigenwert λ die Anzahl der $m \times m$-Jordankästchen zum Eigenwert λ nur von φ abhängt.

Satz 5.8 (a) Für die Anzahl $a(\lambda)$ der Jordankästchen J_α zum Eigenwert λ (d.h., mit $J_\alpha = J_m(\lambda)$ für ein m) gilt

$$a(\lambda) = \dim V(\lambda)\,.$$

(b) Für die Anzahl $a(\lambda, m)$ der Jordankästchen J_α der Größe m zum Eigenwert λ gilt

$$\begin{aligned} a(\lambda, m) &= 2\dim V(\lambda)^{(m)} - \dim V(\lambda)^{(m+1)} - \dim V(\lambda)^{(m-1)} \\ &= \mathrm{rg}\,(\varphi - \lambda)^{m+1} + \mathrm{rg}(\varphi - \lambda)^{m-1} - 2\ \mathrm{rg}(\varphi - \lambda)^m\,. \end{aligned}$$

(Hierbei ist $V(\lambda)^{(0)} := \{0\}$).

Bemerkungen 5.9 (1) Für einen Endomorphismus $\psi : V \to V$ und eine darstellende Matrix $A = M_b^b(\psi)$ von ψ gilt $\dim\ker\psi = \dim\ker A$, denn der Isomorphismus $\varphi_b : K^n \xrightarrow{\sim} V$ induziert einen Isomorphismus $\ker A \xrightarrow{\sim} \ker\psi$ (Beweis?).

(2) Ist weiter $V = \bigoplus\limits_{\alpha=1}^{s} V_\alpha$ eine direkte Summe von ψ-invarianten Unterräumen V_α, so gilt offenbar $\ker\psi = \bigoplus\limits_{\alpha=1}^{s} \ker\psi_\alpha$ für $\psi_\alpha = \psi_{|V_\alpha} : V_\alpha \to V_\alpha$ (Beweis?), also $\dim\ker\psi = \sum\limits_{\alpha=1}^{s} \dim\ker\psi_\alpha$.

(3) Dies können wir mit der ersten Bemerkung auch so sehen: Wählen wir gemäß Lemma 4.16 Basen $b(\alpha)$ für die V_α und ist $A_\alpha = M_{b(\alpha)}^{b(\alpha)}(\psi_\alpha)$, und nehmen wir die Basis $b = (b(1), \dots, b(s))$, so erhalten wir nach 4.16 die Blockdiagonalmatrix

$$(5.9.1) \qquad A = \begin{pmatrix} \boxed{A_1} & & & & \\ & \boxed{A_2} & & 0 & \\ & & \ddots & & \\ & 0 & & \ddots & \\ & & & & \boxed{A_s} \end{pmatrix}.$$

In dieser Situation sieht man nun ganz allgemein $\ker A = \bigoplus\limits_{\alpha=1}^{s} \ker A_\alpha$ (Das ist eigentlich derselbe Schluss wie in (2), nur konkreter), und damit

$$(5.9.2) \qquad \dim\ker A = \sum_{\alpha=1}^{s} \dim\ker A_\alpha\,.$$

(4) Die letzte Gleichung können wir auch so sehen: Aus der Blockdiagonalgestalt (5.9.1) sehen wir sofort

$$(5.9.3) \qquad \mathrm{rg}\ A = \sum_{\alpha=1}^{s}\ \mathrm{rg}\ A_\alpha\,.$$

Ist A eine $(n \times n)$-Matrix und jeweils A_α eine $(n_\alpha \times n_\alpha)$-Matrix, so gilt $n = \sum\limits_{\alpha=1}^{s} n_\alpha$, und aus den Rangformeln $\dim\ker A = n -\ \mathrm{rg}\ A$, $\dim\ker A_\alpha = n_\alpha -\ \mathrm{rg}\ A_\alpha$ folgt (5.9.2).

(5) Angewandt auf den Endomorphismus $(\varphi - \lambda)^i$ (welcher die V_α ebenfalls in sich

überführt und durch die Blockdiagonalmatrix

$$(A - \lambda E)^i = \begin{pmatrix} (A_1 - \lambda E_{n_1})^i & & & \\ & (A_2 - \lambda E_{n_2})^i & & \\ & & \ddots & \\ & & & (A_s - \lambda E_{n_s})^i \end{pmatrix}$$

dargestellt wird) folgt wegen $V(\varphi, \lambda)^{(i)} = \ker(\varphi - \lambda)^i$

$$V(\varphi, \lambda)^{(i)} = \bigoplus_{\alpha=1}^{s} V(\varphi_\alpha, \lambda)^{(i)} .$$

Beweis von Satz 5.8: Sei

$$A = \begin{pmatrix} \boxed{J_1} & & & & \\ & \boxed{J_2} & & 0 & \\ & & \ddots & & \\ & 0 & & \ddots & \\ & & & & \boxed{J_s} \end{pmatrix}$$

die Matrixdarstellung von φ, mit Jordankästchen $J_\alpha (\alpha = 1, \dots, s)$.

(a): Offenbar ist für jede Jordanmatrix

$$J_m(\mu) = \begin{pmatrix} \mu & 1 & & & \\ & \ddots & \ddots & 0 & \\ & & \ddots & \ddots & \\ & 0 & & \ddots & 1 \\ & & & & \mu \end{pmatrix} \in M_m(K)$$

der Eigenraum zu μ eindimensional, da

$$\text{rg } (J_m(\mu) - \mu E) = \text{ rg} \begin{pmatrix} 0 & 1 & & & \\ & \ddots & \ddots & 0 & \\ & & \ddots & \ddots & \\ & 0 & & \ddots & 1 \\ & & & & 0 \end{pmatrix} = m - 1 \, ,$$

und der Eigenraum zu $\lambda \neq \mu$ null, da $J_m(\mu) - \lambda E$ dann invertierbar ist. Es folgt

$$\begin{aligned}
\dim V(\varphi, \lambda) &= \dim \ker(\varphi - \lambda \,\mathrm{id}) \\
&= \dim \ker(A - \lambda E) \\
&= \sum_{i=1}^{r} \dim \ker(J_i - \lambda E_{m_i}) \quad (\text{falls } J_i \in M_{m_i}(K)) \\
&= \text{Anzahl der Jordankästchen zum Eigenwert } \lambda\,.
\end{aligned}$$

(b): Es ist für $J_m(\mu) : K^m \to K^m$:

$$\dim V(J_m(\mu), \mu)^{(i)} = \begin{cases} i & , \quad i \leq m \\ m & , \quad i > m \end{cases} .$$

Dies folgt z.B. aus Proposition 5.2, oder da

$$(J_m(\mu) - \mu E_m)^i = \begin{pmatrix} 0 & 1 & & & \\ & \ddots & \ddots & 0 & \\ & & \ddots & \ddots & \\ & 0 & & \ddots & 1 \\ & & & & 0 \end{pmatrix}^i = \begin{array}{c} \scriptstyle (i+1)\text{-te Stelle} \\ \downarrow \\ \begin{pmatrix} 0 & & 1 & & 0 \\ & \ddots & & \ddots & \\ & & & 0 & 1 \\ & & \ddots & & \\ & 0 & & \ddots & \\ & & & & 0 \end{pmatrix} \end{array}$$

Andererseits ist

$$V(J_m(\mu), \lambda)^{(i)} = 0 \text{ für } \lambda \neq \mu\,.$$

Daher gilt

$$\dim V(J_m(\mu), \lambda)^{(i)} - \dim V(J_m(\mu), \lambda)^{(i-1)} = \begin{cases} 1 & , \quad i \leq m \text{ und } \lambda = \mu \\ 0 & , \quad i > m \text{ oder } \lambda \neq \mu \end{cases} .$$

Mit der Vorbemerkung folgt: Die Anzahl $a(\lambda, \geq i)$ der Jordankästchen mit mindestens i Zeilen zum Eigenwert λ ist

$$a(\lambda, \geq i) = \dim V(\varphi, \lambda)^{(i)} - \dim V(\varphi, \lambda)^{(i-1)}\,.$$

Es folgt

$$a(\lambda, i) = a(\lambda, \geq i) - a(\lambda, \geq i+1) = \text{ erste Formel in } (b)\,.$$

Die zweite Formel folgt hieraus, da nach dem Rangsatz

$$\dim V(\varphi, \lambda)^{(i)} = \dim \ker(\varphi - \lambda \,\mathrm{id})^i = \dim V - \operatorname{rg}\,(\varphi - \lambda \,\mathrm{id})^i\,.$$

Für die Matrizen bedeutet 5.7

Satz 5.10 Die Matrix $A \in M_n(K)$ habe ein zerfallendes Minimalpolynom. Dann ist A ähnlich zu einer Matrix in Jordan-Normalform, d.h., es gibt eine invertierbare Matrix

$B \in M_n(K)$ derart, dass

$$B^{-1}AB = \begin{pmatrix} \boxed{J_1} & & & & \\ & \boxed{J_2} & & 0 & \\ & & \ddots & & \\ & 0 & & \ddots & \\ & & & & \boxed{J_s} \end{pmatrix}$$

mit Jordankästchen

$$J_\alpha = J_{m_\alpha}(\lambda_\alpha) = \begin{pmatrix} \lambda_\alpha & 1 & & & \\ & \ddots & \ddots & 0 & \\ & & \ddots & \ddots & \\ & 0 & & \ddots & 1 \\ & & & & \lambda_\alpha \end{pmatrix} \in M_{m_\alpha}(K)\,,$$

($\alpha = 1, \ldots, s$; hierbei sind die Eigenwerte $\lambda_1, \ldots, \lambda_s$ nicht notwendig paarweise verschieden). Die Jordankästchen J_α sind dabei bis auf die Reihenfolge eindeutig bestimmt.

Beweis Dies folgt aus 5.7, da sich bei Übergang von der Standardbasis $e = (e_1, \ldots, e_n)$ zu einer anderen Basis $b = (b_1, \ldots, b_n)$ die Matrix A gerade in die Matrix $B^{-1}AB$ transformiert, wobei B die Matrix mit den Spaltenvektoren $b_1, \ldots, b_n$ ist.

Corollar 5.11 Sei $K = \mathbb{C}$ (oder algebraisch abgeschlossen). Dann ist jede Matrix $A \in M_n(K)$ ähnlich zu einer Matrix in Jordan-Normalform.

(b) Zwei Matrizen $A, A' \in M_n(K)$ sind genau dann ähnlich, wenn $a(i, A, \lambda) = a(i, A', \lambda)$ für alle $i \in \mathbb{N}$ und alle $\lambda \in K$. Dies ist auch genau dann der Fall, wenn $\dim V(i, A, \lambda) = \dim V(i, A', \lambda)$ für alle $i \in \mathbb{N}$ und alle $\lambda \in K$.

§6 Anwendungen und Berechnung der Jordan-Normalform

Sei K ein Körper und $\varphi : V \to V$ ein Endomorphismus eines endlich-dimensionalen K-Vektorrraums.

Proposition 6.1 Sei $K = \mathbb{C}$ oder ein algebraisch abgeschlossener Körper, und seien $\lambda_1, \ldots, \lambda_r \in K$ die verschiedenen Eigenwerte von φ. Sei m_i die algebraische Vielfachheit von λ_i, $m_i^{\text{geom}} = \dim V(\lambda_i)$ die geometrische Vielfachheit von λ_i, und n_i der Nilpotenzgrad von $(\varphi - \lambda_i)_{| \, V[\lambda_i]}$. Sei

$$A = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{pmatrix}$$

eine Matrixdarstellung in Jordan-Normalform für φ (Diese existiert nach 5.11). Wie in Satz 5.8 sei $a(\lambda_i, m)$ die Anzahl der Jordankästchen J_α der Größe m zum Eigenwert λ_i (nach Satz 5.8 hängt $a(\lambda_i, m)$ nur von φ ab).

(a) Es ist $m_i = \dim V[\lambda_i] \geq n_i$.

(b) Das Minimalpolynom von φ ist

$$p_\varphi(x) = \prod_{i=1}^{r} (x - \lambda_i)^{n_i} \, .$$

(c) Es ist

$$\begin{aligned} m_i &= \sum_{m=1}^{\dim V} a(\lambda_i, m) m, \\ m_i^{\text{geom}} &= \sum_{m=1}^{\dim V} a(\lambda_i, m) = \text{ Anzahl aller Jordankästchen } J_\alpha \text{ zum Eigenwert } \lambda_i, \\ n_i &= \max\{m \mid a(\lambda_i, m) \neq 0\} = \text{ maximale Größe der Jordankästchen zum Eigenwert } \lambda_i \, . \end{aligned}$$

Beweis: (a): Sei $\varphi_i = \varphi_{| \, V[\lambda_i]} : V[\lambda_i] \to V[\lambda_i]$. Dann gilt wegen der Zerlegung $V \overset{(*)}{=} \bigoplus_{i=1}^{r} V[\lambda_i]$ für das charakteristische Polynom von φ

$$\chi_\varphi(x) = \prod_{i=1}^{r} \chi_{\varphi_i}(x) \, .$$

Dies folgt aus der Blockdiagonalform für die Matrixdarstellung, die man aus der Zerlegung $(*)$ von V erhält, siehe 4.16. Aus der Jordan-Normalform für φ_i auf $V[\lambda_i]$ (siehe Beweis von 5.7: obere Dreiecksmatrix, alle Einträge auf der Diagonalen gleich λ_i) folgt

$$\chi_{\varphi_i}(x) = (x - \lambda_i)^{\dim V[\lambda_i]} \, .$$

Durch Vergleich mit

$$\chi_\varphi(x) = \prod_{i=1}^{r} (x - \lambda_i)^{m_i}$$

(Definition von m_i) folgt $m_i = \dim V[\lambda_i]$. Die Abschätzung $n_i \leq \dim V[\lambda_i]$ folgt aus Proposition 5.2: $\nu_i = \varphi_i - \lambda_i$ ist nilpotent vom Grad n_i auf $V[\lambda_i]$, und ist $w \in V[\lambda_i]$ ein Vektor mit $\nu_i^{n_i - 1} w \neq 0$, also mit Periode n_i, so gilt nach 5.2

$$n_i = \dim L_{\nu_i}(w) \leq \dim V[\lambda_i] \, .$$

(b): Sei $f(x) = \prod_{i=1}^{r} (x - \lambda_i)^{n_i}$. Dann ist für alle $j = 1, \ldots, r$

$$f(\varphi)_{| \, V[\lambda_j]} = (\varphi_j - \lambda_j)^{n_j} \prod_{i \neq j} (\varphi_j - \lambda_i)^{n_i} = 0 \, ,$$

da $(\varphi_i - \lambda_i)^{n_i} = 0$. Wegen $V = \bigoplus_{j=1}^{r} V[\lambda_j]$ folgt $f(\varphi) = 0$ auf ganz V. Angenommen $g \in K[x]$ ist ein normierter Teiler von f mit $g(\varphi) = 0$ und $g(x) \neq f(x)$. Dann ist

$$g(x) = \prod_{i=1}^{r} (x - \lambda_i)^{\ell_i}$$

mit $\ell_i \in \mathbb{N}_0$, $\ell_i \leq n_i$, und $\ell_j < n_j$ für ein $j \in \{1, \dots, r\}$ (g zerfällt in Linearfaktoren, jede Nullstelle von g ist auch eine von f, die Multiplizitäten der Nullstellen von g (also die ℓ_i) sind höchstens so groß wie für f (also die n_i), und können für $g \neq f$ nicht alle gleich denen für f sein). Dann ist

$$g(\varphi)_{| \, [\lambda_j]} = (\varphi_j - \lambda_j)^{\ell_j} \prod_{i \neq j} (\varphi_j - \lambda_i)^{\ell_i} \,,$$

wobei $(\varphi_j - \lambda_i)^{\ell_j} \neq 0$, da $\ell_j < n_j =$ Nilpotenzgrad von $\varphi_j - \lambda_j$, und jedes $(\varphi_j - \lambda_i)$ für $i \neq j$ invertierbar ist, aufgrund der Jordan-Normalform für φ_j. Also ist $g(\varphi)_{| \, V[\lambda_j]} \neq 0$ und damit $g(\varphi) \neq 0$ – Widerspruch!

(c): Die erste Aussage folgt aus dem Beweis von 5.7 (auf $V[\lambda_i]$ erhalten wir Jordankästchen zum Eigenwert λ_i), und die zweite Aussage folgt aus 5.8 (a): mit den dortigen Bezeichnungen ist

$$m_i^{\text{geom}} = \dim V(\lambda_i) = a(\lambda_i) = \sum_{m=1}^{\dim V} a(\lambda_i, m) \,.$$

Für die dritte Aussage wenden wir die folgende Beobachtung auf den nilpotenten Endomorphismus $\varphi_i - \lambda_i : V[\lambda_i] \to V[\lambda_i]$ an, wobei wir beachten, dass der Nilpotenzgrad von $J_m(\lambda_i) - \lambda_i E_m$ nach dem Beweis von Satz 5.8 gleich m ist: Ist

$$N = \begin{pmatrix} N_1 & & & & \\ & N_2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & N_t \end{pmatrix}$$

eine Blockdiagonalmatrix mit nilpotenten Matrizen N_β, so ist N nilpotent, mit Nilpotenzgrad gleich dem Maximum der Nilpotenzgrade der N_β.

Bemerkung 6.2 Hier haben wir Folgendes benutzt:

Sind

$$A = \begin{pmatrix} A_1 & & & & \\ & A_2 & & 0 & \\ & & \ddots & & \\ & 0 & & \ddots & \\ & & & & A_t \end{pmatrix}, \quad B = \begin{pmatrix} B_1 & & & & \\ & B_2 & & 0 & \\ & & \ddots & & \\ & 0 & & \ddots & \\ & & & & B_t \end{pmatrix}$$

Blockdiagonalmatrizen vom selben Typus, so ist AB wieder vom selben Typus, und für das Produkt gilt

$$AB = \begin{pmatrix} A_1B_1 & & & & \\ & A_2B_2 & & 0 & \\ & & \ddots & & \\ & 0 & & \ddots & \\ & & & & A_tB_t \end{pmatrix}$$

(Wir sagen dabei, dass A den Typus $(n_1, \ldots, n_t)$ hat, wenn A_i eine $n_i \times n_i$-Matrix ist). Dies folgt unmittelbar aus der Formel für das Matrizenprodukt. Ohne Rechnung folgt die Behauptung daraus, dass die Blockdiagonalform einer Zerlegung $K^n = K^{n_1} \oplus \ldots \oplus K^{n_t}$ entspricht, wobei K^{n_i} A-invariant ist und $A_i = A_{|\ K^{n_i}}$. Sind die K^{n_i} auch B-invariant, so auch AB-invariant, und die Behauptung ergibt sich aus der offensichtlichen Beziehung, dass $AB_{|\ K^{n_i}} = A_{|\ K^{n_i}} B_{|\ K^{n_i}}$. Insbesondere gilt für A selbst

$$A^i = \begin{pmatrix} A_1^i & & & & \\ & A_2^i & & 0 & \\ & & \ddots & & \\ & 0 & & \ddots & \\ & & & & A_t^i \end{pmatrix}.$$

Corollar 6.3 (Satz von Hamilton-Cayley) Es gilt (für beliebigen Körper K)

$$\chi_\varphi(\varphi) = 0\,,$$

d.h., das Minimalpolynom teilt das charakteristische Polynom.

Beweis Die Äquivalenz der Behauptungen folgt aus 4.4: für das Minimalpolynom $p_\varphi(x)$ gilt

$$f(\varphi) = 0 \quad \Leftrightarrow \quad p_\varphi \mid f$$

für alle $f \in K[x]$.

Ist nun K algebraisch abgeschlossen, so gilt $\chi_\varphi(\varphi) = 0$ wegen 6.1(a) und (b). Ist K beliebig, so wähle eine Matrixdarstellung $A \in M_n(K)$ für φ und eine algebraisch abgeschlossenen Körper $L \supseteq K$ (dieser existiert, wie in der Algebra gezeigt wird – für $K = \mathbb{R}$ können wir zum Beispiel $L = \mathbb{C}$ nehmen). Es genügt nun, $\chi_A(A) = 0$ zu zeigen, denn für die gewählte Basis b von V mit $M_b^b(\varphi) = A$ ist

$$\begin{array}{rrcl} M_b^b: & \mathrm{End}(V) & \to & M_n(K) \\ & \psi & \mapsto & M_b^b(\psi) \end{array}$$

ein Ringhomomorphismus. Insbesondere gilt für $f(x) \in K[x]$:

$$M_b^b(f(\varphi)) = f(M_b^b(\varphi)) = f(A)$$

und $M_b^b(\psi = 0 \Leftrightarrow \psi = 0$.

Die Aussage $\chi_A(A) = 0$ kann aber nun gezeigt werden, indem wir A als Matrix in $M_n(L)$ auffassen – nach der Definition ändert sich dabei nichts am charakteristischen Polynom.

Damit sind wir wieder im Fall eines algebraisch abgeschlossenen Körpers, wo wir die Behauptung schon bewiesen haben.

Corollar 6.4 Zerfällt das charakteristische Polynom von φ über K, so auch das Minimalpolynom.

Beweis: Dies folgt aus der zweiten Aussage in 6.3.

Insbesondere kann in allen bisher bewiesenen Aussagen die Voraussetzung "Das Minimalpolynom von φ zerfällt" durch die Voraussetzung "Das charakteristische Polynom von φ zerfällt" ersetzt werden!

Satz 6.5 Sei $K = \mathbb{C}$, oder algebraisch abgeschlossen. Für zwei Matrizen $A, A' \in M_n(K)$ sind die folgenden Aussagen äquivalent:

(a) A und A' sind ähnlich.

(b) A und A' haben (bis auf Anordnung der Jordankästchen) dieselbe Jordannormalform.

(c) Für alle $\lambda \in K$ und $m \in \mathbb{N}$ gilt

$$a(A, \lambda, m) = a(A', \lambda, m)\,.$$

Hierbei ist $a(A, \lambda, m)$ die Anzahl der Jordankästchen $J_m(\lambda)$ der Größe m zum Eigenwert λ in einer Jordannormalform von A (siehe 5.8).

(d) Für alle $\lambda \in K$ und $m \in \mathbb{N}$ gilt

$$\dim V(A, \lambda)^{(m)} = \dim V(A', \lambda)^{(m)}\,.$$

Beweis: Die Äquivalenz von (b) und (c) ist klar, und die Implikation (d) $\Rightarrow$ (c) ergibt sich aus Satz 5.8 (b). Umgekehrt ist nach dem Beweis von 5.8 (b)

$$\sum_{j=1}^{\infty} a(\lambda, j) = a(\lambda, \geq i) = \dim V(\lambda)^{(i)} - \dim V(\lambda)^{(i-1)}\,.$$

Weiter gilt

$$\sum_{i=1}^{m} (\dim V(\lambda)^{(i)} - \dim V(\lambda)^{(i-1)}) = \dim V(\lambda)^{(m)}\,.$$

Kennt man also alle $a(\lambda, m)$, so kennt man auch alle $\dim V(\lambda)^{(m)}$. Dies zeigt (c) $\Rightarrow$ (d).

(a) $\Rightarrow$ (b): Zunächst macht (b) Sinn, da die Jordannormalform bis auf Anordnung der Kästchen eindeutig ist. Sei nun $A' = C^{-1}AC$ für eine invertierbare Matrix $C \in M_n(K)$. Seien $B, B' \in M_n(K)$ invertierbar so dass

$$J = B^{-1}AB \quad \text{und} \quad J' = (B')^{-1}A'B^{-1}$$

Jordannormalform haben. Dann ist $J' = (CB')^{-1}A(CB')$ auch eine Jordannormalform von A, nach Satz 5.7 also J und J' bis auf Anordnung der Kästchen gleich.

(b) $\Rightarrow$ (a): Seien $B, B' \in M_n(K)$ invertierbar so dass

$$J = B^{-1}AB \quad \text{und} \quad J' = (B')^{-1}A'B'$$

Jordannormalform haben, wo die Jordankästchen bis auf Anordnung gleich sind. Es ist also

$$J = \begin{pmatrix} J_1 & & & \\ & \ddots & 0 & \\ & 0 & \ddots & \\ & & & J_r \end{pmatrix}$$

mit Jordankästchen J_α und

$$J' = \begin{pmatrix} J_{\sigma(1)} & & & \\ & J_{\sigma(2)} & & \\ & & \ddots & \\ & & & J_{\sigma(r)} \end{pmatrix}$$

für eine Permutation σ aus der symmetrischen Gruppe S_r. Sei $J_i \in M_{m_i}(K)$, und seien

$$b_1^1, \dots, b_{m_1}^1, b_1^2, \dots, b_{m_2}^2, \dots, b_1^r, \dots, b_{m_r}^r$$

die Spalten von B. Diese bilden also die Basis

$$b = (b(1), b(2), \dots, b(r)) \,,$$

$b(i) = (b_1^i, \dots, b_{m_i}^i)$, bezüglich derer A die Jordannormalform J hat, und setzen wir $V_i =< b_1^i, \dots, b_{m_i}^i >_K$, so ist J_i die Matrixdarstellung von $A_{| V_i}$ bezüglich $b(i) : J_i = M_{b(i)}^{b(i)}(A_{| V_i})$. Sei

$$\tilde{b} = (b(\sigma(1)), b(\sigma(2)), \dots, b(\sigma(r)))$$

die Basis mit denselben Vektoren, aber einer anderen Anordnung: die Blöcke $b(i)$ werden gemäß der Permutation σ permutiert. Dann ist offenbar die Matrixdarstellung von A bezüglich $\tilde{b}$ gleich J', denn es ist $J_{\sigma(i)} = M_{b(\sigma(i))}^{b(\sigma(i))}(A_{| V_{\sigma(i)}})$.

Dies zeigt, dass J und J' konjugiert sind – genauer ist $J' = C^{-1}JC$ für die Basiswechselmatrix $C = M_{\tilde{b}}^b$ $(B = M_b^e, J' = (M_{\tilde{b}}^e)^{-1} A M_{\tilde{b}}^e = (M_{\tilde{b}}^b)^{-1}(M_b^e)^{-1} A M_b^e M_{\tilde{b}}^b = C^{-1}JC)$. Damit sind auch A und A' konjugiert $(A' = B'J'(B')^{-1} = (BC(B')^{-1})^{-1} ABC(B')^{-1})$.

Beispiel 6.6 Nach Satz 5.8 braucht man nur eine Teilinformation, um zu wissen, wie die Jordannormalform aussehen muss (insbesondere muss man hierfür keine Jordanbasis konstruieren). Betrachte zum Beispiel

$$A = \begin{pmatrix} -5 & 2 & 0 \\ -8 & 3 & 0 \\ -2 & 1 & -1 \end{pmatrix} \in M_3(\mathbb{R}) \,.$$

Es ist $\chi_A(x) = (x+1)^3$ (siehe Übungsaufgabe 14 (ii)), und für

$$N = A - (-1)E = A + E = \begin{pmatrix} -4 & 2 & 0 \\ -8 & 4 & 0 \\ -2 & 1 & 0 \end{pmatrix}$$

gilt $\mathrm{rg}N = 1$ und damit $\dim V(-1) = \dim\ker N = 3 - 1 = 2$ und weiter

$$N^2 = \begin{pmatrix} -4 & 2 & 0 \\ -8 & 4 & 0 \\ -2 & 1 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

also $V(-1)^{(2)} = \mathbb{R}^3 \quad (= V[-1])$. Mit der Formel in Satz 5.8 (b) folgt nun leicht, dass die Jordannormalform von A gleich

$$\left(\begin{array}{cc|c} -1 & 1 & \\ & -1 & \\ \hline & & -1 \end{array}\right)$$

sein muss. In der Praxis (und wenn man die Formel vergessen hat) schließt man oft so: Die möglichen Jordannormalformen von A sind

$$J = \begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix}, \quad J' = \begin{pmatrix} -1 & 1 & \\ & -1 & \\ & & -1 \end{pmatrix}, \quad J'' = \begin{pmatrix} -1 & 1 & \\ & -1 & 1 \\ & & -1 \end{pmatrix}.$$

Die erste Form (diagonalisierbarer Fall) ist nicht möglich, da dies ein Widerspruch zu $A + E = N \neq 0$ wäre: Ist $J = B^{-1}AB$ mit invertierbarer Matrix B, so gilt für $m \in \mathbb{N}_0$ und $\lambda \in K$

$$(J - \lambda E)^m = 0 \quad \Leftrightarrow \quad (A - \lambda E)^m = 0\,,$$

denn es ist $(A - \lambda E)^m = 0 \quad \Leftrightarrow \quad B^{-1}(A + \lambda E)^m B = 0$, und

$$\begin{aligned} & B^{-1}(A - \lambda E)^m B \\ = \ & B^{-1}(A - \lambda E)(A - \lambda E)\dots(A - \lambda E)B \\ = \ & B^{-1}(A - \lambda E)BB^{-1}(A - \lambda E)B \dots B^{-1}(A - \lambda E)B \quad (m - \text{mal}) \\ = \ & (B^{-1}(A - \lambda E)B)^m = (J - \lambda E)^m\,. \end{aligned}$$

Im ersten Fall ist aber $J + E = 0$, und es würde $A + E = 0$ folgen – Widerspruch!

Der dritte Fall ist ebenfalls ausgeschlossen, denn dann wäre $(A+E)^2 \neq 0$, weil $(J+E)^2 \neq 0$ (siehe Beweis von Satz 5.8: der Nilpotenzgrad von $J_m(\lambda) - \lambda E$ ist m). Also bleibt nur noch der zweite Fall.

Das Verfahren, eine tatsächliche Jordanbasis zu finden (eine Basis, in der der Endomorphismus Jordannormalform hat) ist komplizierter, aber immer noch **konstruktiv**.

Konstruktion 6.7 (Bestimmung einer Jordanbasis)

1) Berechne das charakteristische Polynom $\chi_\varphi(x)$. Dies liefert die Eigenwerte $\lambda_1, \dots, \lambda_r$ und ihre algebraischen Multiplizitäten $m_1, \dots, m_r$.

2) Für jeden Eigenwert λ_i berechne die Dimension der verallgemeinerten Eigenräume. Dies liefert nach 5.8 schon, wie die Jordannormalform aussehen muss. Außerdem liefert es die Stufe n_i des nilpotenten Endomorphismus $N_i = \varphi - \lambda_{i \mid V[\lambda_i]}$. Berechne eine Basis von

$$V[\lambda_i] = \ker(\varphi - \lambda_i)^{n_i}\,.$$

3) Betrachte nun für jeden Eigenwert $\lambda = \lambda_i$ den nilpotenten Endomorphismus $N = (\varphi - \lambda)_{|\, V[\lambda]} : V[\lambda] \to V[\lambda]$ (am besten in der Basisdarstellung zur gefundenen Basis von $V[\lambda]$) und finde einen Vektor $w_1 \in V[\lambda]$ mit

$$N^{n-1} w_1 \neq 0\,,$$

wobei n $(= n_i)$ der Nilpotenzgrad von N ist.

4) Die Vektoren $N^{n-1}w_1, N^{n-2}w_1, \ldots, Nw_1, w_1$ sind linear unabhängig (5.2); ergänze sie zu einer Basis von $V[\lambda]$. In dieser Basis sieht N so aus

$$A_1 = \begin{array}{r} n^{(1)} := n \left\{ \vphantom{\begin{matrix} 0\\ \ddots \\ 0 \\ 0 \end{matrix}} \right. \\ \vphantom{\begin{matrix} 0\\0\\0\\0 \end{matrix}} \end{array} \left(\begin{array}{cccc|c} 0 & 1 & & & \\ & \ddots & \ddots & & * \\ & 0 & \ddots & 1 & \\ & & & 0 & \\ \hline & & & & \\ & & 0 & & A_2 \\ & & & & \end{array} \right)$$

Dann ist A_2 wieder nilpotent. Mache mit A_2 weiter (Stufe $n^{(2)}$ von A_2, w_2 zu A_2 mit Periode $n^{(2)}$ usw.) und bringe A_2 auf die Gestalt

$$\begin{array}{r} n^{(2)} \left\{ \vphantom{\begin{matrix} 0\\ \ddots \\ 0 \\ 0 \end{matrix}} \right. \\ \vphantom{\begin{matrix} 0\\0\\0\\0 \end{matrix}} \end{array} \left(\begin{array}{cccc|c} 0 & 1 & & & \\ & \ddots & \ddots & & * \\ & 0 & \ddots & 1 & \\ & & & 0 & \\ \hline & & & & \\ & & 0 & & A_3 \\ & & & & \end{array} \right)$$

5) Ändere nun w_2 durch ein Element aus $L_N(w_1) =< N^{n-1}w_1, \ldots, w_1 >$ ab, dass

$$N^{n^{(2)}} w_2 = 0\,.$$

Dies ist immer möglich: Nach der Form der erhaltenen Matrix ist

$$N^{n^{(2)}} w_2 = a_0 w_1 + a_1 N w_2 + \ldots + a_{n-1} N^{n-1} w_2 \in L_N(w_1)$$

Damit schließen wir:
$0 = N^n w_2 = N^{n-n^{(2)}}(N^{n^{(2)}} w_1) = a_0 N^{n-n^{(2)}} w_1 + \ldots + N^{n-1} a_{n^{(2)}-1} w_1$

$\Rightarrow a_0 = \ldots = a_{n^{(2)}-1} = 0$ (lineare Unabhängigkeit der $N^\nu w_1$)
$\Rightarrow N^{n^{(2)}} w_2 = N^{n^{(2)}} \tilde{w}$ mit $\tilde{w} = a_{n^{(2)}} w_1 + a_{n^{(2)}+1} N w_1 + \ldots \in L_N(w_1)$
$\Rightarrow N^{n^{(2)}}(w_2 - \tilde{w}) = 0$.

Wir können also w_2 durch $w_2^{\text{neu}} = w_2 - \tilde{w}$ ersetzen. Dann erreichen wir die Gestalt

$$
\begin{matrix}
n^{(1)}\left\{\vphantom{\begin{matrix}0\\ \ddots\\ 0\\ 0\end{matrix}}\right. \\
n^{(2)}\left\{\vphantom{\begin{matrix}0\\ \ddots\\ 0\\ 0\end{matrix}}\right. \\
\vphantom{\begin{matrix}\\ \\ A_3\\ \\ \end{matrix}}
\end{matrix}
\begin{pmatrix}
\boxed{\begin{matrix}0 & 1 & & \\ & \ddots & \ddots & \\ & 0 & \ddots & 1 \\ & & & 0\end{matrix}} & 0 & * \\
& \boxed{\begin{matrix}0 & 1 & & \\ & \ddots & \ddots & \\ & 0 & \ddots & 1 \\ & & & 0\end{matrix}} & \\
0 & & \boxed{\begin{matrix}\\ \quad A_3 \quad \\ \\ \end{matrix}}
\end{pmatrix}
$$

Fahre nun mit A_3 entsprechend fort.

Dies muss man für jeden verallgemeinerten Eigenraum $V[\lambda_i]$ machen.

Beispiel 6.8 Betrachte die Matrix

$$A = \begin{pmatrix} -5 & 2 & 0 \\ -8 & 3 & 0 \\ -2 & 1 & -1 \end{pmatrix} \in M_3(\mathbb{R}),$$

mit dreifachem Eigenwert -1 (aus Beispiel 6.6). Es ist also $V = \mathbb{R}^3 = V[-1]$. Für $N = A + E$ gilt $N \neq 0\,, N^2 = 0$, also ist die Nilpotenzstufe $n = 2$. Für

$$N = \begin{pmatrix} -4 & 2 & 0 \\ -8 & 4 & 0 \\ -2 & 1 & 0 \end{pmatrix}$$

ist $w_1 = e_2$ ein Vektor mit $Nw_1 = \begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix} =: v \neq 0$ und $L_N(w_1) =< v, e_2 >_{\mathbb{R}}$. Ergänze (v, e_2) durch e_1 zu einer Basis (v, e_2, e_1) von $\mathbb{R}^3$. In dieser Basis hat N die Form

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

wegen $Nv = 0$, $Ne_2 = v$ und

$$Ne_1 = \begin{pmatrix} -4 \\ -8 \\ -2 \end{pmatrix} = -2v\,.$$

Es ist $-2v = -2Ne_2 = N(-2e_2)$. Dann ist

$$w_2 = e_1 + 2e_2 \neq 0$$

mit $Nw_2 = 0$. In der Basis

$$Nw_1 = v = \begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix}, w_1 = e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, w_2 = e_1 + 2e_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$$

hat nun A Jordannormalform

$$\left(\begin{array}{cc|c} -1 & 1 & \\ & -1 & \\ \hline & & -1 \end{array}\right)$$

(vergleiche Beispiel 6.7).

§7 Familien und kartesische Produkte

Familien

Wie "indiziert" man mathematisch korrekt und allgemein?

$$x_1, \dots, x_n \quad ; \quad \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} ; \quad x_i \ (i \in I)$$

Definition 7.1 Sei M eine Menge und I eine weitere Menge. Eine **Familie in M über I** (oder **mit Indexmenge I**) ist eine Abbildung

$$a : I \to M\,.$$

Statt $a(i)$ schreiben wir auch a_i. Weiter schreiben wir auch oft $(a_i)_{i\in I}$ für die Familie. Die Menge aller Familien in M über I wird mit M^I bezeichnet. Es ist also $M^N := \mathrm{Abb}(N, M)$ für Mengen M und N.

Beispiel 7.2 Für die folgenden Spezialfälle gibt es eigene Namen:

(a) Ist $I = \mathbb{N}$, so spricht man von einer (unendlichen) Folge. Zum Beispiel hat man die Folge

$$1, 1, 2, 3, 5, 8, 13, \dots$$

der Fibonacci-Zahlen, genau definiert als die Folge

$$(a_n)_{n\in\mathbb{N}} \quad \text{mit} \quad a_1 = 1, a_2 = 1, a_{n+1} = a_{n-1} + a_n\,.$$

(b) Für $I = \underline{n} := \{1, 2, \ldots, n\}$ heißt die Familie ein n-Tupel, und man schreibt $(a_1, \ldots, a_n)$. Damit ist $M^{\underline{n}} = M^n$, wie in LA I definiert. Speziell sagt man Paar für $n = 2$, Tripel für $n = 3$, Quadrupel für $n = 4$ usw.

Gleichheit von Familien ist die Gleicheit von Abbildungen. Für zwei Familien $(a_i)_{i \in I}, (b_i)_{i \in I}$ in M gilt also

$$(a_i)_{i \in I} = (b_i)_{i \in I} \quad \Leftrightarrow \quad \underset{i \in I}{\forall} \, a_i = b_i \,.$$

Die zwei Tripel von natürlichen Zahlen $(1, 3, 4)$ und $(1, 4, 3)$ sind also verschieden.

Beachte: Man muß zwischen einer Familie $(a_i)_{i \in I}$ (manchmal auch geschrieben als $(a_i \mid i \in I)$) und der assoziierten Menge $\{a_i \mid i \in I\} \subseteq M$ unterscheiden (= Unterschied zwischen Abbildung und Bildmenge). Im obigen Beispiel ist die assoziierte Menge gleich für beide Tupel:

$$\{1, 3, 4\} = \{1, 4, 3\} \,.$$

Definition 7.3 Sei $(M_i)_{i \in I}$ eine Familie von Mengen. Die Vereinigung dieser Familie ist definiert als

$$\bigcup_{i \in I} M_i := \{x \mid \underset{i \in I}{\exists} \, x \in M_i\} \,.$$

Der Durchschnitt der Familie ist definiert als

$$\bigcap_{i \in I} M_i := \{x \mid \underset{i \in I}{\forall} \, x \in M_i\} \,.$$

kartesische Produkte

Wir wollen auch "Familien mit Einträgen in verschiedenen Mengen" betrachten, also z.B. Paare (a, b) mit $a \in M$ und $b \in N$, oder "Ausdrücke" $(a_i)_{i \in I}$ mit $a_i \in M_i$ für eine Familie $(M_i)_{i \in I}$ von Mengen.

Definition 7.4 Sei $(M_i)_{i \in I}$ eine Familie von Mengen. Unter dem **(kartesischen) Produkt** dieser Familie verstehen wir die Menge

$$\prod_{i \in I} M_i := \{\text{Familien } a : I \to \bigcup_{i \in I} M_i \mid \underset{i \in I}{\forall} \, a_i \in M_i\} \,.$$

Für n-Tupel $(M_1, \ldots, M_n)$ schreiben wir auch

$$M_1 \times \ldots \times M_n \quad \text{oder} \quad \prod_{i=1}^{n} M_i \,.$$

Dies ergibt dieselbe Definition wie in LA I.

Beispiele 7.4 (a) Insbesondere erhalten wir für zwei Mengen M und N – also das Paar (M, N) – wieder das (kartesische) Produkt

$$M \times N \,,$$

dessen Elemente wir als Paare (m, n) mit $m \in M$ und $n \in N$ schreiben.

(b) Für die konstante Familie $(M)_{i\in I}$ ist offenbar gerade $\prod\limits_{i\in I} M = M^I$.

§8 Relationen

Definition 8.1 Seien M und N Mengen. Eine **Relation von M nach N** ist eine Teilmenge $R \subseteq M \times N$. Statt $(x, y) \in R$ schreibt man oft xRy (x steht in der Relation R zu y). Ist $M = N$, also $R \subseteq M \times M$, so spricht man auch von einer **Relation auf M**.

Beispiel 8.2 Ist $f : M \to N$ eine Abbildung, so ist der **Graph von f**,

$$\Gamma = \Gamma_f := \{(x, y) \in M \times N \mid y = f(x)\}$$

eine Relation von M nach N. Es ist dann $x\Gamma y \;\Leftrightarrow\; y = f(x)$.

Nicht jede Relation ist von dieser Form, zum Beispiel nicht die Relation $K = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ (der Einheitskreis):

denn zu einem vorgegebenen $x \in]-1, 1[$ gibt es mehrere y mit $(x, y) \in K$; das kann bei dem Graph einer Funktion nicht passieren, da der Funktionswert $f(x)$ eindeutig ist. Man kann also Relationen von M nach N als "mehrdeutige Funktionen" auffassen.

Oft verwenden wir andere Symbole für eine Relation, zum Beispiel $x \sim y, x \leq y, x \,\square\, y \ldots$

Definition 8.3 Eine Relation R auf einer Menge M heißt

(a) **reflexiv**, wenn xRx für alle $x \in M$,

(b) **transitiv**, wenn für alle $x, y, z \in M$ gilt

$$xRy \wedge yRz \Rightarrow xRz\,,$$

(c) **symmetrisch**, wenn für alle $x, y \in M$ gilt

$$xRy \Rightarrow yRx\,,$$

(d) **antisymmetrisch**, wenn für alle $x, y \in M$ gilt

$$xRy \text{ und } yRx \Rightarrow x = y\,,$$

(e) **Äquivalenzrelation**, wenn sie reflex, transitiv und symmetrisch ist, und

(f) **Ordnungsrelation** (oder Ordnung auf M), wenn sie reflexiv, transitiv und antisymmetrisch ist.

Bemerkungen 8.4 (a) Für Ordnungsrelationen benutzt man gern Zeichen wie $\leq, <$ usw.

(b) Für Äquivalenzrelationen benutzt man gern Symbole wie $\sim, \equiv$ usw.

Beispiele 8.5 (a) Die Gleicheit ist eine Äquivalenzrelation.

(b) Sei $f : M \to N$ eine Abbildung. Definiere eine Relation $\sim_f$ auf M (!) durch

$$x \sim_f y \quad \Leftrightarrow \quad f(x) = f(y)\,.$$

Dann ist $\sim_f$ eine Äquivalenzrelation:

$$\begin{array}{rcl} f(x) & = & f(x) \\ f(x) & = & f(y) \wedge f(y) = f(z) \Rightarrow f(x) = f(z) \\ f(x) & = & f(y) \Rightarrow f(y) = f(x)\,. \end{array}$$

(c) Die übliche $\leq$-Relation auf $\mathbb{R}$ ist eine Ordnungsrelation (und keine Äquivalenzrelation).

(d) Sei X eine Menge. Auf der Menge $\mathfrak{P}(X)$ aller Teilmengen von X (Potenzmenge von X) ist die Inklusion $\subseteq$ eine Ordnungsrelation.

(e) Die Teilbarkeit $\mid$ auf $\mathbb{Z}$ ($a \mid b \;\Leftrightarrow\; \exists n \in \mathbb{Z}$ mit $a \cdot n = b$) ist keine Äquivalenzrelation ($3 \mid 6$ aber $6 \nmid 3$) und keine Ordnungsrelation ($3 \mid -3$ und $-3 \mid 3$, aber $3 \neq -3$). Aber die Einschränkung auf $\mathbb{N}$ ist eine Ordnungsrelation.

(f) Sei $a \in \mathbb{Z}, a \neq 0$. Für $x, y \in \mathbb{Z}$ definiere

$$x \equiv y \pmod a \quad :\Leftrightarrow \quad a \mid x - y$$

(gesprochen: x kongruent (zu) y modulo a).

Dies liefert eine Äquivalenzrelation auf $\mathbb{Z}$:

1) $a \cdot 0 = 0 \quad \Rightarrow \quad a \mid 0 = x - x \quad \Rightarrow \quad x \equiv x (\operatorname{mod} a)$.

2) $x \equiv y \ (\operatorname{mod} a)$ und $y \equiv z \ (\operatorname{mod} a)$
$\Rightarrow a \mid x - y$ und $a \mid y - z$
$\Rightarrow \ \exists\, m, n \in \mathbb{Z} : am = x - y, an = y - z$
$\Rightarrow x - z = x - y + y - z = am + an = a(m + n)$
$\Rightarrow x \equiv z \ (\operatorname{mod} a)$.

3) $x \equiv z \ (\operatorname{mod} a)$
$\Rightarrow a \mid x - y$
$\Rightarrow a \mid -(x - y) = y - x$
$\Rightarrow y \equiv x \ (\operatorname{mod} a)$.

Sei zum Beispiel $a = 4$. Dann ist für die Kongruenz modulo 4

$$7 \equiv 3, \quad 14 \equiv 2, \quad \text{sowie } 2 \not\equiv 3\,.$$

§9 Äquivalenzrelationen

Definition 9.1 Sei $\sim$ eine Äquivalenzrelation auf einer Menge M. Für $x \in M$ heißt die Menge

$$\overline{x} := \{y \in M \mid y \sim x\} \subseteq M$$

die **Äquivalenzklasse von** x **bezüglich** $\sim$ (oder **modulo** $\sim$). Die Menge

$$M/_{\sim} := \{\overline{x} \mid x \in M\}$$

aller Äquivalenzklassen modulo $\sim$ heißt die **Quotientenmenge von** M **nach** $\sim$ (oder **modulo** $\sim$).

Die Abbildung

$$\begin{aligned} \pi : M &\rightarrow M/_{\sim} \\ x &\mapsto \overline{x} \end{aligned}$$

heißt die **Quotientenabbildung** zu $\sim$.

Bemerkung 9.2 Die obige Bezeichnung $\overline{x}$ ist üblich und praktisch, macht aber nicht deutlich, dass die Äquivalenzklasse von $\sim$ abhängt. Deswegen schreibt man manchmal $\overline{x}^{\sim}$. Andere Bezeichnungen für Äquivalenzklassen: $[x], [x]_{\sim}$.

Lemma 9.3 Sei $\sim$ eine Äquivalenzrelation auf einer Menge M. Seien $x, y \in M$ und $\overline{x}, \overline{y}$ die zugehörigen Äquivalenzklassen modulo $\sim$. Dann gilt

$$y \sim x \quad \Leftrightarrow \quad y \in \overline{x} \quad \Leftrightarrow \quad \overline{y} = \overline{x}\,.$$

Beweis Die erste Äquivalenz gilt nach Definition. Wir zeigen nun noch

$$y \sim x \Leftrightarrow \overline{y} = \overline{x}$$

"$\Rightarrow$": Es sei $y \sim x$. Dann gilt

$$z \in \overline{y} \overset{\text{Def.}}{\Rightarrow} z \sim y \overset{\text{Trans.}}{\Rightarrow} z \sim x \overset{\text{Def.}}{\Rightarrow} z \in \overline{x}\,,$$

also $\overline{y} \subseteq \overline{x}$. Da auch $x \sim y$ gilt (Symmetrie) folgt auch $\overline{x} \subseteq \overline{y}$.

"$\Leftarrow$" Sei $\overline{y} = \overline{x}$. Da $y \in \overline{y}$ (Reflexivität), folgt $y \in \overline{x}$, also $y \sim x$.

Corollar 9.4 Sei

$$\pi : M \rightarrow M/_{\sim}$$

die Quotientenabbildung zu $\sim$, und sei $\sim_\pi$ die zu π gehörige Äquivalenzrelation (siehe 8.5 (b)), also

$$x \sim_\pi y \;\Leftrightarrow\; \pi(x) = \pi(y)\,.$$

Dann ist $\sim_\pi = \sim$. Mit anderen Worten: jede Äquivalenzrelation kann durch die Konstruktion 8.5 (b) erhalten werden.

Beweis $\pi(x) = \pi(y) \overset{\text{Def.}}{\Leftrightarrow} \overline{x} = \overline{y} \overset{9.3}{\Leftrightarrow} x \sim y$.

Wir geben nun eine weitere Konstruktion von Äquivalenzrelationen an.

Definition 9.5 Sei M eine Menge, und sei P eine Menge von Teilmengen von M (also $P \subseteq \mathfrak{P}(M)$). P heißt **Partition** (oder **Klasseneinteilung**) von M, wenn gilt

(a) $\emptyset \notin P$

(b) Die Mengen in P sind paarweise disjunkt, d.h., für $U, V \in P$ mit $U \neq V$ gilt $U \cap V = \emptyset$.

(c) M wird durch die Mengen in P überdeckt, d.h., $\bigcup\limits_{U \in P} U = M$.

Satz 9.6 Sei M eine Menge.

(a) Sei $\sim$ eine Äquivalenzrelation auf M. Dann bilden die Äquivalenzklassen modulo $\sim$ eine Partition von M.

(b) Umgekehrt sei P eine Partition von M. Dann gibt es eine eindeutig bestimmte Äquivalenzrelation $\sim$ auf M, so dass P gerade aus den Äquivalenzklassen modulo $\sim$ besteht.

Beweis (a): Wegen $x \in \overline{x}$ ist $\overline{x}$ nichtleer. Ist weiter $\overline{x} \cap \overline{y} \neq \emptyset$, etwa $z \in \overline{x} \cap \overline{y}$, so ist $z \in \overline{x}$ und $z \in \overline{y}$ also $z \sim x$ und $z \sim y$, also $x \sim y$, nach 9.3 also $\overline{x} = \overline{y}$. Schließlich liegt trivialerweise jedes $x \in M$ in einer Äquivalenzklasse, nämlich in $\overline{x}$.

(b): Für die Äquivalenzrelation muß offenbar gelten

$$(*) \qquad x \sim y \quad \Leftrightarrow \quad \exists U \in P : x \in U \wedge y \in U\,.$$

Wir nehmen dies als Definition von $\sim$, das gibt eine wohldefinierte Relation auf M.

Behauptung: $\sim$ ist eine Äquivalenzrelation

Beweis Reflexivität: Sei $x \in M$. Wegen $\bigcup\limits_{U \in P} U = M$ existiert ein $U \in P$ mit $x \in U$. Damit ist $x \sim x$.

Transivität: Sei $x \sim y$ und $y \sim z$. Dann existiert ein $U \in P$ mit $x \in U$ und $y \in U$, und es existiert ein $V \in P$ mit $y \in V$ und $z \in V$. Es folgt $y \in U \cap V$ und damit $U = V$, da P eine Partition ist. Folglich gilt $x \sim z$.

Symmetrie: ist klar.

Schließlich besteht P gerade aus den Äquivalenzklassen bezüglich $\sim$:

1) Sei $U \in P$. Wegen $U \neq \emptyset$ gibt es ein $x \in U$.

Behauptung $U = \overline{x}$

Beweis $y \in \overline{x} \overset{\text{Def.}}{\Rightarrow} y \sim x \overset{\text{Def.}}{\Rightarrow} \exists V \in P : y \in V \wedge x \in V \overset{x \in U}{\Rightarrow} U = V$ und damit $y \in U$. Also gilt $\overline{x} \subseteq U$.

Umgekehrt gilt: $y \in U \Rightarrow y \in U \wedge x \in U \overset{\text{Def.}}{\Rightarrow} y \sim x \overset{\text{Def.}}{\Rightarrow} y \in \overline{x}$. Also gilt auch $U \subseteq \overline{x}$.

2) Sei $x \in M$ Wegen $\bigcup\limits_{U \in P} U = M$ gibt es dann ein $U \in P$ mit $x \in U$. Nach der in 1) bewiesenen Behauptung gilt dann $U = \overline{x}$.

Dass $\sim$ eindeutig ist, haben wir schon bemerkt $(*)$.

Das Denken in Äquivalenzklassen ist ganz natürlich:

Beispiele 9.7 (a) Auf der Menge aller Menschen definiere eine Äquivalenzrelation

$$x \sim y \;:\Leftrightarrow\; x \text{ und } y \text{ haben dasselbe Geschlecht}$$

Es gibt genau zwei Äquivalenzklassen: {Frauen}, {Männer}. Wir können diese Äquivalenzrelation durch die folgende Abbildung erhalten:

$$\begin{array}{rcl} \{\text{Menschen}\} & \to & \text{Menge der Geschlechter} = \{\text{weiblich, männlich}\} \\ x & \mapsto & \text{Geschlecht von } x. \end{array}$$

(b) Auf der Menge aller Schüler definiere die Äquivalenzrelation

$$x \sim y \quad \Leftrightarrow \quad x \text{ und } y \text{ gehen in dieselbe Schule}\,.$$

Wir können dies durch die folgende Abbildung erhalten

$$\begin{array}{rcl} \{\text{Schüler}\} & \to & \{\text{Schulen}\} \\ x & \mapsto & \text{Schule von } x\,. \end{array}$$

(c) Also: jede Einführung von Klassen im täglichen Leben

Menschen —— Berufe
Menschen —— Religionen
Menschen —— Parteizugehörigkeit

Obst —— Sorten

Tiere —— Gattungen —— Arten

liefert eine Klasseneinteilung im Sinne der Mathematik und entsprechend auch eine Äquivalenzrelation, d.h., eine "Gleichsetzung" bzw. "Vergröberung" (wie "die Säugetiere").

Definition 9.8 Sei $\sim$ eine Äquivalenzrelation auf einer Menge M, und sei c eine Äquivalenzklasse bezüglich $\sim$. Ein Element $x \in M$ heißt **Repräsentant** von c, wenn $x \in c$, also wenn $\overline{x} = c$.

Beispiel 9.10 Auf $\mathbb{Z}$ betrachte die Kongruenz modulo 4 (vergleiche 8.5 (f)), also die Äquivalenzrelation

$$x \equiv y \pmod 4 \quad \Leftrightarrow \quad x - y \text{ ist durch 4 teilbar}.$$

Jede ganze Zahl ist kongruent zu $0, 1, 2$ oder 3, (da $x \pm 4 \sim x$), und diese Zahlen sind paarweise inkongruent modulo 4. Also ist

$$\mathbb{Z}/_{\equiv (\mathrm{mod}\ 4)} = \{\overline{0}, \overline{1}, \overline{2}, \overline{3}\},$$

besteht also aus 4 Elementen. Allgemeiner haben wir für $a \in \mathbb{N}$ genau a Äquivalenzklassen modulo a (d.h., bezüglich der Kongruenz modulo a), nämlich $\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{a-1}$.

Anders ausgedrückt: jede Äquivalenzklasse modulo a hat einen eindeutig bestimmten Repräsentanten in $\{0, \ldots, a-1\}$, nämlich den kleinsten nicht-negativen Repräsentanten. Diesen erhält man als Rest beim Teilen durch a; deswegen heißt die Äquivalenzklasse modulo a auch **Restklasse modulo** a.

Für $m \in \mathbb{Z}$ wird die Restklasse $\overline{m}$ modulo a auch mit $m \pmod a$ bezeichnet.

§10 Quotientengruppen und Quotientenräume

Sei $(A, +)$ eine *kommutative* (abelsche) Gruppe

Lemma/Definition 10.1 Sei $B \subseteq A$ eine Untergruppe. Definiere die Relation auf A

$$a \equiv a' \mod B \quad :\Leftrightarrow \quad a - a' \in B.$$

Wir sagen hierzu "a kongruent (zu) a' modulo B" und schreiben auch $a \equiv_B a'$. Dies ist eine Äquivalenzrelation auf A.

Beweis: Schreibe $\equiv$ für $\equiv_B$.

Reflexivität: $a \equiv a$, da $a - a = 0 \in B$.

Transivität: $a \equiv b$ und $b \equiv c \;\Rightarrow\; a - b, b - c \in B \;\Rightarrow\; a - c = (a - b) + (b - c) \in B$.

Symmetrie: $a \equiv b \;\Rightarrow\; a - b \in B \;\Rightarrow\; b - a = -(a - b) \in B \;\Rightarrow\; b \equiv a$.

Lemma 10.2 Die Äquivalenzklasse von $a \in A$ bezüglich $\equiv_B$ ist die Menge

$$\overline{a} = a + B \;:=\; \{a + b \mid b \in B\}.$$

Diese heißt die **Nebenklasse von** a **modulo** B.

Beweis "$\subseteq$": $a' \equiv_B a \;\Rightarrow\; a' - a \in B$, etwa $a' - a = b \in B \;\Rightarrow\; a' = a + b \in a + B$.
"$\supseteq$": $a' = a + b \in a + B \quad (b \in B) \;\Rightarrow\; a' - a = b \in B \;\Rightarrow\; a' \equiv_B a$.

Bemerkung 10.3 Nach 9.3 gilt

$$a \equiv_B a' \;\Leftrightarrow\; a \in a' + B \;\Leftrightarrow\; a + B = a' + B.$$

Definition 10.4 Die Menge der Äquivalenzklassen bezüglich $\equiv_B$ wird mit A/B bezeichnet. Es ist also

$$A/B = \{a + B \mid a \in A\}$$

die Menge der Nebenklassen von A modulo B.

Beispiel 10.5 Sei $A = \mathbb{Z}$ und $B = a\mathbb{Z}$ für $a \neq 0$. Dann sehen wir: $m \in B \Leftrightarrow m = na$ für ein $n \in \mathbb{Z} \Leftrightarrow a \mid n$. Daher gilt

$$a \equiv a' \mod a\mathbb{Z} \quad \Leftrightarrow \quad a \equiv a' (\operatorname{mod} a)\,,$$

und es ist

$$\mathbb{Z}/a\mathbb{Z} = \mathbb{Z}/ \equiv (\operatorname{mod} a)\,.$$

Satz 10.6 Sei B eine Untergruppe von A. Dann gibt es genau eine Gruppenstruktur auf A/B derart, dass die Quotientenabbildung

$$\begin{aligned} \pi : A &\to A/B \\ a &\mapsto \overline{a} = a + B \end{aligned}$$

ein Gruppenhomomorphismus ist. Diese ist gegeben durch die Verknüpfung

$$(*) \qquad \overline{a} + \overline{a'} = \overline{a + a'}\,.$$

Die so erhaltene Gruppe A/B heißt die **Quotientengruppe** (oder **Faktorgruppe) von** A **modulo** B (oder **nach** B). Sie ist wieder kommutativ. Die Surjektion π heißt der **kanonische Epimorphismus**.

Beweis Ist π ein Gruppenhomomorphismus, so muss gelten $\pi(a + a') = \pi(a) + \pi(a')$. Dies bedeutet aber gerade $(*)$. Wir haben also zu zeigen, dass diese Verknüpfung $(*)$ wohldefiniert ist, d.h., nicht von der Auswahl der Repräsentanten abhängt. Sei $\overline{a} = \overline{c}$ und $\overline{a'} = \overline{c'}$. Dann ist $a - c \in B$ und $a' - c' \in B$. Es folgt $a - c + a' - c' \in B$ und damit

$$a + a' - (c + c') \in B$$

(Hier wurde die Kommutativität von A benutzt). Es folgt $\overline{a + a'} = \overline{c + c'}$.

Weiter ist noch zu zeigen, dass $(*)$ die Gruppenaxiome erfüllt. Dies folgt aber sofort aus den Gruppenaxiomen für A. Zum Beispiel ist $\overline{0}$ das neutrale Element von A/B, wenn $0 \in A$ das neutrale Element von A ist, denn es ist $\overline{a} + \overline{0} = \overline{a + 0} = \overline{a}$. Das Inverse von $\overline{a}$ ist $\overline{-a}$. Schließlich ist A/B kommutativ: $\overline{a} + \overline{a'} = \overline{a + a'} = \overline{a' + a} = \overline{a'} + \overline{a}$.

Beispiel 10.7 Damit wird für jedes $a \in \mathbb{N}$

$$\mathbb{Z}/a\mathbb{Z} = \{\overline{0}, \overline{1}, \ldots, \overline{a-1}\}$$

eine abelsche Gruppe mit a Elementen. Zum Beispiel sei $a = 5$. Dann ist

$$\overline{2} + \overline{4} = \overline{6} = \overline{1} \quad \text{und} \quad \overline{2} + \overline{3} = \overline{5} = \overline{0},\ \text{also}\ -\overline{2} = \overline{3}\,.$$

Für $a = 0$ ist $0 \cdot \mathbb{Z} = \{0\}$, und die Relation $\equiv_{\{0\}}$ ist die Gleichheit. Jede Äquivalenzklasse besteht daher nur aus einem Element. Der kanonische Epimorphismus $\mathbb{Z} \to \mathbb{Z}/\{0\}$ ist also ein Isomorphismus.

Wir haben schon gesehen, dass für einen Homomorphismus $\varphi : G \to G'$ von beliebigen Gruppen der Kern $\ker(\varphi) \subseteq G$ und das Bild $\mathrm{im}(\varphi) \subseteq G'$ Untergruppen sind. Der wichtigste Satz über Homomorphismen ist der Homomorphiesatz. In unserer Situation lautet er:

Satz 10.8 (Homomorphiesatz) Sei $\varphi : A \to A'$ ein Homomorphismus von abelschen Gruppen.

(a) Dann induziert φ einen Monomorphismus

$$\tilde{\varphi} : \quad A/\ker(\varphi) \quad \to \quad A',$$

der das Diagramm

$$\begin{tikzcd} A \arrow[rr, "\varphi"] \arrow[dr, "\pi"'] & & A' \\ & A/\ker(\varphi) \arrow[ur, "\tilde{\varphi}"'] & \end{tikzcd}$$

kommutativ macht, (d.h., für den $\varphi = \tilde{\varphi} \circ \pi$ gilt). Hierbei ist π der kanonische Epimorphismus.

(b) Weiter induziert $\tilde{\varphi}$ einen Isomorphismus

$$A/\ker(\varphi) \xrightarrow{\sim} \mathrm{im}(\varphi)\,.$$

Beweis (a): *Eindeutigkeit*: Haben wir $\tilde{\varphi}$ mit $\varphi = \tilde{\varphi} \circ \pi$, so gilt für alle $x \in A : \varphi(x) = \tilde{\varphi}(\pi(x)) = \tilde{\varphi}(\overline{x})$. Also muss gelten

$$\tilde{\varphi}(\overline{x}) = \varphi(x)\,.$$

Hierdurch ist $\tilde{\varphi}$ eindeutig bestimmt.

Existenz: Definiere $\tilde{\varphi}$ durch $\tilde{\varphi}(\overline{x}) = \varphi(x)$. Dies ist wohldefiniert (d.h., unabhängig von den Repräsentanten): $\overline{x} = \overline{y} \Rightarrow x - y \in \ker(\varphi) \Rightarrow 0 = \varphi(x-y) = \varphi(x) - \varphi(y) \Rightarrow \varphi(x) = \varphi(y)$.

$\tilde{\varphi}$ ist Homomorphismus: $\tilde{\varphi}(\overline{x} + \overline{y}) = \tilde{\varphi}(\overline{x+y}) = \varphi(x+y) = \varphi(x) + \varphi(y) = \tilde{\varphi}(\overline{x}) + \tilde{\varphi}(\overline{y})$.

$\tilde{\varphi}$ ist injektiv: $\tilde{\varphi}(\overline{x}) = 0 \Rightarrow \varphi(x) = 0 \Rightarrow x \in \ker(\varphi) \Rightarrow \overline{x} = 0$.

(b): Offenbar liegt das Bild von $\tilde{\varphi}$ in $\mathrm{im}(\varphi)$, wir erhalten aus $\tilde{\varphi}$ also auch einen Monomorphismus $\tilde{\varphi} : A/\ker(\varphi) \to \mathrm{im}(\varphi)$. Dieser ist auch surjektiv: Iat $a' \in \mathrm{im}(\varphi)$, so gibt es ein $x \in A$ mit $\varphi(x) = a'$. Dann ist $\tilde{\varphi}(\overline{x}) = \varphi(x) = a'$.

Quotientenräume

Sei K ein Körper und V ein K-Vektorraum. Ein Untervektorraum $W \subseteq V$ ist insbesondere eine Untergruppe der kommutativen Gruppe $(V, +)$, und wir können die Quotientengruppe V/W bilden.

Satz 10.9 Es gibt genau eine Skalarmultiplikation

$$\begin{array}{ccccc} K & \times & V/W & \to & V/W \\ & (\lambda, \overline{u}) & & \mapsto & \lambda \cdot \overline{u}\,, \end{array}$$

die V/W zu einem K-Vektorraum macht und für die der kanonische Epimorphismus

$$\pi : V \to V/W$$

ein Vektorraum-Homomorphismus ist. Diese Verküpfung ist

$$\lambda \cdot \overline{v} = \overline{\lambda v} \qquad \text{für } \lambda \in K, v \in V\,.$$

Beweis Ist π eine lineare Abbildung, so muss gelten

$$\pi(\lambda v) = \lambda \cdot \pi(v)\,,$$

also gerade (10.9.1). Die Skalarmultiplikation ist also, wenn sie existiert, eindeutig bestimmt.

Es bleibt noch zu zeigen, dass (10.9.1) wohldefiniert ist; dann können wir es als Definition der Skalarmultiplikation nehmen. Sei $\lambda \in K$, und seien $v, v' \in V$ mit $\overline{v} = \overline{v'}$. Dann ist $v - v' \in W$ und damit auch $\lambda(v - v') \in W$, d.h., $\lambda v - \lambda v' \in W$. Es folgt $\overline{\lambda v} = \overline{\lambda v'}$.

Die Vektorraum-Axiome (siehe LA I 5.1 (i)-(v)) für V/W (und die Sklalarmultiplikation (10.9.1)) folgen nun sofort aus der Gültigkeit der Axiome für V: Zum Beispiel ist $\lambda \cdot (\overline{v_1} + \overline{v_2}) = \lambda \cdot \overline{v_1 + v_2} = \overline{\lambda(v_1 + v_2)} = \overline{\lambda v_1 + \lambda v_2} = \overline{\lambda v_1} + \overline{\lambda v_2} = \lambda \cdot \overline{v_1} + \lambda \cdot \overline{v_2}$.

Für die Dimension von Quotientenräumen haben wir:

Satz 10.10 Sei V endlich-dimensional und $W \subseteq V$ ein Unterraum. Dann gilt

$$\dim V/W \quad = \quad \dim V - \dim W\,.$$

Beweis Seien $w_1 \ldots, w_r \in W$, so dass $(w_1, \ldots, w_r)$ eine Basis von W ist. Ergänze dies mit $v_1, \ldots, v_s \in V$ zu einer Basis $(w_1, \ldots, w_r, v_1, \ldots, v_s)$ von V. Damit ist $\dim W = r$ und $\dim V = r + s$. Wir behaupten nun, dass $(\overline{v_1}, \ldots, \overline{v_s})$ eine Basis von V/W ist; dann folgt $\dim V/W = s$ und die Behauptung. Da $w_1, \ldots, w_r, v_1, \ldots, v_s$ ein Erzeugendensystem von V ist und $\pi : V \to V/W$ surjektiv ist, ist $\overline{w_1}, \ldots, \overline{w_r}, \overline{v_1}, \ldots, \overline{v_s}$ ein Erzeugendensystem von V/W, wegen $\overline{w_1} = \ldots = \overline{w_r} = 0$, also $\overline{v_1}, \ldots, \overline{v_s}$ ein Erzeugendensystem. Wir zeigen nun, dass diese Vektoren linear unabhängig sind. Seien $\alpha_1, \ldots, \alpha_s \in K$ mit

$$0 = \alpha_1\overline{v_1} + \ldots + \alpha_s\overline{v_s} = \overline{\alpha_1 v_1 + \ldots + \alpha_s v_s}\,.$$

Dann ist also

$$\alpha_1 v_1 + \ldots + \alpha_s v_s = \beta_1, w_1 + \ldots + \beta_r w_r \in W\,.$$

Aus der linearen Unabhängigkeit von $w_1, \ldots, w_r, v_1, \ldots, v_s$ folgt nun $\alpha_1 = \ldots = \alpha_s = 0 = \beta_1 = \ldots = \beta_r$. Damit sind $\overline{v_1}, \ldots, \overline{v_s}$ linear unabhängig.

Bemerkung 10.11 Dies gibt einen neuen Beweis der Rangformel: Sei $\varphi : V \to W$ eine lineare Abbildung zwischen endlich-dimensionalen K-Vektorräumen. Dann gibt der Homomorphiesatz einen Isomorphismus

$$V/\ker(\varphi) \xrightarrow{\sim} \mathrm{im}(\varphi)\,.$$

Es folgt

$$\mathrm{rg}(\varphi) = \dim\ \mathrm{im}(\varphi) = \dim V/\mathrm{im}(\varphi) = \dim V - \dim\ker(\varphi)\,.$$

§11 Ordnungsrelationen

Ist M eine Menge und $\leq$ eine Ordnungsrelation auf M, so nennen wir das Paar $(M, \leq)$ eine **geordnete Menge**. Oft unterdrücken wir das Symbol $\leq$ in der Bezeichnung.

Definition 11.1 Eine geordnete Menge $(M, \leq)$ heißt **total geordnet** (oder **angeordnet** oder **Kette**), wenn für alle $x, y \in M$ gilt

$$x \leq y \quad \vee \quad y \leq x\,.$$

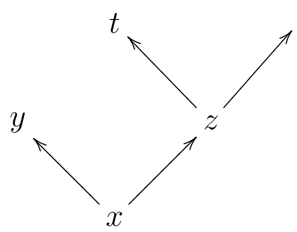
Die Ordnungsrelation $\leq$ heißt dann eine **Totalordnung** (oder **Anordnung**).

Beispiele 11.2 (a) $\leq$ auf $\mathbb{R}$ ist eine Totalordnung.

(b) Sei M eine Menge mit mindestens 2 Elementen. Dann ist die Inklusion $\subseteq$ auf $\mathfrak{P}(M)$ keine Totalordnung: sind $x, y \in M$ mit $x \neq y$, so ist $\{x\} \nsubseteq \{y\}$ und $\{y\} \nsubseteq \{x\}$.

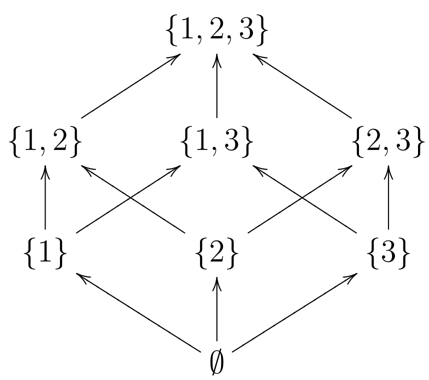
(c) Die Teilbarkeitsrelation $|$ auf $\mathbb{N}$ ist keine Totalordnung: $2 \nmid 3$ und $3 \nmid 2$.

11.3 Man kann Ordnungen durch gerichtete Graphen visualisieren:



heißt $x \leq y, x \leq z, z \leq t$, aber $y \nleq t$.

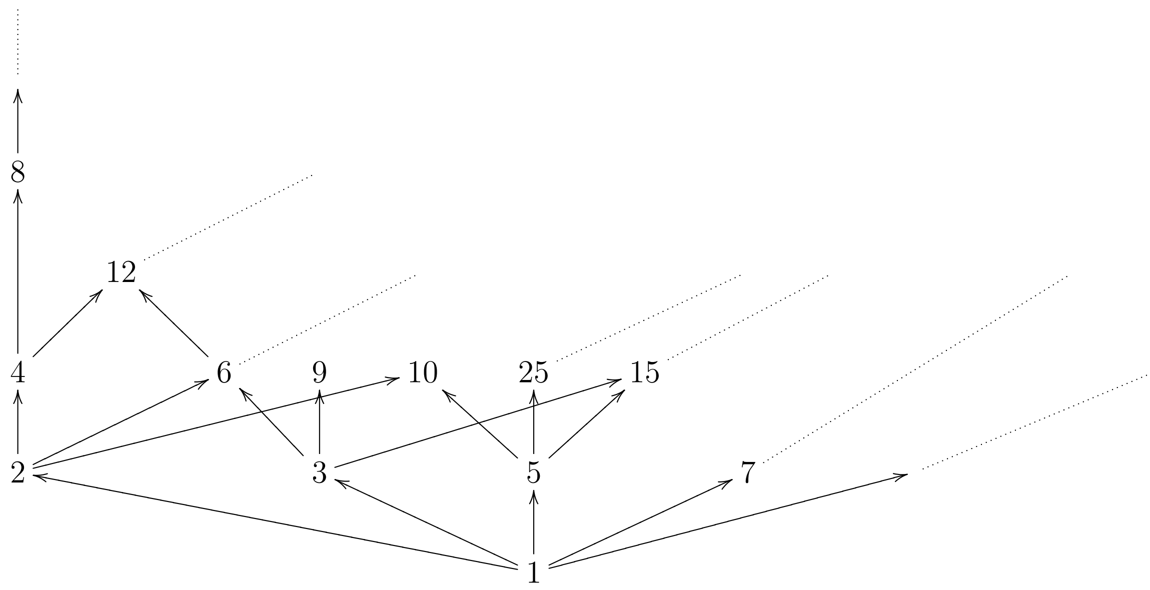
Beispiele: $(\mathfrak{P}(\{1,2,3\}), \subseteq)$:



$(\mathbb{N}, \leq)$

$$\begin{array}{c} \vdots \\ \uparrow \\ 5 \\ \uparrow \\ 4 \\ \uparrow \\ 3 \\ \uparrow \\ 2 \\ \uparrow \\ 1 \end{array}$$

$(\mathbb{N}, |)$



Definition 11.4 Sei $(M, \leq)$ eine geordnete Menge, und sei $U \subseteq M$ eine Teilemenge. Ein Element $x \in U$ heißt

(a) **minimales Element** von U, wenn gilt

$$\text{für alle } y \in U: \quad y \leq x \quad \Rightarrow \quad y = x\,.$$

(b) **kleinstes Element** von U, wenn gilt

$$\text{für alle } y \in U: \quad x \leq y\,.$$

(c) **maximales Element** von U, wenn gilt

$$\text{für alle } y \in U: \quad y \geq x \quad \Rightarrow \quad y = x\,.$$

(d) **größtes Element** von U, wenn gilt

$$\text{für alle } y \in U: \quad x \geq y\,.$$

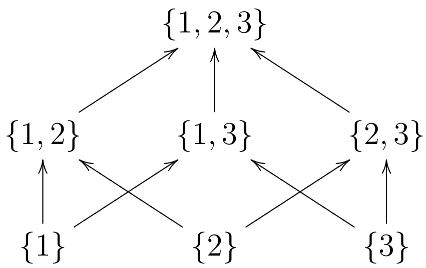
Bemerkungen 11.5 (a) Die Begriffe "maximal" und "minimal" sowie "größtes" und "kleinstes" sind **dual** zueinander, in dem Sinne, dass sie ineinander übergehen, wenn man $\leq$ und $\geq$ vertauscht.

(b) "minimal" $\Leftrightarrow$ es gibt kein echt kleineres Element.
"maximal" $\Leftrightarrow$ es gibt kein echt größeres Element.

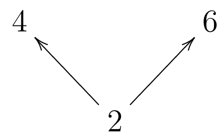
Beispiele 11.6 (a) In $(\mathbb{N}, \leq)$ ist 1 kleinstes und minimales Element; es gibt kein größtes oder maximales Element.

(b) In $(\mathfrak{P}(M), \subseteq)$ ist $\emptyset$ kleinstes und minimales Element, M ist größtes und maximales Element.

(c) In $\mathfrak{P}(\{1,2,3\}) \smallsetminus \{\emptyset\}$ gibt es kein kleinstes Element; $\{1\}, \{2\}$ und $\{3\}$ sind minimale Elemente



In $(\{2,3,6\}, |)$ ist 2 kleinstes und minimales Element, 4 und 6 sind maximale Elemente, und es gibt kein größtes Element.



Satz 11.7 Sei $(M, \leq)$ eine geordnete Menge.

(a) Es gibt höchstes ein größtes Element.

(a) Ist $x \in M$ ein größtes Element, so ist x auch maximal, und jedes maximale Element ist gleich x. Entsprechendes gilt für den dualen Fall minimal/kleinstes.

Beweis Dies folgt aus den folgenden 2 Behauptungen

1. Behauptung: x größtes Element $\quad\Rightarrow\quad$ x maximal

Beweis: x größtes Element und $y \geq x \quad\Rightarrow\quad y = x$, da $y \leq x$.

2. Behauptung: x größtes Element und y maximal $\quad\Rightarrow\quad y = x$.

Beweis: x größtes Element und y maximal $\quad\Rightarrow\quad x \geq y$ und y maximal $\Rightarrow \quad y = x$.
q.e.d.

Bemerkung 11.8 Wie wir in den Beispielen gesehen haben, kann es mehrere maximale Elemente geben, und diese sind nicht unbedingt größte Elemente. Alle Aussagen gelten entsprechend für minimale/kleinste Elemente.

Satz 11.9 Ist $(M \leq)$ eine Kette, so gibt es höchstens ein maximales Element, das dann auch das größte Element von M ist. (Entsprechendes gilt dual für minimal/kleinstes).

Beweis Sei $x \in M$ maximal und sei $y \in M$. Nach Voraussetzung gilt $y \geq x$ oder $y \leq x$. Ist $y \geq x$, so gilt $y = x$ wegen der Maximalität von x. Also gilt immer $y \leq x$, also ist x größtes Element.

Definition 11.10 Sei $(M, \leq)$ eine geordnete Menge und sei $N \subseteq M$ eine Teilmenge. Ein Element $m \in M$ heißt

(a) obere Schranke von N in M, wenn gilt:

$$\text{für alle } x \in N: \quad x \leq m\,,$$

(b) Suprenum (oder obere Grenze) von N in M, wenn m die kleinste obere Schranke von N in M ist, Bezeichnung: $\mathrm{Sup}(N)$

(c) untere Schranke von N in M, wenn gilt:

$$\text{für alle } x \in N: \quad x \geq m\,,$$

(d) Infimum (oder untere Grenze) von N in M, wenn m die größte untere Schranke von N in M ist. Bezeichnung: $\mathrm{Inf}(N)$ (Die Begriffe "obere Schranke" und "untere Schranke" sind also dual zueinander, ebenso wie "Supremum" und "Infimum").

Beispiele 11.11 (a) Betrachte $\mathbb{R}$ mit der natürlichen Ordnung $\leq$ und darin das Intervall $I = [2, 3[$. Dann gilt
4 ist eine obere Schranke von I,
3 ist Supremum von I,
I hat kein größtes Element,
2 ist Infimum von I,
2 ist auch kleinstes Element.

(b) Sei M eine Menge und seien $N_1, N_2 \subseteq M$. Das Supremum von $\{N_1, N_2\}$ in $(\mathfrak{P}(M), \subseteq)$ ist $N_1 \cup N_2$, das Infimum ist $N_1 \cap N_2$.

(c) In $(\mathfrak{P}(\{1, 2\}) \setminus \{\emptyset\}, \subseteq)$ hat die Menge $\{\{1\}, \{2\}\}$ kein Infimum.

(d) In $(\mathbb{N}, |)$ ist $\sup(m, n) = \ \mathrm{kg}\,V(M, n)$ und $\inf(m, n) = ggT(m, n)$.

Lemma 11.12 Sei $(M, \leq)$ eine geordnete Menge, und sei $N \subseteq M$.

(a) N hat höchstens ein Supremum.

(b) Hat N ein größtes Element, so ist es auch das Supremum von N in M.

(c) Wenn N eine obere Schranke $m \in M$ mit $m \in N$ hat, so ist m größtes Element von N und das Supremum von N in M.

Für Infima gelten die dualen Aussagen.

Beweis (a): Ist m ein Supremum von N in M, so ist m das kleinste Element der geordneten Menge $\{m \in M \mid m \text{ obere Schranke von } N \text{ in } M\}$ und daher eindeutig nach 11.7 (duale Version).

(b): Sei $n \in N$ größtes Element von $(N, \leq)$. Dann ist n obere Schranke von N in M. Ist $m \in M$ eine weitere obere Schranke von N in M, so ist insbesondere $n \leq m$. Also ist n kleinste obere Schranke.

(c): Die erste Aussage ist trivial, und die zweite folgt aus (b). q.e.d.

Lemma 11.13 (Zornsches Lemma) Sei $(M, \leq)$ eine nichtleere geordnete Menge. Hat jede Teilmenge $A \subseteq M$, die eine Kette bezüglich $\leq$ ist, eine obere Schranke in M, so gibt es zu jedem $x \in M$ ein maximales Element m von M mit $x \leq m$.

Dieses Lemma kann man nicht "beweisen", sondern es kann bei einem exakten Aufbau der Mengenlehre nur als ein (unbeweisbares) Axiom hinzugenommen werden. Eine äquivalente Aussage ist das

Auswahlaxiom 11.14 Sei $(M_i)_{i \in I}$ eine Familie von nicht-leeren Mengen, Dann ist das kartesische Produkt $\prod\limits_{i \in I} M_i$ nicht leer.

Diese Aussage, dass man eine Familie $(x_i)_{i \in I}$ hat, also **gleichzeitig** zu jedem $i \in I$ ein $x_i \in M_i$ auswählen kann, ist plausibel, aber ebenso wenig beweisbar wie das Zornsche Lemma. Eine weitere äquivalente Formulierung ist

Auswahlaxiom' 11.15 Ist $f : M \to N$ eine surjektive Abbildung, so gibt es eine Abbildung $g : N \to M$ mit $f \circ g = \mathrm{id}_N$.

Lemma 11.16 Jede nichtleere, **endliche** Kette besitzt ein größtes und ein kleinstes Element.

Beweis (für "größtes", da "kleinstes" dual) durch Induktion:

Sei $(A, \leq)$ eine Kette mit $|A| \in \mathbb{N}$.

$|\mathbf{A}| = \mathbf{1}$: trivial.

$|\mathbf{A}| \geq \mathbf{2}$: Sei $a \in A$ beliebig. Nach Induktionsvoraussetzung hat die Kette $(A \setminus \{a\}, \leq)$ ein größtes Element b. Da A Kette ist, gilt $a \leq b \ \vee \ b \leq a$.

1. Fall: $a \leq b$
Dann ist b offenbar größtes Element von A.

2. Fall: $b \leq a$
Ist $c \in A \setminus \{a\}$, so gilt $c \leq b \leq a$, also $c \leq a$, d.h., a ist in diesem Fall größtes Element von A. q.e.d.

"Anwendung"

Satz 11.17 Jeder Vektorraum hat eine Basis.

Beweis Sei K ein Körper und V ein K-Vektorraum, sei $M := \{\mathfrak{u} \in \mathfrak{P}(V) \mid \mathfrak{u}$ linear unabhängig$\}$, geordnet mit $\subseteq$. Wir zeigen zunächst, dass M die Voraussetzung des Zornschen Lemmas erfüllt: Sei also $A \subseteq M$ eine Kette bezüglich $\subseteq$.

Wenn wir zeigen können, dass

$$S := \bigcup_{\mathfrak{u} \in A} \mathfrak{u}$$

linear unabhängig ist, ist S offenbar obere Schranke von A in M. Sei also

$$\lambda_1 v_1 + \ldots + \lambda_s v_s = 0 \in V \text{ mit } \lambda_i \in K, v_i \in S\,.$$

Es ist $v_1 \in \mathfrak{u}_1, \ldots, v_s \in \mathfrak{u}_s$ für geeignete $\mathfrak{u}_1, \ldots, \mathfrak{u}_s \in A$. Nach Lemma 11.16 hat die Kette $\{\mathfrak{u}_1, \ldots, \mathfrak{u}_s\} \subseteq A$ ein größtes Element $\mathfrak{u}$, d.h., $\mathfrak{u}_1, \ldots, \mathfrak{u}_s \subseteq \mathfrak{u} \in A$, also $v_1, \ldots v_s \in \mathfrak{u} \in A \subseteq M$. Da $\mathfrak{u} \in M$, ist $\mathfrak{u}$ linear unabhängig; insbesondere sind $v_1, \ldots, v_s$ linear unabhängig und demnach $\lambda_1 = \ldots = \lambda_s = 0$.

Wir können also das Zornsche Lemma auf M anwenden und schließen, dass M ein maximales Element $\mathfrak{b}$ enthält. Wir behaupten, dass $\mathfrak{b}$ eine Basis von V ist.

"linear unabhängig": klar, da $\mathfrak{b} \in M$.

"Erzeugendensystem": Sei $v \in V$ beliebig. Ist $v \in \mathfrak{b}$, so ist nichts zu zeigen, sei also $v \notin \mathfrak{b}$. Wegen $\mathfrak{b} \subsetneq \mathfrak{b} \cup \{v\}$ und der Maximalität von $\mathfrak{b}$ ist $\mathfrak{b} \cup \{v\}$ linear abhängig, d.h., es gibt $\lambda, \lambda_w (w \in \mathfrak{b})$ in k, fast alle, aber nicht alle gleich 0, mit

$$\lambda v + \sum_{w \in \mathfrak{b}} \lambda_w w = 0\,.$$

Da $\mathfrak{b}$ linear unabhängig ist, muß $\lambda \neq 0$ gelten, d.h.,

$$v = \sum_{w \in \mathfrak{b}} -\tfrac{\lambda_w}{\lambda} w\,.$$

§12 Freier Vektorraum über einer Menge und Tensorprodukt

Sei K ein Körper. Der folgende Satz verallgemeinert LA I 8.31.

Satz 12.1 (Universelle Eigenschaft einer Basis). Sei V ein K-Vektorraum und sei $(b_i)_{i \in I}$ eine Basis von V. Zu jedem K-Vektorraum W und zu jeder Familie $(w_i)_{i \in I}$ von Vektoren in W gibt es dann genau eine lineare Abbildung

$$\varphi : V \to W$$

mit $\varphi(b_i) = w_i$ für alle $i \in I$.

("Eine lineare Abbildung kann man beliebig auf einer Basis vorgeben").

Beweis Jeder Vektor $v \in V$ hat eine eindeutige Darstellung

$$v = \sum_{i \in I} a_i b_i \qquad (12.1.1)$$

mit $a_i \in K$, wobei die Summe endlich ist: für fast alle $i \in I$ ist $a_i = 0$ und der entsprechende Summand $a_i b_i = 0$. Es muss dann gelten

$$\varphi(v) = \varphi(\sum_{i \in I} a_i b_i) = \sum_{i \in I} a_i \varphi(b_i) = \sum_{i \in I} a_i w_i$$

(die Summe ist wieder endlich), wegen der Linearität von φ. Wir können aber φ hierdurch definieren, weil die Darstellung (12.1.1) eindeutig ist. Dann folgt $\varphi(b_i) = w_i$, und φ ist linear, wie man sofort nachrechnet.

Lemma/Definition 12.2 Sei I eine Menge. Dann wird die Menge der Familien in K über I,

$$K^I = \{(a_i)_{i \in I} \mid a_i \in I\},$$

ein K-Vektorraum durch die Addition

$$(a_i)_{i \in I} + (b_i)_{i \in I} = (a_i + b_i)_{i \in I}$$

und die Skalarmultiplikation

$$\lambda (a_i)_{i \in I} = (\lambda a_i)_{i \in I}\,.$$

Beweis Alle Vektorraum-Axiome folgen sofort aus der Gültigkeit dieser Axiome im Vektorraum K.

Bemerkung 12.3 Für $I = \{1, \ldots, n\} = \underline{n}$ erhalten wir wieder den Vektorraum $K^{\underline{n}} = K^n$.

Lemma/Definition 12.4 Sei I eine Menge. Dann ist die Menge

$$K^{(I)} := \{(a_i)_{i \in I} \in K^I \mid a_i = 0 \text{ für fast alle } i \in I\}$$

ein Untervektorraum von K^I und heißt der **freie Vektorraum über** I.

Beweis Es ist klar, dass sich die Beziehung "$a_i = 0$ für fast alle $i \in I$" bei Addition und Skalarprodukt erhält.

Bemerkung 12.5 Ist I eine **endliche** Menge, so gilt offenbar $K^{(I)} = K^I$, insbesondere gilt für $n \in \mathbb{N} : K^{(\underline{n})} = K^{\underline{n}} = K^n$.

Die Bezeichnung "freier Vektorraum" rechtfertigt sich durch die beiden folgenden Resultate.

Lemma 12.6 Für $i \in I$ sei der Vektor $e_i \in K^{(I)}$ definiert durch

$$e_i = (a_j)_{j \in I} \text{ mit } a_j = \begin{cases} 1 & , \quad j = i, \\ 0 & , \quad \text{sonst.} \end{cases}$$

Dann ist $(e_i)_{i \in I}$ eine Basis von $K^{(I)}$.

Beweis Für $a = (a_i)_{i \in I} \in K^{(I)}$ gilt

$$a = \sum_{i \in I} a_i e_i$$

(beachte, dass diese Summe in Wirklichkeit endlich ist!); und diese Darstellung ist eindeutig.

Satz 12.7 (universelle Eigenschaft des freien Vektorraums über I) Sei W ein K-Vektorraum. Für jede Familie $(w_i)_{i\in I}$ in W gibt es genau eine lineare Abbildung

$$\varphi : K^{(I)} \to W$$

mit $\varphi(e_i) = w_i$ für alle $i \in I$.

Beweis Dies folgt aus 12.6 und 12.1. Explizit: Für $a = (a_i)_{i\in I} = \sum\limits_{i\in I} a_i e_i$ (endliche Summe!) muss gelten

$$\begin{aligned} \varphi(a) = \varphi(\sum_{i\in I} a_i e_i) &= \sum_{i\in I} a_i \varphi(e_i) \\ &= \sum_{i\in I} a_i w_i \,. \end{aligned}$$

Dies zeigt die Eindeutigkeit. Umgekehrt können wir φ hierdurch definieren, und es ist dann $\varphi(e_i) = w_i$.

Wir kommen nun zum Tensorprodukt.

Satz/Definition 12.8 Seien V und W zwei K-Vektorräume. Dann gibt es einen Vektorraum $V \otimes W := V \otimes_K W$ und eine bilineare Abbildung

$$\alpha_{\text{univ}} : V \times W \to V \otimes W$$

mit der folgenden universellen Eigenschaft:

Für jede bilineare Abbildung

$$\alpha : V \times W \to X$$

in einen dritten K-Vektorraum X gibt es genau eine lineare Abbildung $\tilde{\alpha} : V \otimes W \to X$, die das Diagramm

$$\begin{tikzcd} V \times W \arrow[rr, "\alpha_{\text{univ}}"] \arrow[dr, "\alpha"'] & & V \otimes W \arrow[dl, "\tilde{\alpha}"] \\ & X & \end{tikzcd}$$

kommutativ macht.

$V \otimes W$ heißt das **Tensorprodukt** von V und W. (Es ist universell für bilineare Abbildungen $V \times W \to$?)

Beweis *Konstruktion*: Betrachte den freien Vektorraum über $V \times W$

$$K^{(V\times W)}\,.$$

Er hat die Basis $(e_{(v,w)})_{(v,w)\in V\times W}$. Wir haben eine Abbildung

$$\begin{aligned} \alpha_0 \;:\; V \times W &\to K^{(V\times W)} \\ (v,w) &\mapsto e_{(v,w)} \end{aligned}$$

die nicht bilinear ist. Sie wird aber bilinear, wenn wir zum folgenden Quotienten

$$V \otimes W := K^{(V\times W)}/U$$

übergehen: Sei

$$\begin{array}{rcl} U & = & < e_{(\lambda v_1+\mu v_2,w)} - \lambda e_{(v_1,w)} - \mu e_{(v_2,w)}, \\ & & e_{(v,\lambda w_1+\mu w_2)} - \lambda e_{(v,w_1)} - \mu e_{(v,w_2)} \\ & & (v, v_1, v_2 \in V, w, w_1, w_2 \in W, \lambda, \mu \in K) >_K , \end{array}$$

und sei

$$\alpha_{\text{univ}} : V \times W \xrightarrow{\alpha_0} K^{(V\times W)} \to V \otimes W$$

die Komposition. Bezeichne

$$v \otimes w := \alpha_{\text{univ}}(v, w) = \overline{e_{(v,w)}} \in V \otimes W ,$$

also die Klasse von $e_{(v,w)}$ in $V \otimes W$. Dann gilt

$$(\lambda v_1 + \mu v_2) \otimes w \quad = \quad \lambda\, v_1 \otimes w + \mu\, v_2 \otimes w \tag{12.7.1}$$

in $V \otimes W$, denn nach Definition von U gilt modulo U:

$$\begin{array}{rcl} 0 & = & \overline{e_{(\lambda v_1+\mu v_2,w)} - \lambda e_{(v_1,w)} - e_{(v_2,w)}} \\ & = & \overline{e_{(\lambda v_1+\mu v_2,w)}} - \lambda \overline{e_{(v_1,w)}} - \mu \overline{e_{(v_2,w)}} \\ & = & (\lambda v_1 + \mu v_2) \otimes w - \lambda(v_1 \otimes w) - \mu(v_2 \otimes w) . \end{array}$$

Entsprechend gilt wegen der Elemente vom 2. Typ in U

$$v \otimes (\lambda w_1 + \mu w_2) = \lambda\, v \otimes w_1 + \mu\, v \otimes w_2 . \tag{12.7.2}$$

Hieraus folgt, dass α_{univ}, also die Abbildung $(v, w) \mapsto v \otimes w$, bilinear ist ("das Tensorprodukt ist bilinear").

Universelle Eigenschaft: Sei

$$\alpha : V \times W \to X$$

bilinear.

Existenz von $\tilde{\alpha}$: Definiere

$$\tilde{\alpha} : V \otimes W \to X$$

wie folgt: Zunächst haben wir nach Satz 12.6 eine eindeutig bestimmte lineare Abbildung

$$\alpha' : K^{(V\times W)} \to X$$

mit $\alpha'(e_{(v,w)}) = \alpha(v, w)$. Wir zeigen nun, dass $\alpha'(u) = 0$ für alle $u \in U$; dann ist die induzierte Abbildung

$$\begin{array}{rcl} \tilde{\alpha} : V \otimes W & = & K^{(V\times W)}/U \to X \\ \tilde{\alpha}(\overline{a}) & := & \alpha'(a) \text{ für } a \in K^{(V\times W)} \end{array}$$

wohldefiniert: $\overline{a_1} = \overline{a_2} \Rightarrow a_1 - a_2 \in U \Rightarrow 0 = \alpha'(a_1 - a_2) = \alpha'(a_1) - \alpha'(a_2) \Rightarrow \alpha'(a_1) = \alpha'(a_2)$.

Es genügt zu zeigen, dass $\alpha'(u) = 0$ für alle Erzeugenden von U. Für die Erzeugenden vom 1. Typ haben wir wegen Linearität und Definition von α'

$$\alpha'(e_{(\lambda v_1+\mu v_2,w)} - \lambda\, e_{(v_1,w)} - \mu\, e_{(v_2,w)}) = \alpha(\lambda v_1 + \mu v_2, w) - \lambda\alpha(v_1, w) - \mu\alpha(v_2, w) = 0\,,$$

da α bilinear ist. Entsprechend zeigt man $\alpha'(u) = 0$ für die Erzeugenden vom 2. Typ.

Mit dieser Definition von $\tilde{\alpha}$ ist nun das Diagramm

(12.7.3)
$$\begin{tikzcd} V \times W \arrow[rr, "\alpha_{\mathrm{univ}}"] \arrow[dr, "\alpha"'] & & V \otimes W \arrow[dl, "\tilde{\alpha}"] \\ & X & \end{tikzcd}$$

kommutativ , d.h., $\tilde{\alpha} \circ \alpha_{\mathrm{univ}} = \alpha$, denn es gilt

$$\begin{aligned} \tilde{\alpha}(\alpha_{\mathrm{univ}}(v, w)) &= \tilde{\alpha}(v \otimes w) = \tilde{\alpha}(\overline{e_{(v,w)}}) \\ &= \alpha'(e_{(v,w)}) = \alpha(v, w)\,. \end{aligned}$$

Eindeutigkeit von $\tilde{\alpha}$: Ist das Diagramm (12.7.3) kommutativ mit einer beliebigen linearen Abbildung $\tilde{\alpha}$, so muss gelten $\tilde{\alpha}(\alpha_{\mathrm{univ}}(v, w)) = \alpha(v, w)$, also

$$\tilde{\alpha}(v \otimes w) = \alpha(v, w)\,.$$

Hierdurch ist $\tilde{\alpha}$ eindeutig bestimmt, da die Elemente $v \otimes w$ ein Erzeugendensystem von $V \otimes W$ bilden (als Bilder der $e_{(v,w)}$, die ein Erzeugendensystem von $K^{(V\times W)}$ bilden).

Lemma 12.8 Das Paar $(V \otimes W, \alpha_{\mathrm{univ}})$ ist bis auf eindeutige Isomorphie eindeutig bestimmt.

Beweis Sei $(V \otimes' W, \alpha'_{\mathrm{univ}})$ eine andere Lösung des universellen Problems. Dann erhalten wir ein Diagramm

(12.8.1)
$$\begin{tikzcd} & V \otimes W \arrow[d, shift right, "\tilde{\alpha}'"'] \\ V \times W \arrow[ur, "\alpha_{\mathrm{univ}}"] \arrow[dr, "\alpha'_{\mathrm{univ}}"'] & \\ & V \otimes' W \arrow[u, shift right, "\tilde{\alpha}"'] \end{tikzcd}$$

mit den folgenden Eigenschaften: Nach der universellen Eigenschaft von $(V \otimes W, \alpha_{\mathrm{univ}})$, und wegen der Bilinearität von $\alpha' = \alpha'_{\mathrm{univ}}$ gibt es genau eine lineare Abbildung $\widetilde{\alpha'}$ mit $\widetilde{\alpha'}\,\alpha_{\mathrm{univ}} = \alpha'_{\mathrm{univ}}$. Umgekehrt gibt es wegen der angenommenen universellen Eigenschaft von $(V \otimes' W, \alpha'_{\mathrm{univ}})$ und der Bilinearität von $\alpha = \alpha_{\mathrm{univ}}$ genau eine lineare Abbildung $\tilde{\alpha}$ mit $\tilde{\alpha}\,\alpha'_{\mathrm{univ}} = \alpha_{\mathrm{univ}}$.

Es folgt

$$\tilde{\alpha}\,\widetilde{\alpha'}\alpha_{\mathrm{univ}} = \tilde{\alpha}\,\alpha'_{\mathrm{univ}} = \alpha_{\mathrm{univ}}\,.$$

Andererseits gibt natürlich auch

$$\mathrm{id}_{V\otimes W}\,\alpha_{\mathrm{univ}} = \alpha_{\mathrm{univ}}\,.$$

Wegen der universellen Eigenschaft von $(V \otimes W, \alpha_{\text{univ}})$ folgt

$$\tilde{\alpha}\,\widetilde{\alpha'} = \mathrm{id}_{V\otimes W}\,,$$

denn es gibt nur eine lineare Abbildung, die

$$\begin{tikzcd}
V \times W \arrow[rr, "\alpha_{\text{univ}}"] \arrow[dr, "\alpha_{\text{univ}}"'] & & V \otimes W \arrow[dl, "?"] \\
& V \otimes W &
\end{tikzcd}$$

kommutativ macht (wegen der universellen Eigenschaft von α_{univ} und der Bilinearität von α_{univ}). Genauso folgt

$$\widetilde{\alpha'}\tilde{\alpha} = \mathrm{id}_{V\otimes' W}\,.$$

Daher ist $\widetilde{\alpha'}$ ein Isomorphismus, mit Inversem $\tilde{\alpha}$. Weiter sind $\tilde{\alpha}$ und $\widetilde{\alpha'}$ die einzigen linearen Abbildungen, die das Diagramm (12.8.1) kommutativ machen. Das ist die gemeinte Eindeutigkeit der Isomorphismen $\tilde{\alpha}$ und $\widetilde{\alpha'}$.

Bemerkungen 12.9 (a) Es ist eine ganz allgemeine Eigenschaft, dass Objekte, die universelle Eigenschaften erfüllen, hierdurch bis auf kanonische Isomorphie eindeutig sind.

(b) Die Elemente von $V \otimes W$ heißen auch **Tesoren**. Elemente in $V \otimes W$ von der Form $v \otimes w$ heißen **Tensorprodukte**. Nicht jeder Tensor ist von dieser Form (siehe unten). Da aber die Tensorprodukte ein Erzeugendensystem bilden (siehe oben) und immer gilt

$$(\lambda v) \otimes w = \lambda(v \otimes w) = v \otimes (\lambda w) \quad (\lambda \in K)\,,$$

ist jedes Element in $V \otimes W$ von der Form

$$\sum_{i=1}^{r} v_i \otimes w_i\,.$$

(c) Die universelle Eigenschaft des Tensorprodukts wird oft in der folgenden Weise benutzt: Hat man eine bilineare Zuordnung $(v, w) \mapsto \alpha(v, w)$, so beschreibt man die zugehörige lineare Abbildung auf $V \otimes W$ einfach durch die Zuordnung

$$v \otimes w \mapsto \alpha(v, w)\,. \tag{12.9.1}$$

Wie wir eben bemerkt haben, sind nicht alle Elemente $V \otimes W$ von der Form $v \otimes w$, und die Darstellung $v \otimes w$ ist auch nicht eindeutig (zum Beispiel ist $\lambda v \otimes w = v \otimes \lambda w$). Aber die universelle Eigenschaft sagt gerade, dass es genau eine lineare Abbildung mit dieser Eigenschaft (12.9.1) gibt (vorausgesetzt, die Abbildung $(v, w) \mapsto \alpha(v, w)$ ist bilinear!).

Satz 12.10 Sei $(b_i)_{i\in I}$ eine Basis von V und $(c_j)_{j\in J}$ eine Basis von W. Dann ist die Familie der Tensorprodukte

$$(b_i \otimes c_j)_{(i,j)\in I\times J}$$

eine Basis von $V \otimes W$.

Beweis Sei $v \in V$ und $w \in W$. Dann gibt es $(\alpha_i)_{i\in I} \in K^{(I)}$ und $(\beta_j)_{j\in J} \in K^{(J)}$ mit

$$v = \sum_{i\in I} \alpha_i \beta_i \quad , \quad w = \sum_{j\in J} \beta_j e_j$$

(endliche Summen!). Dann ist

$$v \otimes w = (\sum_{i\in I} \alpha_i b_i) \otimes (\sum \beta_j c_j) = \sum_{i,j} \alpha_i \beta_j \; b_i \otimes c_j \,.$$

Da jedes Element von $V \otimes W$ Summe von Tensorprodukten ist, folgt dass die $b_i \otimes c_j$ ein Erzeugendensystem von $V \otimes W$ bilden.

Angenommen, es gibt eine endliche Linearkombination

(12.10.1) $$\sum_{i,j} \lambda_{ij} \; b_i \otimes c_j = 0 \,,$$

mit $\lambda_{ij} \in K$, nicht alle null. Sei etwa $\lambda_{rs} \neq 0$, $(r,s) \in I \times J$. Wir betrachten nun die Abbildung

$$\begin{array}{rrcl} \alpha : & V \times W & \to & K \\ & (v,w) & \mapsto & \alpha_r \cdot \beta_s \,, \end{array}$$

falls $v = \sum_{i\in I} \alpha_i b_i$, $w = \sum_{j\in J} \beta_j c_j$ mit $(\alpha_i)_{i\in I} \in K^{(I)}$ und $(\beta_j)_{i\in J} \in K^{(J)}$. Da diese Darstellungen von v und w eindeutig sind, ist α wohldefiniert. Weiter ist α offenbar bilinear, da die Abbildungen

$$\begin{array}{rcl} V & \to & K \\ v & \mapsto & \alpha_r \end{array} \qquad\qquad \begin{array}{rcl} W & \to & K \\ w & \mapsto & \beta_s \end{array}$$

(Darstellungen von v und w wie oben) linear sind. Nach der universellen Eigenschaft gibt es also eine lineare Abbildung

$$\begin{array}{rrcl} \tilde{\alpha} : & V \otimes W & \to & K \\ & \tilde{\alpha}(v \otimes w) & = & \alpha(v,w) \,. \end{array}$$

Offenbar gilt

$$\tilde{\alpha}(b_i \otimes c_j) = \alpha(b_i, c_j) = \begin{cases} 1 & , \quad (i,j) = (r,s), \\ 0 & , \quad \text{sonst.} \end{cases}$$

Angewandt auf (12.10.1) folgt

$$0 = \tilde{\alpha}(\sum_{i,j} \lambda_{ij} b_i \otimes c_j) = \sum_{i,j} \lambda_{ij} \tilde{\alpha}(b_i \otimes c_j) = \lambda_{rs} \,,$$

im Widerspruch zur Annahme. Die $b_i \otimes c_j$ sind also linear unabhängig.

Corollar 12.11 Ist $\dim V = m$ und $\dim W = n$, so gilt

$$\dim V \otimes W = m \cdot n \,.$$

Beweis Ist $e_1, \ldots, e_m$ eine Basis von V und $f_1, \ldots, f_n$ eine Basis von W, so hat die Basis

$$(e_i \otimes f_j)_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$$

die Mächtigkeit mn.

Bemerkung 12.12 In der Situation von Corollar 12.11 und seinem Beweis hat also jedes Element in $V \otimes W$ eine Darstellung

$$\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}\ e_i \otimes f_j$$

mit eindeutig bestimmten $a_{ij} \in K$. Fixiert man also Basen, so wird ein Tensor durch das System (= die Matrix) der Koeffizienten

$$a_{ij}$$

beschrieben. In dieser Form werden die Tensoren oft in der Physik eingeführt. Da man doch manchmal die Basis wechseln muss oder die Wirkung von Symmetriegruppen verstehen muss, wird dann noch oft angegeben, "wie sich die a_{ij} transformieren" (unter linearen Abbildungen nämlich). Da man zum Ausgangsraum V auch oft noch den Dualraum V^* betrachtet, unterscheidet man "kovariante" und "kontravariante" Transformationen und macht dies durch untere und obere Indizis deutlich. Dies hängt auch mit der Einsteinschen Summenkonvention zusammen. Siehe Bröckers Buch, S. 212-219.

Satz 12.13 Sind V und W endlich-dimensional, so hat man einen kanonischen Isomorphismus von K-Vektorräumen

$$\phi : V^* \otimes W \xrightarrow{\sim} \mathrm{Hom}(V, W)$$

Hierbei ist $V^* = \mathrm{Hom}(V, K)$ der Dualraum von V.

Beweis Die Abbildung wird beschrieben durch

$$\chi \otimes w \quad \mapsto \quad (v \mapsto \chi(v)w)\,.$$

Dies ist im Sinne der Bemerkung 12.9 (c) zu verstehen: Die Abbildung ist diejenige, die (vermöge der universellen Eigenschaft des Tensorprodukts) zur bilinearen Abbildung

$$V^* \times W \to \mathrm{Hom}(V, W)$$

$$(\chi, w) \mapsto (v \mapsto \chi(v)w)$$

assoziiert ist.

Um die Bijektivität von ϕ zu zeigen, benutzen wir Basen. Seien $e = (e_1, \ldots, e_m)$ und $f = (f_1, \ldots, f_n)$ Basen von V bzw. W, und sei $e^* = (e_1^*, \ldots, e_m^*)$ die Dualbasis zu e von V^*, charakterisiert durch

$$e_i^*(e_j) = \delta_{ij} \quad \text{(Kronecker-Symbol)}\,.$$

Dann gilt für ϕ

$$\phi(e_i^* \otimes f_j)(e_k) = e_i^*(e_k) f_j = \delta_{ik} f_j\,,$$

also

$$\phi(\sum_{i,j} a_{ij}\ e_i^* \otimes f_j)(e_k) = \sum_{i,j} a_{ij}\delta_{ik} f_j = \sum_{j} a_{kj} f_j$$

Daher bildet ϕ den Tensor $\sum_{i,j} a_{ij} e_i^* \otimes f_j$ auf die lineare Abbildung mit der Matrixdarstellung $(a_{ij})^t$ bezüglich der Basen e und f ab (siehe LA I Definition 10.21). Dies zeigt, dass ϕ ein Isomorphismus ist.

Es gibt viele weitere Isomorphismen für Tensorprodukte:

Satz 12.14 Seien U, V, W K-Vektorräume. Dann gibt es kanonische Isomorphismen

$$
\begin{array}{llcllcll}
\text{(a)} & V \otimes W & \xrightarrow{\sim} & W \otimes V & \text{mit} & v \otimes w & \mapsto & w \otimes v, \\
\text{(b)} & K \otimes_K V & \xrightarrow{\sim} & V & \text{mit} & \lambda \otimes v & \mapsto & \lambda v, \\
\text{(c)} & U \otimes (V \otimes W) & \xrightarrow{\sim} & (U \otimes V) \otimes W & \text{mit} & (u \otimes v) \otimes w & \mapsto & u \otimes (v \otimes w), \\
\text{(d)} & (U \oplus V) \otimes W & \xrightarrow{\sim} & (U \otimes W) \oplus (V \otimes W). & & & &
\end{array}
$$

Beweis Dies erhält man über die universelle Eigenschaft: Wir beweisen nur einen Teil:

(a): Die Abbildung

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

ist offenbar bilinear. Nach der universellen Eigenschaft des Tensorprodukts induziert sie also eine lineare Abbildung

$$\varphi : V \otimes W \to W \otimes V, \quad \text{mit } v \otimes w \mapsto w \otimes v\,.$$

Mit demselben Argument erhält man eine lineare Abbildung

$$\psi : W \otimes V \to V \otimes W, \quad \text{mit } w \otimes v \mapsto v \otimes w\,.$$

Es ist offenbar $\psi \circ \varphi = \mathrm{id}$, denn $\psi(\varphi(v \otimes w)) = \psi(w \otimes v) = v \otimes w$, und die $v \otimes w$ bilden ein Erzeugendensystem von $V \otimes W$. Genauso folgt $\varphi \circ \psi = \mathrm{id}$. Daher ist φ ein Isomorphismus (mit Inversem ψ).

(b): Die Abbildung

$$K \times V \to V, \quad (\lambda, v) \mapsto \lambda v$$

ist offenbar bilinear, induziert also eine lineare Abbildung

$$\varphi : K \otimes V \to V, \quad \text{mit } \lambda \otimes v \mapsto \lambda v\,.$$

Die Umkehrabbildung ist die lineare Abbildung(!)

$$\psi : V \to K \otimes V, \quad v \mapsto 1 \otimes v\,.$$

Denn es ist $\psi(\varphi(\lambda \otimes v)) = \psi(\lambda v) = 1 \otimes \lambda v = \lambda \otimes v$ und $\varphi(\psi(v)) = \varphi(1 \otimes v) = 1v = v$.

(c): selbst!

(d): Übungsaufgabe!

Weiter haben wir:

Satz 12.15 Sind V und W endlich-dimensional, so gibt es einen kanonischen Isomorphismus

$$
\begin{array}{rcl}
V^* \otimes W^* & \xrightarrow{\sim} & (V \otimes W)^* \\
\chi \otimes \nu & \mapsto & (v \otimes w \mapsto \chi(v) \cdot \nu(w))\,.
\end{array}
$$

Beweis(skizze): Seien $\chi \in V^*$ und $\nu \in W^*$ lineare Funktionale. Dann erhalten wir eine bilineare Abbildung

$$\begin{array}{rcl} V \times W & \to & K \\ (v,w) & \mapsto & \chi(v) \mapsto \nu(w)\,. \end{array}$$

Diese induziert eine lineare Abbildung (also ein lineares Funktional auf $V \otimes W$)

$$V \otimes W \to K \quad , v \otimes w \mapsto \chi(v)\nu(w)\,,$$

die wir $\chi\underline{\otimes}\nu$ nennen.

Mit dieser Definition ist dann

$$\begin{array}{rcl} V^* \times W^* & \to & (V \otimes W)^* \\ (\chi,\nu) & \mapsto & \chi\underline{\otimes}\nu \end{array}$$

bilinear (nachrechnen!), induziert also eine lineare Abbildung

$$\begin{array}{rrcl} \psi: & V^* \otimes W^* & \to & (V \otimes W)^* \\ & \chi \otimes \nu & \mapsto & \chi\underline{\otimes}\nu \end{array}$$

Dass diese für endlich-dimensionale V und W ein Isomorphismus ist, folgt mit Basen $(e_1,\dots,e_m)$ von V und $(f_1,\dots,f_n)$ von W: Sind $(e_1^*,\dots,e_n^*)$ und $(f_1^*,\dots,f_n^*)$ die zugehörigen Dualbasen von V^* bzw. W^*, und ist $((e_i \otimes f_j)^*)_{i,j}$ die Dualbasis zur Basis $(e_i \otimes f_j)_{i,j}$ von $V \otimes W$, so rechnet man nach, dass ψ gerade $e_i^* \otimes f_j^*$ auf $(e_i \otimes f_j)^*$ abbildet. Da $(e_i^* \otimes f_j^*)_{i,j}$ eine Basis von $V^* \otimes W^*$ bildet, folgt, dass ψ ein Isomorphismus ist.

Bemerkung 12.16 Nehmen wir die Regeln 12.13 bis 12.15 zusammen, so erhalten wir viele weitere Formeln wie zum Beispiel

$$\mathrm{Hom}(U,V) \otimes W \cong (U^* \otimes V) \otimes W \cong U^* \otimes (V \otimes W) \cong \mathrm{Hom}(U, V \otimes W)$$

oder

$$\begin{aligned} &\mathrm{Hom}(V_1,W_1) \otimes \mathrm{Hom}(V_2,W_2) \cong V_1^* \otimes W_1 \otimes V_2^* \otimes W_2 \cong V_1^* \otimes V_2^* \otimes W_1 \otimes W_2 \\ &\cong (V_1 \otimes V_2)^* \otimes (W_1 \otimes W_2) \cong \mathrm{Hom}(V_1 \otimes V_2, W_1 \otimes W_2)\,, \end{aligned}$$

für *endlich-dimensionale* Vektorräume.

§13 Skalarerweiterungen und die Komplexifizierung eines reellen Vektorraums

Wir geben noch die folgende Anwendung des Tensorprodukts.

Sei K ein Körper und V ein K-Vektorraum. Sei L eine **Körpererweiterung** von K, d.h., $K \subseteq L$, L ist ein Körper und K ist ein Teilkörper von L (Die Verknüpfungen $+$ und $\cdot$ in L setzen die Verknüpfungen $+$ und $\cdot$ von K fort). Man sagt auch L/K ist eine Körpererweiterung. Das besondere Beispiel, das wir im Auge haben, ist die Körpererweiterung $\mathbb{C}/\mathbb{R}$.

Dann ist insbesondere L ein K-Vektorraum, und wir können den K-Vektorraum

$$L \otimes_K V$$

bilden.

Lemma/Definition 13.1 $L \otimes_K V$ wird zu einem L-Vektorraum durch die Skalarmultiplikation

$$\lambda(\mu \otimes v) = \lambda\mu \otimes v\,. \tag{13.1.1}$$

Dieser L-Vektorraum heißt die **Skalarerweiterung** von V (bezüglich L/K, oder von K nach L). Im Fall eines reellen Vektorraums V heißt der komplexe Vektorraum $\mathbb{C} \otimes_{\mathbb{R}} V$ auch die **Komplexifizierung** von V.

Beweis Die Definition (13.1.1) ist wie üblich zu verstehen: für jedes $\lambda \in L$ induziert die (K-)bilineare(!) Abbildung

$$\begin{array}{rccl} \varphi_\lambda : & L \times V & \to & L \otimes_K V \\ & \mu \times v & \mapsto & \lambda\mu \otimes v \end{array}$$

eine eindeutig bestimmte lineare Abbildung

$$\begin{array}{rccl} \psi_\lambda\,(=\lambda\cdot): & L \otimes_K V & \to & L \otimes_K V \\ \text{mit} \qquad & \mu \otimes v & \mapsto & \lambda\mu \otimes v\,. \end{array}$$

Wir setzen nun $\lambda \cdot w = \psi_\lambda(w)$ für $\lambda \in L$ und $w \in L \otimes_K V$. Also ist tatsächlich $\lambda \cdot (\mu \otimes v) = \lambda\mu \otimes v$ für $\mu \in L$ und $v \in V$, und für ein beliebiges Element $\sum\limits_{i=1}^{r} \lambda_i \otimes v_i \in L \otimes_K V$ und $\lambda \in L$ gilt

$$\lambda(\sum_{i=1}^{r} \lambda_i \otimes v_i) = \sum_{i=1}^{r} \lambda\lambda_i \otimes v_i\,. \tag{13.1.2}$$

Die Vektorraum-Axiome lassen sich nun leicht nachrechnen. Z.B. folgt die Formel

$$\lambda(w_1 + w_2) = \lambda w_1 + \lambda w_2 \quad (\lambda \in K, w_1, w_2 \in L \otimes V)$$

aus der Linearität von ψ_λ (oder der Formel (13.1.2)).

Bemerkung 13.2 Oft schreibt man auch kurz V_L für $L \otimes_K V$, also zum Beispiel $V_{\mathbb{C}}$ für die Komplexifizierung eines reellen Vektorraums V.

Lemma 13.3 Ist $(e_i)_{i\in I}$ eine (K-) Basis von V, so ist $(1 \otimes e_i)_{i\in I}$ eine (L-)Basis von $L \otimes_K V$. Insbesondere gilt

$$\dim_K V = \dim_L V_L\,.$$

Beweis $(1 \otimes e_i)_{i\in I}$ ist ein Erzeugendensystem von V_L, denn für $\lambda \in K$ und $v \in V$ gibt es $(a_i)_{i\in I} \in K^{(I)}$ mit $v = \sum\limits_i a_i e_i$, und dann gilt

$$\lambda \otimes v = \lambda \otimes \sum_i a_i e_i = \sum_i a_i \lambda \otimes e_i = \sum_i a_i \lambda \cdot (1 \otimes e_i)\,.$$

Jedes Element in V_L ist aber von der Form

$$\sum_{i=1}^{r} \lambda_i \otimes v_i$$

mit $\lambda_i \in L$ und $v_i \in V$ (siehe 12.9 (b)).

Weiter sind die $1 \otimes e_i$ linear unabhängig: Sei $(\lambda_i)_{i\in I} \in L^{(I)}$ mit

$$0 = \sum_i \lambda_i \cdot (1 \otimes e_i) = \sum_i \lambda_i \otimes e_i \,.$$

Sei $(\ell_j)_{j\in J}$ eine K-Basis von L (existiert!), und sei

$$\lambda_i = \sum_{j\in J} a_{ij}\ell_j$$

mit $(a_{ij})_j \in K^{(J)}$ (für alle $i \in I$). Dann folgt

$$0 = \sum_i \lambda_i \otimes e_i = \sum_{i\in I}\sum_{j\in J} a_{ij}\ell_j \otimes e_i \,.$$

Es folgt $a_{ij} = 0$ für alle $(i,j) \in I \times J$ (da die $\ell_j \otimes e_i$ eine Basis von $L \otimes_K V$ bilden) und damit $\lambda_i = 0$ für alle $i \in I$.

Bemerkungen 13.4 (a) Wenn wir e_i mit $1 \otimes e_i$ identifizieren, hat also V_L "diesselbe Basis" wie V (wobei man bei V eine K-Basis und bei V_L eine L-Basis hat). Dies wird auch manchmal als (unkanonische) Definition der Skalarerweiterung genommen.

(b) Durch Einschränkung ist V_L natürlich auch ein K-Vektorraum (das ist gerade die K-Vektorraumstruktur von $L \otimes V$). Die Abbildung

$$\begin{array}{ccc} V & \to & L \otimes_K V \\ v & \mapsto & 1 \otimes v \end{array}$$

ist eine injektive K-lineare Abbildung (injektiv, da sie eine Basis $(e_i)_i$ von V auf die K-linear unabhängige Familie $(1 \otimes e_i)_i$ abbildet).

(c) Es gilt nach 12.10

$$\dim_K V_L = \dim_K L \otimes_K V = \dim_K L \cdot \dim_K V = [L : K] \cdot \dim_K V \,,$$

wobei $[L : K] := \dim_K L$ der **Grad** der Körpererweiterung L/K ist.

Zum Beispiel gilt für einen reellen Vektorraum V

$$\dim_{\mathbb{C}} V_{\mathbb{C}} = \dim V, \quad \dim_{\mathbb{R}} V_{\mathbb{C}} = 2 \dim V \,,$$

da $[\mathbb{C} : \mathbb{R}] = 2$ (eine $\mathbb{R}$-Basis von $\mathbb{C}$ ist $(1, i = \sqrt{-1})$). Die Bedeutung der Skalarerweiterung liegt in den folgenden beiden Resultaten.

Proposition 13.5 Sei $\varphi : V \to V'$ eine lineare Abbildung von K-Vektorräumen. Dann induziert φ eine kanonische lineare Abbildung von L-Vektorräumen

$$\varphi_L : V_L \to V_L \qquad \text{mit } \lambda \otimes v \mapsto \lambda \otimes \varphi(v) \,.$$

Beweis Die Abbildung

$$\begin{array}{rcl} L \times V & \to & L \otimes_K V \\ (\lambda, v) & \mapsto & \lambda \otimes_K \varphi(v) \end{array}$$

ist K-bilinear und induziert die angegebene K-lineare Abbildung $\varphi_L : L \otimes V \to L \otimes V$. Diese ist auch L-linear, wie man leicht nachrechnet.

Satz 13.6 Seien V und W endlich-dimensionale K-Vektorräume, mit Basen $e = (e_1, \ldots, e_m)$ und $f = (f_1, \ldots, f_n)$, und sei $\varphi : V \to W$ eine $(K\text{-})$lineare Abbildung mit Matrixdarstellung $A = M_e^f(\varphi) = M(m \times n, K)$ in diesen Basen. Dann ist A auch die Darstellung von $\varphi_L : V_L \to W_L$ bezüglich der Basen $e_L = (1 \otimes e_1, \ldots, 1 \otimes e_m)$ und $f_L = (1 \otimes f_1, \ldots, 1 \otimes f_n)$.

Beweis $(a_{ij}) = M_e^f(\varphi) \quad \Leftrightarrow \quad \varphi(e_j) = \sum a_{ij} f_i$ für alle $i \quad \Rightarrow \quad \varphi_L(1 \otimes e_j) = \sum a_{ij}(1 \otimes e_i)$ für alle $i \quad \Leftrightarrow \quad (a_{ij}) = M_{e_L}^{f_L}(\varphi_L)$.

Diese beiden Resultate verallgemeinern die Tatsache, dass man eine Matrix in $M(m \times n, K)$ auch als Matrix in $M(m \times n, L)$ auffassen kann, auf die allgemeinere Situation von Vektorräumen und linearen Abbildungen. Gleichzeitig zeigt Satz 13.6, dass dies im Fall von Matrizen das Bekannte wiedergibt.

§14 Die Normalform von unitären und orthogonalen Matrizen

Satz 14.1 Jede unitäre Matrix $U \in M_n(\mathbb{C})$ ist diagonalisierbar. Genauer gibt es eine unitäre Matrix T mit

$$T^* U T = T^{-1} U T = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & 0 & \\ & 0 & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

diagonal. Für die Eigenwerte λ gilt $|\lambda| = 1$.

Wir betrachten zunächst allgemeiner die folgende Situation. Sei $\mathbb{K} = \mathbb{C}$ und $(V, <,>)$ ein unitärer Vektorraum (d.h., V ein $\mathbb{C}$-Vektorraum und $<,>$ ein hermitesches Skalarprodukt), oder $\mathbb{K} = \mathbb{R}$ und $(V, <,>)$ ein euklidischer Vektorraum (d.h., V ein $\mathbb{R}$-Vektorraum und $<,>$ ein euklidisches Skalarprodukt).

Für $\lambda \in \mathbb{K}$ ist $\overline{\lambda}$ das komplex Konjugierte (also $\overline{\lambda} = \lambda$ wenn $\mathbb{K} = \mathbb{R}$).

Erinnerung: Für einen $\mathbb{K}$-Vektorraum $U \subseteq V$ war sein orthogonales Komplement

$$U^\perp = \{v \in V \mid < u, v >= 0 \quad \forall\, u \in U\},$$

und es galt

$$V = U \perp U^\perp$$

(orthogonale Summe). Insbesondere gilt $V = U \oplus U^\perp$ und somit

$$\dim V = \dim U + \dim U^\perp .$$

(LA I 19.24).

Lemma 14.2 Sei $\varphi : V \to V$ eine unitäre (bzw. orthogonale) Abbildung (also φ linear und $< \varphi(v), \varphi(v') >=< v, v' >$ für alle $v, v' \in V$). Sei $U \subseteq V$ ein Unterraum mit $\varphi(U) \subseteq U$. Dann gilt $\varphi(U^\perp) \subseteq U^\perp$.

Beweis Wir beweisen zuerst, dass φ injektiv ist: $\varphi(v) = 0 \Rightarrow 0 =< \varphi(v), \varphi(v') >=< v, v' >$ für alle $v' \in V \Rightarrow v = 0$. Die injektive lineare Abbildung

$$\varphi : U \to U$$

ist dann ein Isomorphismus, aus Dimensionsgründen. Ist nun $v \in U^\perp$ und $u \in U$, so gibt es ein $u' \in U$ mit $\varphi(u') = u$, und damit folgt

$$< u, \varphi(v) > = < \varphi(u'), \varphi(v) > = < u', v > = 0\,,$$

also $\varphi(v) \in U^\perp$, da $u \in U$ beliebig war.

Beweis von Satz 14.1: Für die ersten Aussagen haben wir zu zeigen, dass $V = \mathbb{C}^n$ eine Orthonormalbasis $(v_1, \ldots, v_n)$ aus Eigenvektoren von U hat. Dann ist nämlich die Matrix $T = (v_1 \mid \ldots \mid v_n)$ mit den Spalten v_i unitär, und es gilt

$$T^{-1}UT = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & 0 & \\ & 0 & \ddots & \\ & & & \lambda_n \end{pmatrix}\,,$$

wenn λ_i der Eigenwert zu v_i ist (siehe die Argumente beim Spektralsatz für *hermitesche* Matrizen). Satz 14.1 folgt also aus

Satz 14.3 Sei $(V, <,>)$ ein endlich-dimensionaler unitärer Raum, und sei $\varphi : V \to V$ eine unitäre Abbildung. Dann besitzt V eine Orthonormalbasis aus Eigenvektoren für φ. Ist $\lambda \in \mathbb{C}$ ein Eigenwert von φ, so gilt $|\lambda| = 1$.

Beweis Die letzte Aussage ist klar: Ist v ein Eigenvektor von φ zum Eigenwert λ, so gilt

$$< v, v > = < \varphi(v), \varphi(v) > = < \lambda v, \lambda v > = \lambda\overline{\lambda} < v, v > \;.$$

Wegen $< v, v > > 0$ folgt $1 = \lambda\overline{\lambda} = |\lambda|^2$, also $|\lambda| = 1$. Für die ersten Aussagen führen wir Induktion über $\dim V$, wobei der Fall $\dim V = 0$ trivial ist. Für $\dim V > 0$ sei $\lambda \in \mathbb{C}$ ein Eigenwert von φ (ein solcher existiert, da über $\mathbb{C}$ das charakteristische Polynom von φ eine Nullstelle hat). Sei v ein Eigenvektor zum Eigenwert λ, und sei $U =< v >_{\mathbb{C}} = \mathbb{C} \cdot v$ der 1-dimensionale Unterraum von V, der von v erzeugt wird. Wegen $\varphi(v) = \lambda v$ gilt dann $\varphi(U) \subseteq U$, nach Lemma 14.2 also auch $\varphi(U^\perp) \subseteq U^\perp$. Durch Einschränkung induziert φ also eine unitäre Abbildung

$$\varphi : U^\perp \to U^\perp$$

des unitären Vektorraums $(U^\perp, <,>)$. Weiter gilt nach den Vorbemerkungen $\dim U^\perp = \dim V - \dim U = \dim V - 1$. Nach Induktionsannahme besitzt also $U^\perp$ eine Orthonormalbasis $v_2, \ldots, v_n$ aus Eigenvektoren für φ, und wir erhalten die gewüschte Orthonormalbasis für V durch $v_1 = \frac{1}{\|v\|}v, v_2, \ldots, v_n$ (beachte dass $< v_1, v_i >= 0$ für $i = 2, \ldots, n$ weil $v_i \in U^\perp$).

Wir wenden uns nun den orthogonalen Abbildungen/Matrizen zu.

Satz 14.4 Sei $S \in M_n(\mathbb{R})$ eine orthogonale Matrix. Dann gibt es eine orthogonale Matrix $T \in M_n(\mathbb{R})$ mit

$$(14.4.1) \qquad T^tST = T^{-1}ST = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & 0 & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & 0 & & & & A_1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & A_s \end{pmatrix}$$

(Blockdiagonalform), wobei

$$A_i = \begin{pmatrix} \cos\varphi_i & -\sin\varphi_i \\ \sin\varphi_i & \cos\varphi_i \end{pmatrix} \in M_2(\mathbb{R})$$

mit $\varphi_i \in \,]0, 2\pi[\,$, $\varphi_i \neq \pi$. Wir nennen dies die **Normalform** einer orthogonalen Matrix.

Wie oben folgt dieser Satz aus

Satz 14.5 Sei $(V, <, >)$ ein endlich-dimensionaler euklidischer Vektorraum und $\varphi : V \to V$ eine orthogonale Abbildung. Dann gibt es eine Orthogonalbasis $v_1, \ldots, v_n$ von V, bezüglich derer die Matrixdarstellung von φ die obige Normalform hat.

Beweis von Satz 14.5 Die Strategie ist, wie im Beweis für unitäre Abbildungen zu verfahren (insbesondere mit Induktion über die Dimension). Weil φ aber möglicherweise keinen reellen Eigenwert hat, betrachten wir eine zugehörige unitäre Abbildung. Sei $V_{\mathbb{C}}$ die Komplexifizierung von V und $\varphi_{\mathbb{C}} : V_{\mathbb{C}} \to V_{\mathbb{C}}$ die von φ induzierte $\mathbb{C}$-lineare Abbildung ($\varphi(\lambda \otimes v) = \lambda \otimes \varphi(v)$ für $\lambda \in \mathbb{C}$ und $v \in V$). Jedes Element $x \in V_{\mathbb{C}} = \mathbb{C} \otimes_{\mathbb{R}} V$ lässt sich eindeutig schreiben als

$$(14.5.1) \qquad x = 1 \otimes v + i \otimes w$$

mit $v, w \in V$ weil $(1, i)$ eine $\mathbb{R}$-Basis von $\mathbb{C}$ ist (vergleiche auch die Übungsaufgabe 36 (i) und 40):

Ist nämlich $(e_1, \ldots, e_n)$ eine Basis von V, so bilden nach 12.10 die Tensorprodukte

$$1 \otimes e_j,\ i \otimes e_j \quad (j = 1, \ldots, n)$$

eine $\mathbb{R}$-Basis von $V_{\mathbb{C}} = \mathbb{C} \otimes V$. Für jedes $x \in V_{\mathbb{C}}$ gibt es also eindeutig bestimmte $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{R}$ mit

$$(14.5.2) \qquad \begin{aligned} x &= \sum_{j=1}^n a_j(1 \otimes e_j) + \sum_{j=1}^n b_j(i \otimes e_j) \\ &= \sum_{j=1}^n 1 \otimes a_je_j + \sum_{j=1}^n i \otimes b_je_j \\ &= 1 \otimes v + i \otimes w\,, \end{aligned}$$

wobei

$$v = \sum_{j=1}^{n} a_j e_j,\ w = \sum_{j=1}^{n} b_j e_j \in V\,. \tag{14.5.3}$$

Dies zeigt die Existenz der Darstellung. Ist umgekehrt $x = 1 \otimes v + i \otimes w$ mit $v, w \in V$, so finden wir $a_j, b_j \in \mathbb{R}$ mit (14.5.3), und die Gleichung (14.5.2) zusammen mit 12.10 zeigt, dass die a_j und b_j und damit auch v und w eindeutig sind.

Wir behaupten nun, dass es auf dem komplexen Vektorraum $V_{\mathbb{C}}$ ein *hermitesches* Skalarprodukt $<,>_{\mathbb{C}}$ gibt mit der Eigenschaft

$$< \lambda \otimes v, \mu \otimes w >_{\mathbb{C}} = \lambda\overline{\mu} < v, w > \tag{14.5.4}$$

für $\lambda, \mu \in \mathbb{C}$ und $v, w \in V$, wobei rechts das gegebene Skalarprodukt von V steht. In der Tat, definiere $<,>_{\mathbb{C}}$ durch

$$< 1 \otimes v + i \otimes w,\ 1 \otimes v' + i \otimes w' >_{\mathbb{C}} = < v, v' > \ + \ < w, w' > \ . \tag{14.5.5}$$

Wegen der eindeutigen Darstellung (14.5.1) ist dies wohldefiniert (und ergibt sich zwingend aus der gewünschten Eigenschaft (14.5.4)). Man sieht leicht, dass die Eigenschaft (14.5.4) erfüllt ist (nachrechnen!) und dass $<,>_{\mathbb{C}}$ additiv in beiden Argumenten ist. Hieraus folgt, wiederum leicht, dass $<,>_{\mathbb{C}}$ ein hermitesches Skalarprodukt ist. Dass $<,>_{\mathbb{C}}$ positiv definiert ist, folgt aus der Positiv-Definiertheit von $<,>$: Ist $x = 1 \otimes v + i \otimes w \neq 0$, so ist $v \neq 0$ oder $w \neq 0$, also $< x, x >_{\mathbb{C}} = < v, v > + < w, w > > \ 0$.

Weiter ist $\varphi_{\mathbb{C}}$ unitär bezüglich $<,>_{\mathbb{C}}$, denn es ist

$$\begin{aligned} < \varphi(\lambda \otimes v), \varphi(\mu \otimes w) > \ &= \ < \lambda \otimes \varphi(v), \mu \otimes \varphi(w) > = \lambda\overline{\mu} < \varphi(v), \varphi(w) > \\ &= \ \lambda\overline{\mu} < v, w > = < \lambda \otimes v, \mu \otimes w > \end{aligned}$$

für $\lambda, \mu \in \mathbb{C},\ v, w \in V$ (dies verallgemeinert die Tatsache, dass jede reelle orthogonale Matrix als komplexe Matrix aufgefasst unitär ist).

Wir beweisen nun 14.5 durch Induktion über $\dim V$, wobei der Fall $\dim V = 0$ trivial ist.

Für $\dim V > 0$ gibt es einen Eigenvektor $x \in V_{\mathbb{C}}$ von $\varphi_{\mathbb{C}}$. Sei

$$x = 1 \otimes v + i \otimes w$$

$(v, w \in V)$ die eindeutige Darstellung, und sei der zugehörige Eigenwert $\lambda = \alpha + \beta i \quad (\alpha, \beta \in \mathbb{R})$. Dann folgt

$$\begin{aligned} &1 \otimes \varphi(v) + i \otimes \varphi(w) = \varphi(x) = \lambda x = (\alpha + \beta i)(1 \otimes v + i \otimes w) \\ &= \alpha \otimes v + \alpha i \otimes w + \beta i \otimes v - \beta \otimes w = 1 \otimes (\alpha v - \beta w) + i \otimes (\beta v + \alpha w)\,. \end{aligned}$$

Wegen der Eindeutigkeit der Darstellung gilt

$$\begin{aligned} \varphi(v) &= \alpha v - \beta w \\ \varphi(w) &= \beta v + \alpha w\,. \end{aligned}$$

Es ist also $U = < v, w >_{\mathbb{R}} \subseteq V$ ein Unterraum von mit $\varphi(U) \subseteq U$.

Ist $\dim U = 1$, so ist jeder Vektor $0 \neq u \in U$ ein Eigenvektor zu einem reellen Eigenwert $\lambda \in \mathbb{R}$. Es muss dann $\lambda = 1$ oder $\lambda = -1$ sein, wegen

$$< u, u >_{\mathbb{C}} = < \varphi(u), \varphi(u) >_{\mathbb{C}} = < \lambda(u), \lambda(u) >_{\mathbb{C}} = \lambda^2 < u, u >_{\mathbb{C}} \; .$$

Weiter können wir annehmen, das u normiert ist.

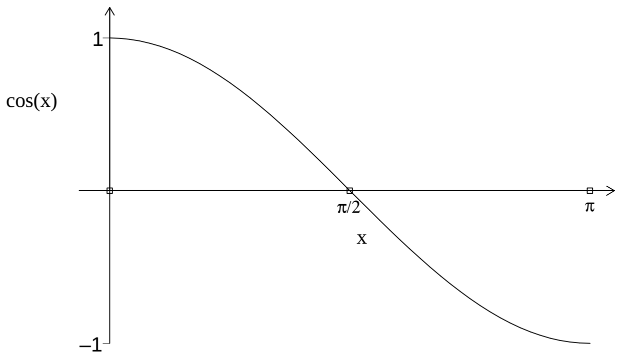
Ist $\dim U = 2$, so sei (u_1, u_2) eine Orthonormalbasis von U. Wegen der Orthogonalität von $\varphi : U \to U$ ist dann die Matrixdarstellung von φ auf U bezüglich (u_1, u_2) eine orthogonale Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} .$$

Es muss also gelten $a^2 + c^2 = 1 = b^2 + d^2$ sowie $ab + cd = 0$. Es folgt dann aus der Analysis, dass es einen eindeutig bestimmten Winkel $\varphi \in [0, 2\pi[$ gibt mit

$$a = \cos \varphi \quad , \quad c = \sin \varphi \, .$$

Denn es ist $|a| \leq 1$, und daher gibt es ein eindeutiges $\alpha \in [0, \pi]$ mit $a = \cos \alpha$.



Es ist $c^2 = 1 - a^2 = 1 - \cos^2 \alpha = \sin^2 \alpha$, also $c = \pm \sin \alpha$. Ist $c = \sin \alpha$, so setze $\varphi = \alpha$. Ist $c = -\sin \alpha \neq 0$, so setze $\varphi = 2\pi - \alpha$; dann ist $\cos \varphi = \cos(-\alpha) = \cos \alpha = a$ und $\sin \varphi = \sin(-\alpha) = -\sin \alpha = c$, und diese Wahl von φ ist eindeutig.

Wegen $\begin{pmatrix} a \\ c \end{pmatrix} \cdot \begin{pmatrix} b \\ d \end{pmatrix} = 0$ (übliches Skalarprodukt in $\mathbb{R}^2$) gilt

$$\begin{pmatrix} b \\ d \end{pmatrix} = \pm \begin{pmatrix} -c \\ a \end{pmatrix} ,$$

denn für den 1-dimensionalen Raum $U = \mathbb{R} \begin{pmatrix} a \\ c \end{pmatrix} \subseteq \mathbb{R}^2$ gilt $\dim U^\perp = 1$ und $\begin{pmatrix} -c \\ a \end{pmatrix} \in U^\perp$, also

$$\begin{pmatrix} b \\ d \end{pmatrix} = \lambda \begin{pmatrix} -c \\ a \end{pmatrix} \text{ mit } \lambda \in \mathbb{R} \, .$$

Da beide Vektoren den Betrag 1 haben ($a^2 + c^2 = 1 = b^2 + d^2$) folgt $\lambda = \pm 1$.

Ist $\lambda = 1$, so ist also

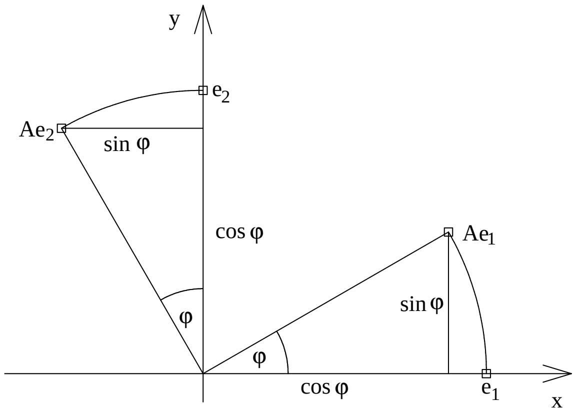
$$A = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} ,$$

wobei für $\varphi = 0, \pi$ die Matrizen

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

herauskommen.

Geometrisch entspricht A einer Drehung um den Winkel φ



Ist $\lambda = -1$, so ist

$$A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} .$$

Das charakteristische Polynom von A ist dann

$$x^2 - \cos^2\varphi - \sin^2\varphi = x^2 - 1 .$$

Es folgt, dass A die Eigenwerte $+1, -1$ hat, und dasselbe gilt dann auch für $\varphi : U \to U$.

Seien v_1 und v_2 normierte Eigenvektoren zu den Eigenwerten $+1$ beziehungsweise -1. Nach dem folgenden Lemma ist dann (v_1, v_2) eine Orthonormalbasis, und die Darstellungsmatrix in dieser Basis ist

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

Wir haben also in jedem Fall eine Orthonormalbasis von U gefunden, so dass $\varphi_{|\,U}$ in dieser Basis die gewünschte Normalform (14.4.1) hat. Weiter gilt $\dim U^\perp = \dim V - \dim U < \dim V$ und $\varphi(U^\perp) \subseteq U^\perp$, und nach Induktionsvoraussetzung hat $\varphi : U^\perp \to U^\perp$ die gewünschte Normalform bezüglich einer Orthonormalbasis von $U^\perp$.

Wegen $V = U \perp U^\perp$ erhalten wir insgesamt eine Orthonormalbasis von V, in der φ die gewünschte Normalform hat (nach eventueller Umordnung der Basis). (Vergleiche Bemerkung 5.9)

Lemma 14.6 Ist $\varphi : V \to V$ eine unitäre oder orthogonale Abbildung und sind $\lambda \neq \mu$ zwei verschiedene Eigenwerte von φ, so gilt für die zugehörigen Eigenräume

$$V(\lambda) \perp V(\mu) ,$$

d.h., $V(\lambda)$ ist orthogonal zu $V(\mu)$.

Beweis (vergleiche LA I 19.19) Sei $v \in V(\lambda)$ und $w \in V(\mu)$. Dann gilt

$$< v, w > = < \varphi(v), \varphi(w) > = < \lambda v, \mu w > = \lambda\overline{\mu} < v, w > ,$$

also $< v, w > (1 - \lambda\overline{\mu}) = 0$. Wegen $|\mu| = 1$ gilt $\overline{\mu} = \mu^{-1}$, und aus $\lambda\overline{\mu} = 1$ würde $\lambda = \mu$ folgen, im Widerspruch zur Annahme. Also ist $(1 - \lambda\overline{\mu}) \neq 0$ und daher $< v, w >= 0$.

§15 Die orthogonale Gruppe

Definition 15.1 Für $n \in \mathbb{N}$ sei $O(n)$ die Gruppe der orthogonalen $(n \times n)$-Matrizen, also

$$O(n) = \{A \in M_n(\mathbb{R}) \mid A^t A = E\} .$$

Sie heißt die **orthogonale Gruppe** der Ordnung n.

Es handelt sich wirklich um eine Gruppe (unter dem Matrixprodukt), nämlich eine Untergruppe der Gruppe der invertierbaren Matrizen: Erstens ist $(A \cdot B)^t = B^t \cdot A^t$; für orthogonale Matrizen A und B ist also $(AB)^t AB = B^t A^t AB = B^t B = E$, also AB orthogonal. Zweitens ist für jede orthogonale Matrix A auch

$$AA^t = E ,$$

da aus der Beziehung $A^t A = E$ die Beziehung $A^t = A^{-1}$ folgt. Wegen $(A^t)^t = A$ ist also $(A^t)^t A^t = AA^t = E$, also $A^{-1} = A^t$ orthogonal.

Bemerkungen 15.2 (a) Sei A eine orthogonale Matrix. Die Beziehung $A^t A = E$ bedeutet, dass die Spalten von A eine Orthonormalbasis von $\mathbb{R}^n$ bilden. Wegen $AA^t = E$ bilden dann auch die Zeilen von A eine Orthonormalbasis.

(b) Dass $O(n)$ eine Gruppe ist, lässt sich auch daran sehen, dass $O(n)$ aus den Matrizen besteht, die das Standardskalarprodukt $<, >$ erhalten.

Wir wollen nun die orthogonalen Gruppen näher beschreiben und verstehen.

Lemma 15.3 Für eine orthogonale Matrix A gilt $\det A = \pm 1$.

Beweis $\quad A^t A = E \quad \Rightarrow \quad 1 = \det E = \det(A^t A) = \det(A^t) \cdot \det(A) = (\det(A))^2$.

Definition 15.4 Die Untergruppe

$$SO(n) = \{A \in O(n) \mid \det(A) = 1\}$$

der orthogonalen Matrizen mit Determinante 1 heißt die **spezielle orthogonale Gruppe**.

Wir betrachten nun $O(2)$ und $SO(2)$:

Satz 15.5 (a) Jedes $A \in SO(2)$ ist von der Form

$$A = A_\varphi = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

für ein eindeutig bestimmtes $\varphi \in [0, 2\pi[$. Das charakteristische Polynom von A_φ ist

$$(x - \cos\varphi)^2 + \sin^2\varphi\,;$$

es hat die komplexen Nullstellen

$$\cos\varphi \pm i \cdot \sin\varphi\,.$$

(b) Jedes $A \in O(2) \smallsetminus SO(2)$ ist von der Form

$$A'_\varphi = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} = A_\varphi \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

für ein eindeutig bestimmtes $\varphi \in [0, 2\pi[$. Das charakteristische Polynom von A'_φ ist

$$x^2 - 1 = (x+1)(x-1)\,.$$

Beweis Dies folgt aus dem Beweis von Satz 14.5: Die erste Spalte von A war von der Form

$$v = \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}$$

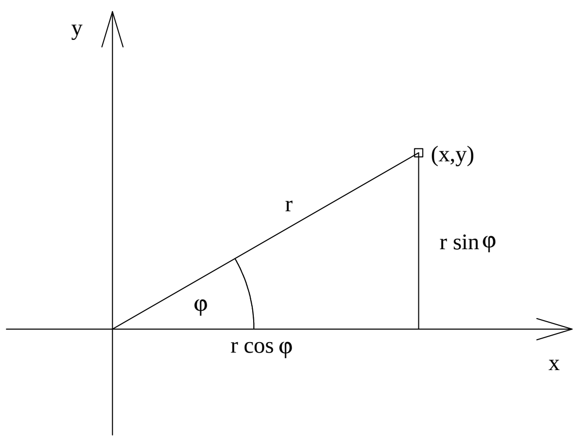
mit eindeutigem $\varphi \in [0, 2\pi[$, und die zweite Spalte von A war von der Form v' oder $-v'$, mit

$$v' = \begin{pmatrix} -\sin\varphi \\ \cos\varphi \end{pmatrix}\,.$$

Im ersten Fall ist $\det(A) = 1$ und im zweiten Fall ist $\det(A) = -1$.

Bemerkungen 15.6 (a) Jedes $v = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ lässt sich schreiben als

(15.6.1) $$v = r \cdot \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}$$



(b) Für

$$A_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

gilt

$$\begin{aligned} A_\alpha v &= r \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix} \\ &= r \begin{pmatrix} \cos\alpha\cos\varphi - \sin\alpha\sin\varphi \\ \sin\alpha\cos\varphi + \cos\alpha\sin\varphi \end{pmatrix} \\ &= r \begin{pmatrix} \cos(\alpha+\varphi) \\ \sin(\alpha+\varphi) \end{pmatrix} \end{aligned}$$

nach den Additionssätzen für sin und cos: Diese lassen sich am besten durch die **Euler-Formel**

$$e^{i\varphi} = \cos\varphi + i\sin\varphi \quad \in \mathbb{C} \quad \text{für } \varphi \in \mathbb{R} \tag{15.6.2}$$

und die Funktionalgleichung

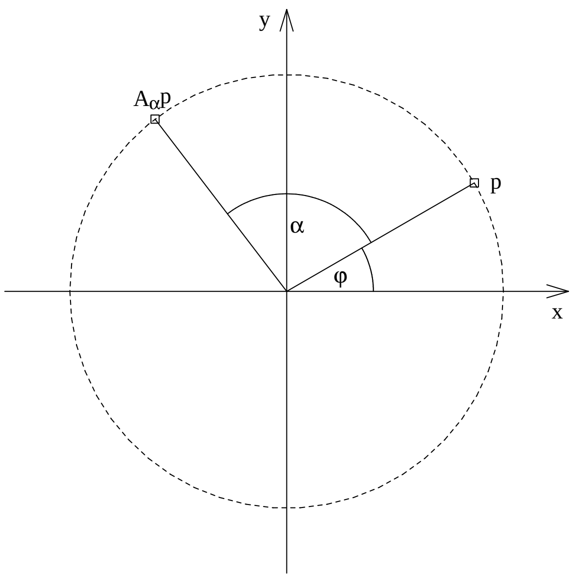
$$e^{z_1+z_2} = e^{z_1} \cdot e^{z_2} \quad \text{für } z_1, z_2 \in \mathbb{C}, \text{ also insbesondere} \tag{15.6.3}$$

$$e^{i(\varphi_1+\varphi_2)} = e^{i\varphi_1} e^{i\varphi_2} \quad \text{für } \varphi_1, \varphi_2 \in \mathbb{R}$$

für die komplexe Exponentialfunktion sehen:

$$\begin{aligned} & \cos(\alpha+\varphi) + i\sin(\alpha+\varphi) \\ = \ & e^{i(\alpha+\varphi)} = e^{i\alpha} \cdot e^{i\varphi} = (\cos\alpha + i\sin\alpha)(\cos\varphi + i\sin\varphi) \\ = \ & \cos\alpha\cos\varphi - \sin\alpha\sin\varphi + i(\sin\alpha\cos\varphi + \cos\alpha\sin\varphi)\,. \end{aligned}$$

Dies zeigt, dass A_α eine Darstellung um den Winkel α im mathematisch positiven Sinne (also “gegen den Uhrzeiger”) bewirkt, wie im Beweis von Satz 14.5 behauptet.



Dies liefert die folgende explizite Beschreibung von $SO(2)$ als Gruppe.

Satz 15.7 Sei $S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ der Einheitskreis in der komplexen Zahlenebene $\mathbb{C} = \mathbb{R} + i\mathbb{R}$. Dies ist eine Gruppe unter der Multiplikation, und wir haben Isomorphismen von Gruppen

$$\begin{array}{ccccc} \mathbb{R}/2\pi\mathbb{Z} & \overset{\varphi_1}{\underset{\sim}{\longrightarrow}} & S^1 & \overset{\varphi_2}{\underset{\sim}{\longrightarrow}} & SO(2) \\ \overline{\varphi} & \mapsto & e^{i\varphi} & \mapsto & A_\varphi\,. \end{array}$$

Insbesondere ist $SO(2)$ kommutativ.

Beweis Es ist klar, dass $(S^1, \cdot)$ eine Gruppe ist (wegen $|z_1 z_2| = |z_1|\,|z_2|$). Nach der Euler-Formel (15.6.2) (oder der Polardarstellung $z = |z| \cdot e^{i\varphi}$ für komplexe Zahlen) ist die Abbildung

$$\begin{array}{ccc} \mathbb{R} & \to & S^1 \\ \varphi & \mapsto & e^{i\varphi} \end{array}$$

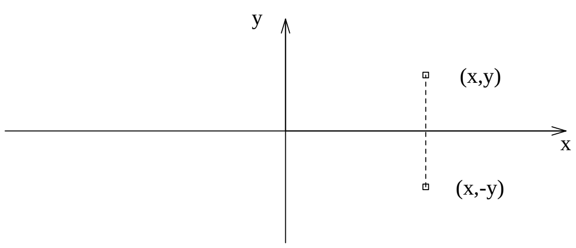
surjektiv. Weiter ist diese Abbildung nach der Funktionalgleichung (15.6.3) ein Gruppenhomomorphismus von $(\mathbb{R}, +)$ nach $(S^1, \cdot)$. Sein Kern ist $2\pi\mathbb{Z}$, weil $e^{i\varphi} = 1$ genau dann, wenn $\varphi \in 2\pi\mathbb{Z}$. Der erste Isomorphismus φ_1 ergibt sich also aus dem Homomorphiesatz 10.8 (b).

Die Additionssätze für cos und sin zeigen weiter, dass φ_2 ein Homomorphismus ist. Dieser ist bijektiv wegen 15.5 (a) und 15.6 (a).

Wir betrachten nun $O(2) \setminus SO(2)$ näher. Eine orthogonale Abbildung $A \in O(2) \setminus SO(2)$ ist nach Satz 15.5 (b) von der Form

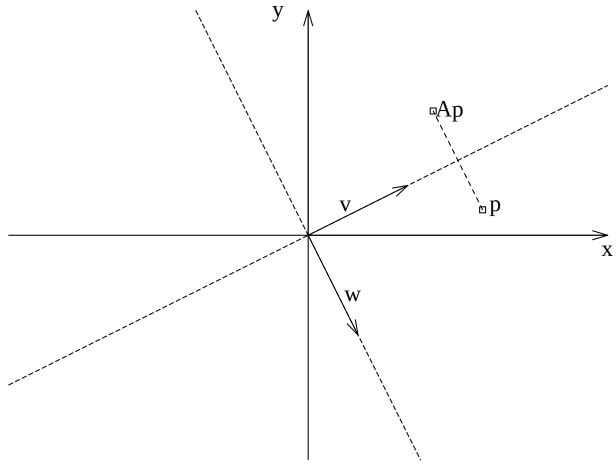
$$A'_\varphi = A_\varphi \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} A_{-\varphi}\,.$$

Weiter ist $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ offenbar die Spiegelung an der x-Achse



Also ist A'_φ die Komposition der Spiegelung an der x-Achse und der Drehung mit dem Winkel φ, oder der Drehung von $-\varphi$ und der Spiegelung an der x-Achse.

Wir können mehr sagen: Wir wissen, dass A'_φ die Eigenwerte $+1, -1$ hat. Es wird also ein Vektor $v \in \mathbb{R}^2$ festgelassen und ein zu v orthogonaler Vektor $w \neq 0$ auf $-w$ abgebildet (siehe 14.6: die Eigenräume zu 1 und -1 sind orthogonal!). Dies bedeutet: A_φ ist die Spiegelung an der Geraden $\mathbb{R}v$



Explizit können wir

$$v = \begin{pmatrix} \cos \frac{\varphi}{2} \\ \sin \frac{\varphi}{2} \end{pmatrix}, \; w = \begin{pmatrix} \sin \frac{\varphi}{2} \\ -\cos \frac{\varphi}{2} \end{pmatrix}$$

nehmen. Dies können wir am besten in der komplexen Schreibweise sehen: Identifizieren wir $\mathbb{R}^2$ mit $\mathbb{C}$, so ist A_φ die Abbildung

$$z \mapsto e^{i\varphi} \cdot z \,,$$

und $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, die Spiegelung an der reellen Achse, ist die Abbildung

$$z \mapsto \overline{z}\,.$$

Es ist also

$$A'_\varphi : z \mapsto e^{i\varphi}\overline{z}$$

Ein Fixvektor ist $z = e^{i\alpha}$ mit $e^{i\alpha} = e^{i(\varphi-\alpha)}$; dies gilt für $2\alpha = \varphi$, also $\alpha = \varphi/2$.

Wir betrachten nun $SO(3)$ und $O(3)$, also die orthogonalen Abbildungen des $\mathbb{R}^3$.

Sei $B \in O(3)$. Nach Satz 14.4/14.5 gibt es nach eventueller Umnummerierung der dort gefundenen Basis eine Orthonormalbasis von $\mathbb{R}^3$, so dass B in dieser Basis die Darstellung

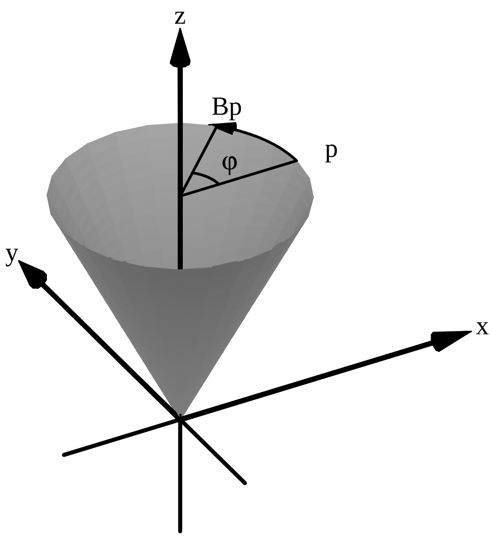
$$B^\pm_\varphi = \left(\begin{array}{cc|c} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & \\ \hline 0 & & \pm 1 \end{array}\right)$$

mit $\varphi \in [0, 2\pi[$ hat. Dies schließt die Fälle

$$\begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}, \begin{pmatrix} 1 & & \\ & 1 & \\ & & -1 \end{pmatrix}, \begin{pmatrix} 1 & & \\ & -1 & \\ & & -1 \end{pmatrix}, \begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix}$$

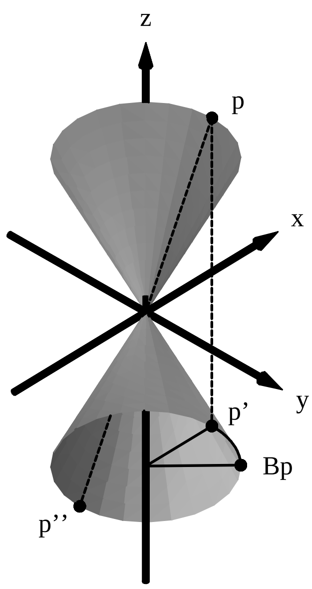
aus 14.4 ein: die erste Matrix ist B^+_0, die zweite ist B^-_0, die vierte ist B^-_π, und die dritte ist nach Umnummerierung der Basis gleich B^+_π.

Wir betrachten nun zuerst den Fall, dass B schon selbst gleich $B^\pm_\varphi$ ist. Dann sehen wir: $B = B^+_\varphi$ lässt die z-Achse fest und ist eine Drehung mit dem Winkel φ um diese Achse.



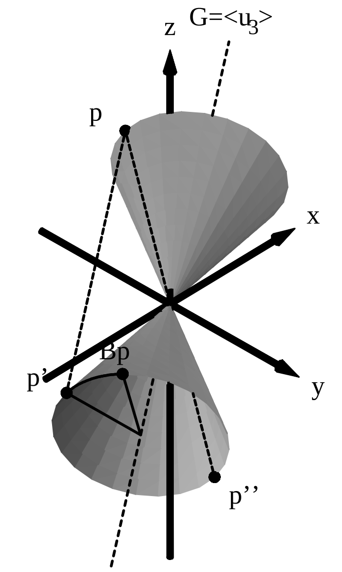
Dies ist der Fall $\det B = 1$.

$B = B_\varphi^-$ ist die Komposition aus einer Spiegelung an der $x-y$-Ebene $(z \mapsto -z)$ und einer Drehung um den Winkel φ um die z-Achse



Dies kann man auch erhalten, indem man die Komposition einer Spiegelung am Ursprung $(p \mapsto -p)$ und einer Drehung um die z-Achse mit dem Winkel $\pi + \varphi$ betrachtet (wegen $e^{i\varphi} = -e^{i(\pi+\varphi)}$).
Dies ist der Fall $\det B = -1$.

Im allgemeinen gilt diese Beschreibung von B in dem betrachteten Orthonormalsystem u_1, u_2, u_3). Man sieht dann: Ist $\det B = 1$, so ist B eine Drehung um eine Gerade (nämlich $\mathbb{R}u_3$). Ist $\det B = -1$, so ist B die Komposition einer solchen Drehung mit einer Spiegelung an der zur Geraden senkrechten Ebene (nämlich $< u_1, u_2 >_{\mathbb{R}}$).



Dies kann man auch direkt daraus herleiten, dass es für $n = 3$ immer eine Gerade G gibt, die in sich überführt wird (da das charakteristische Polynom einen reellen Eigenwert hat), und dass dann die zu G senkrechte Ebene $G^{\perp}$ ebenfalls in sich überführt wird. Für den Fall $n \geq 4$ kann man aber nicht so einfach argumentieren und man muss wirklich zeigen, dass es immer invariante Teilräume der Dimension 1 oder 2 gibt, wie in 14.5 bewiesen.

Im höherdimensionalen Fall ist dies genau die geometrische Aussage: $\mathbb{R}^n$ ist die orthogonale Summe von 1 oder 2-dimensionalen Unterräumen, die von $A \in O(n)$ respektiert werden, und A setzt sich aus entsprechenden Drehungen in diesen Ebenen und Spiegelungen in Richtung einiger dieser Geraden zusammen. Dabei gilt

Lemma 15.8 Ist n ungerade, so gibt es zu jedem $A \in O(n)$ eine invariante Gerade L (also $A(L) \subseteq L$), und jedes $A \in SO(n)$ lässt einen Vektor $v \neq 0$ fest.

Beweis Dies folgt aus der Normalform

$$\begin{pmatrix} 1 &&&&&&&& \\ & \ddots &&&&&&& \\ && 1 &&&&&& \\ &&& -1 &&&&& \\ &&&& \ddots &&&& \\ &&&&& -1 &&& \\ &&&&&& A_1 && \\ &&&&&&& \ddots & \\ &&&&&&&& A_s \end{pmatrix} \quad \text{mit } A_i \in SO(2) \subseteq M_2(\mathbb{R})\,.$$

Ist n ungerade, so muss es einen Eigenwert $+1$ oder -1 geben, also eine invariante Gerade. Ist $\det(A) = 1$, so ist wegen $\det(A_i) = 1 \quad (i = 1, \ldots, s)$ die Anzahl der Einträge -1 gerade; es muss also einen Eigenwert 1 geben.

§16 Bilinearformen

Sei K ein Körper und V ein K-Vektorraum.

Definition/Lemma 16.1 Sei $\mathrm{Bil}(V)$ die Menge der Bilinearformen auf V, also der bilinearen Abbildungen

$$\psi : V \times V \to K$$

(vergl. LA I 13.4). Dies ist ein Untervektorraum von $\mathrm{Abb}(V \times V, K)$.

Beweis Für Abbildungen $\psi, \psi' : V \times V \to K$ und $\lambda \in K$ sind $\psi + \psi'$ und $\lambda\psi$ definiert durch

$$\begin{aligned} (\psi + \psi')(v, v') &= \psi(v, v') + \psi'(v, v') \\ (\lambda\psi)(v, v') &= \lambda\psi(v, v') \end{aligned}$$

für alle $(v, v') \in V \times V$. Sind ψ und ψ' bilinear, so ist offensichtlich, dass $\psi + \psi'$ und $\lambda\psi$ wieder bilinear sind.

Definition 16.2 (vergl. LA I Def. 18.3) Sei V endlich-dimensional und $b = (b_1, \ldots, b_n)$ eine Basis von V. Für $\psi \in \mathrm{Bil}(V)$ heißt dann

$$B = M_b(\psi) := (\psi(b_i, b_j))_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}} \in M_n(K)$$

die **Fundamentalmatrix von ψ bezüglich** B.

Lemma 16.3 Für eine feste Basis b von V ist die Abbildung

$$\begin{aligned} M_b \ : \ \mathrm{Bil}(V) &\to M_n(K) \\ \psi &\mapsto M_b(\psi) \end{aligned}$$

ein Vektorraum-Isomorphismus.

Beweis (vergl. LA I 18.5) Es ist klar, dass diese Abbildung linear ist: $M_b(\psi + \psi') = M_b(\psi) + M_b(\psi')$ und $M_b(\lambda\psi) = \lambda M_b(\psi)$. Weiter erhalten wir eine Umkehrabbildung

$$\Psi_b : M_n(K) \to \mathrm{Bil}(V)\,,$$

indem wir eine Matrix $B = (b_{ij})$ auf die folgende Bilinearform $\psi = \Psi_b(B)$ abbilden:

$$\psi\left(\sum_{i=1}^n x_i b_i, \sum_{j=1}^n y_j b_j\right) = \sum_{i,j=1}^n x_i b_{ij} y_j = x^t B y\,,$$

wobei

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \ y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in K^n\,.$$

Denn: Da jedes $v \in V$ eindeutig als $v = \sum\limits_{i=1}^n \alpha_i b_i$ geschrieben werden kann, mit $\alpha_1, \ldots, \alpha_n \in K$, ist ψ hierdurch wohldefiniert. Weiter sieht man sofort, dass ψ bilinear ist. Schließlich ist die Fundamentalmatrix von ψ gleich B, da $\psi(b_i, b_j) = e_i^t B e_j = b_{ij}$, also $M_b \circ \Psi_b = \mathrm{id}$. Ebenso ist $\psi_b \circ M_b = \mathrm{id}$, denn für eine beliebige Bilinearform auf V gilt

$$\psi\left(\sum_{i=1}^n x_i b_i,\ \sum_{j=1}^n y_j b_j\right) = \sum_{i,j=1}^n x_i y_j \psi(b_i, b_j)\,.$$

Bemerkung 16.4 Ein eleganterer Beweis ergibt sich so: Wir haben durch die universelle Eigenschaft des Tensorprodukts eine bijektive Abbildung

$$\begin{array}{rcl} \alpha : \mathrm{Bil}(V) & \to & \mathrm{Hom}(V \otimes V, K) \\ \psi & \mapsto & \tilde{\psi} \text{ mit } \tilde{\psi}(v \otimes v') = \psi(v, v') \end{array}$$

Man sieht weiter, dass diese Abbildung *linear* ist, also ein Vektorraum-Isomorphismus.

Da $(b_i \otimes b_j)_{i,j=1,\ldots,n}$ eine Basis von $V \otimes V$ ist, folgt aus der universellen Eigenschaft von Basen (12.1), dass die Abbildung

$$\begin{array}{rcl} \beta \ : \ \mathrm{Hom}(V \otimes V, K) & \to & M_n(K) \\ \varphi & \mapsto & (\varphi(e_i \otimes e_j))_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}} \end{array}$$

bijektiv ist. Offenbar ist aber auch β *linear*, also ebenfalls ein Vektorraumisomorphismus. Schließlich folgt aus den Definitionen, dass

$$M_b = \beta \circ \alpha\,.$$

Beispiel 16.5 Auf K^n ist die **Standard-Bilinearform** definiert durch

$$\psi(x, y) = x^t y = \sum_{i=1}^n x_i y_i\,.$$

Ihre Fundamentalmatrix bezüglich der Standardbasis $e = (e_1, \ldots, e_n)$ ist die Einheitsmatrix. Jede andere Bilinearform ψ auf K^n ist von der Form

$$\psi_B(x, y) = x^t B y = \sum_{i=1}^n x_i b_{ij} y_j$$

für eine eindeutig bestimmte Matrix $B \in M_n(K)$ (B ist die Fundamentalmatrix von ψ_B bezüglich e).

Fundamentalmatrizen transformieren sich bei Basiswechsel anders als Darstellungsmatrizen von linearen Abbildungen:

Lemma 16.6 Sei $b' = (b'_1, \ldots, b'_n)$ eine weitere Basis von V, und sei

$$T = M_{b'}^b$$

die Transformationmatrix von b und b'. Sei $\psi \in \mathrm{Bil}(V)$, und seien B und B' die Fundamentalmatrizen von ψ bezüglich b und b'. Dann ist

$$B' = T^t B T\,.$$

Beweis Ist $T = (t_{ij})$, so gilt nach Definition

$$b'_j = \sum_{i=1}^{n} t_{ij} b_i \qquad (j = 1, \dots, n)\,.$$

Ist nun $B = (b_{ij})$ und $B' = (b'_{ij})$, so gilt

$$\begin{aligned}
B'_{k\ell} = b'_{k\ell} &= \psi(b'_k, b'_\ell) = \psi(\sum_{i=1}^{n} t_{ik} b_i, \sum_{j=1}^{n} t_{j\ell} b_j) \\
&= \sum_{i,j=1}^{n} t_{ik} t_{j\ell} \psi(b_i, b_j) \\
&= \sum_{i,j=1}^{n} t_{ik} b_{ij} t_{j\ell} \\
&= \sum_{i=1}^{n} t_{ik} \sum_{j=1}^{n} b_{ij} t_{j\ell} = \sum_{i=1}^{n} t_{ik} (BT)_{i\ell} \\
&= (T^t BT)_{k\ell}\,.
\end{aligned}$$

Lemma 16.7 Die Abbildung

$$\begin{array}{rcl}
\mathrm{Bil}(V) & \to & \mathrm{Hom}(V, V^*) \\
\psi & \mapsto & \begin{array}{rcll} \varphi_\psi : & V & \to & V^* \\ & w & \mapsto & \left(\begin{array}{rrcl} \psi(-, w) : & V & \to & K \\ & v & \mapsto & \psi(v, w) \end{array} \right) \end{array}
\end{array}$$

ist ein Vektorraum-Isomorphismus (wobei $V^* = \mathrm{Hom}(V, K)$ der Dualraum von V ist).

Beweis Die Abbildung ist linear, und die Umkehrabbildung ist

$$\begin{array}{rcl}
(\varphi : V \to V^*) & \mapsto & \psi_\varphi : V \times V \to K \\
& & \psi_\varphi(v, w) = \varphi(w)(v)\,.
\end{array}$$

Bemerkung 16.8 Sei V endlich-dimensional, mit Basis $b = (b_1, \dots, b_n)$, und sei $b^* = (b_1^*, \dots, b_n^*)$ die Dualbasis von V^*. Ist $B = M_b(\psi)$ die Fundamentalmatrix von ψ bezüglich b, so ist B auch die darstellende Matrix von $\varphi_\psi : V \to V^*$ bezüglich der Basen b und b^*, also $B = M_b^{b^*}(\varphi_\psi)$. Es gilt nämlich für $v = \sum_{i=1}^{n} x_i b_i$ und $B = (b_{ij})$:

$$\begin{aligned}
\varphi_\psi(b_j)(v) &= \psi(v, b_j) = \psi(\sum_{i=1}^{n} x_i b_i, b_j) \\
&= \sum_{i=1}^{n} \psi(b_i, b_j) x_i = \sum_{i=1}^{n} b_{ij} x_i \\
&= \sum_{i=1}^{n} b_{ij} b_i^*(v)\,,
\end{aligned}$$

da $b_r^*(v) = x_r$. Weil v beliebig war, folgt

$$\varphi_\psi(b_j) = \sum_{i=1}^{n} b_{ij} b_i^*\,,$$

also $M_b^{b^*}(\varphi_\psi) = B$.

Lemma/Definition 16.8 Sei V endlich-dimensional. Eine Bilinearform $\psi : V \times V \to K$ heißt nicht-ausgeartet, wenn die folgenden äquivalenten Bedingungen gelten:

(i) Für jedes $w \in V \smallsetminus \{0\}$ gibt es ein $v \in V$ mit

$$\psi(v, w) \neq 0 \,.$$

(ii) Für jedes $v \in V \smallsetminus \{0\}$ gibt es ein $w \in V$ mit

$$\psi(v, w) \neq 0 \,.$$

(iii) Die assoziierte lineare Abbildung

$$\varphi_\psi : V \to V^*,\; w \mapsto (v \mapsto \psi(v, w))$$

ist ein Isomorphismus.

(iv) Für eine Basis b von V ist die Fundamentalmatrix $B = M_b(\psi)$ von ψ bezüglich b invertierbar.

(v) Für jede Basis b ist die Fundamentalmatrix von ψ bezüglich b invertierbar.

Beweis der Äquivalenz: Offenbar gilt

$$\text{(i)} \quad \Leftrightarrow \quad \varphi_\psi \text{ ist injektiv}$$

Wegen $\dim V^* = \dim V < \infty$ gilt außerdem

$$\varphi_\psi \text{ ist injektiv} \quad \Leftrightarrow \quad \varphi_\psi \text{ ist surjektiv}\,.$$

Dies zeigt die Äquivalenz von (i) und (iii).

(iii) $\Rightarrow$ (ii): Sei $v \in V \smallsetminus \{0\}$. Dann gibt es ein $\chi \in V^*$ und $\chi(v) \neq 0$: Ergänze v zu einer Basis $(v, f_2, \ldots, f_n)$; dann gibt es ein $\chi : V \to K$ mit $\chi(v) = 1$ und (z. B.) $\chi(f_i) = 0$ für $i = 2, \ldots, n$. Ist φ_ψ surjektiv, so gibt es ein $w \in V$ mit $\varphi_\psi(w) = \psi(-, w) = \chi$, also $\psi(v, w) = \chi(v) \neq 0$.

Damit gilt auch (i) $\Rightarrow$ (ii), und aus Symmetriegründen (Übergang von ψ zu ψ' mit $\psi'(v, w) = \psi(w, v)$) schließen wir auch (ii) $\Rightarrow$ (i):

(iii) $\Leftrightarrow$ (v) folgt aus Bemerkung 16.7, und (v) $\Rightarrow$ (iv) ist trivial. Aber (iv) $\Rightarrow$ (v) folgt aus der Transformationsregel 16.6: Ist B invertierbar, so auch T^tBT für jede invertierbare Matrix T (da T^t wieder invertierbar ist).

Definition 16.9 Sei V endlich-dimensional. Der Rang einer Bilinearform $\psi : V \times V \to K$ wird definiert als der Rang einer zugehörigen Fundamentalmatrix B.

Bemerkung 16.10 (a) Dies ist unabhängig von der Wahl einer Basis, denn für eine invertierbare $(n \times n)$-Matrix T, $n = \dim V$, ist $\mathrm{rg}(B) = \mathrm{rg}(T^tBT)$.

(b) ψ ist nicht-ausgeartet $\Leftrightarrow \quad \psi$ hat Rang $n = \dim V \quad \Leftrightarrow \quad \det(B) \neq 0$.

(c) Die "Gram'sche Determinante" $\det(B)$ (B Fundamentalmatrix von ψ) hängt von der Wahl einer Basis ab: $\det(T^tBT) = \det(T^t) \cdot \det(B) \cdot \det(T) = \det(B) \cdot [\det(T)]^2$, aber das Verschwinden von $\det(B)$ hängt nicht von der Basis ab.

Erinnerung 16.11 (vergl. LA I 13.5) Eine Bilinearform $\psi \in \mathrm{Bil}(V)$ heißt **symmetrisch**, wenn $\psi(v,w) = \psi(w,v)$ für alle $v, w \in V$.

Ist V endlich-dimensional, so ist ψ offenbar genau dann symmetrisch, wenn eine (und damit jede) Fundamentalmatrix von ψ symmetrisch ist. Insbesondere sind die symmetrischen Bilinearform auf K^n von der Form

$$\psi_B(x,y) = x^t By$$

mit $B = B^t \in M_n(K)$.

Symmetrische Bilinearformen hängen eng mit sogenannten quadratischen Formen zusammen:

Definitionen 16.12 Ist $\psi : V \times V \to K$ eine symmetrische Bilinearform, so heißt

$$q = q_\psi : V \to K \quad \text{mit} \quad q(v) = \psi(v,v)$$

die assoziierte **quadratische Form**.

Bemerkung 16.13 Damit ist auch definiert, was eine quadratische Form auf V ist: eine Abbildung $q : V \to K$ von der Form $q = q_\psi$ für eine symmetrische Bilinearforn ψ auf V. Insbesondere ist eine quadratische Form auf K^n eine Abbildung $q : K^n \to K$ mit

$$q(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j \,,$$

wobei $A = (a_{ij})$ eine symmetrische Matrix ist.

Lemma 16.14 Ist die Charakteristik von K ungleich 2 ($\mathrm{char}(K) \neq 2$, siehe LA I, Übungsaufgabe 8), so ist ψ durch die assoziierte quadratische Form q bestimmt; insbesondere gilt $\psi = 0 \Leftrightarrow q = 0$.

Beweis Die Voraussetzung bedeutet, dass $2 \neq 0$ in K , d.h., dass 2 invertierbar ist. Dann gilt aber

$$\psi(v,w) = \frac{1}{2}\left(q(v+w) - q(v) - q(w)\right).$$

Lemma/Definition 16.15 (vergl. LA I §18) Sei $\psi : V \times V \to K$ eine symmetrische Bilinearform.

(a) Zwei Vektoren $v, w \in V$ heißen **orthogonal** (oder **senkrecht**) zueinander (bezüglich ψ), Bezeichnung $v \perp w$, wenn

$$\psi(v,w) = 0\,.$$

(b) Für $v \in V$ heißt
$$v^\perp := \{w \in V \mid v \perp w\}$$
das **orthogonale Komplement von** v (bezüglich ψ). Dies ist ein Unterraum von V.

(c) Sei $U \subseteq V$ ein Unterraum. Dann heißt
$$U^\perp = \{w \in V \mid w \perp v \quad \text{für alle} \quad v \in U\}$$
das **orthogonale Komplement von** U (bezüglich ψ) und ist ein Unterraum von V.

(d) Ist V endlich-dimensional und ψ nicht-ausgeartet, so gilt
$$\dim U + \dim U^\perp = \dim V\,.$$

Beweis der Behauptungen:

(b): selbst!

(c): Dass $U^\perp$ ein Unterraum ist, folgt zum Beispiel daraus, dass $U^\perp$ der Kern der folgenden Komposition von linearen Abbildungen ist
$$\begin{array}{rccccc} \alpha: & V & \xrightarrow{\varphi_\psi} & V^* & \xrightarrow{\beta} & U^* \\ & v & \mapsto & \psi(-,v) & & \\ & & & \chi & \mapsto & \chi_{|\,U}\,. \end{array}$$

(d): Wir benutzen die obige lineare Abbildung α und zeigen, dass sie *surjektiv* ist. Dann folgt mit der Rangformel
$$\dim U^\perp = \dim \ker(\alpha) = \dim V - \dim U^*\,,$$
wegen $\dim U^* = \dim U$ also die Behauptung.

Da φ_ψ nach Voraussetzung ein Isomorphismus ist (16.8), genügt es, die Surjektivität von β zu zeigen. Sei $(b_1, \ldots, b_m)$ eine Basis von U, und ergänze dies zu einer Basis
$$(b_1, \ldots, b_m, c_1, \ldots, c_r)$$
von V. Ist dann $\chi : U \to K$ eine Linearform, so gibt es eine Linearform $\tilde{\chi} : V \to K$ mit
$$\begin{array}{rclcl} \tilde{\chi}(b_i) & = & \chi(b_i) & , & i = 1, \ldots, m, \\ \tilde{\chi}(c_j) & = & 0 & , & j = 1, \ldots, r \end{array}$$
(universelle Eigenschaft einer Basis). Für diese gilt offenbar $\tilde{\chi}_{|\,U} = \chi$. Dies zeigt die Surjektivität von β.

Lemma 16.16 Sei $\psi : V \times V \to K$ eine symmetrische Bilinearform. Für Unterräume $U, W \subseteq V$ gilt dann
$$U \subseteq W \Rightarrow W^\perp \subseteq U^\perp$$
und
$$U \subseteq U^{\perp\perp}\,.$$

Ist $\dim V < \infty$ und ψ nicht-ausgeartet, so gilt $U = U^{\perp\perp}$.

Beweis Die ersten beiden Aussagen sind unmittelbar klar. Ist $\dim V = n < \infty$ und ψ nicht-ausgeartet, so gilt weiter nach 16.15 (d)

$$\dim U^{\perp\perp} = n - \dim U^{\perp} = n - (n - \dim U) = \dim U\,.$$

Definition 16.17 Sei $\psi : V \times V \to K$ eine symmetrische Bilinearform. Eine Basis $(b_i)_{i\in I}$ von V heißt **Orthogonalbasis** bezüglich ψ, wenn

$$\psi(b_i, b_j) = 0 \quad \text{für} \quad i \neq j\,.$$

(Beachte den Unterschied zur *Orthonormal*basis, wie in LA I §18 definiert!)

Satz 16.18 Sei $\dim V < \infty$ und ψ eine symmetrische Bilinearform auf V. Ist $\operatorname{char}(K) \neq 2$, so besitzt V eine Orthogonalbasis bezüglich ψ.

Beweis durch Induktion über $n = \dim V$, wobei die Fälle $n = 0$ oder $n = 1$ trivial sind. Ist die assoziierte quadratische Form q identisch 0, so ist nach 16.14 auch $\psi = 0$ und wir können jede Basis von V nehmen. Sei also $q \neq 0$. Dann gibt es einen Vektor $v \in V$ mit

$$0 \neq q(v) = \psi(v, v)\,.$$

Sei $U =< v >_K$; dies ist ein 1-dimensionaler Unterraum von V. Dann ist $U \cap U^{\perp} = \{0\}$, wegen $\psi(v, v) \neq 0$. Weiter ist $U + U^{\perp} = V$: Sei nämlich $w \in V$ und $\alpha := \psi(v, w) \in K$. Dann ist

$$w' := w - \frac{\alpha}{q(v)} w \in U^{\perp}$$

und $w = \lambda v + w'$ mit $\lambda = \alpha/q(v) \in K$. Es folgt

$$V = U \oplus U^{\perp}\,;$$

insbesondere ist $\dim U^{\perp} = n - 1$. Nach Induktionsvoraussetzung gibt es eine Orthogonalbasis $(v_2, \ldots, v_n)$ von $U^{\perp}$, bezüglich der Einschränkung von ψ auf $U^{\perp}$ (d.h., auf $U^{\perp} \times U^{\perp}$). Dann ist $(v_1 = v, v_2, \ldots, v_n)$ eine Orthogonalbasis von V.

Corollar 16.19 Sei $B \in M_n(K)$ eine symmetrische Matrix, und sei $\operatorname{char}(K) \neq 2$. Dann gibt es eine invertierbare Matrix $T \in M_n(K)$, so dass $T^t B T$ eine Diagonalmatrix ist.

Beweis Dies folgt durch Anwendung von 16.18 auf die Bilinearform $\psi_B : K^n \times K^n \to K$, $\psi_B(x, y) = x^t B y$, unter Berücksichtigung der Transformationsregel 16.6 und der Tatsache, dass die Fundamentalmatrix bezüglich einer *Orthogonalbasis* immer einer *Diagonalmatrix* ist.

Definition 16.20 (a) Sei $GL_n(K)$ die Gruppe der **invertierbaren** $(n \times n)$-Matrizen über K. Sie wird auch die allgemeine lineare Gruppe n-ter Ordnung genannt.

(b) Zwei Matrizen $B, B' \in M_n(K)$ heißen **äquivalent**, wenn es ein $T \in GL_n(K)$ gibt mit

$$B' = T^t B T\,.$$

Dies ist offenbar eine Äquivalenzrelation, und im Allgemeinen verschieden von der Relation der Ähnlichkeit.

Corollar 16.21 Für $\mathrm{char}(K) \neq 2$ ist jede symmetrische Matrix $B \in M_n(K)$ äquivalent zu einer Diagonalmatrix.

Bemerkung 16.22 Für eine Bilinearform ψ sind nach 16.6 alle Fundamentalmatrizen äquivalent.

Definition 16.23 (a) Ein Paar (V, ψ) bestehend aus einem K-Vektorraum V und einer symmetrischen Bilinearform ψ auf V heißt **quadratischer Raum** (über K).

(b) Seien (V_1, ψ_1) und (V_2, ψ_2) quadratische Räume. Eine lineare Abbildung

$$\varphi : V_1 \to V_2$$

heißt **Isometrie** (oder **orthogonale Abbildung**), wenn φ bijektiv ist und

$$\psi_2(\varphi(v), \varphi(w)) = \psi_1(v, w) \quad \forall\, v, w \in V_1 \,.$$

(c) Zwei quadratische Räume (V_1, ψ_1) und (V_2, ψ_2) heißen **isomorph** (und die Bilinearformen ψ_1 und ψ_2 **äquivalent**, bzw. die quadratische Formen **affin äquivalent**), wenn es eine Isometire zwischen (V_1, ψ_1) und (V_2, ψ_2) gibt.

(d) Sei (V, ψ) ein quadratischer Raum. Die Gruppe (!) $O(V, \psi)$ (oder kurz $O(\psi)$) der Isometrien $\varphi : (V, \psi) \to (V, \psi)$ heißt die **orthogonale Gruppe** von (V, ψ) (oder von ψ).

Bemerkungen 16.24 (a) Manche Bücher nennen eine lineare Abbildung $\varphi : V_1 \to V_2$ schon Isometrie, wenn

$$(*) \qquad \psi_2(\varphi(v), \varphi(w)) = \psi_1(v, w) \quad \forall\, v, w \in V_1$$

gilt (keine Bijektivität von φ vorausgesetzt). Ist ψ_1 nicht-ausgeartet, so folgt aus $(*)$ jedenfalls schon die Injektivität von φ:

$$\varphi(v) = 0 \underset{(*)}{\Rightarrow} \psi_1(v, w) = 0 \quad \forall w \in V_1 \underset{\psi_1 \text{ n.a}}{\Rightarrow} v = 0 \,.$$

(b) Zwei symmetrische Bilinearformen $\psi_1 : V_1 \times V_1 \to K$ und $\psi_2 : V_2 \times V_2 \to K$ sind genau dann äquivalent, wenn $\dim V_1 = \dim V_2$, und wenn die Fundamentalmatrizen von ψ_1 und ψ_2 äquivalent sind (es reicht, dies für zwei beliebige Fundamentalmatrizen von ψ_1 und ψ_2 zu prüfen).

(c) Ist K algebraisch abgeschlossen (z.B. $K = \mathbb{C}$), so sind zwei quadratische Räume (V_1, ψ_1) und (V_2, ψ_2) genau dann isomorph, wenn $\dim V_1 = \dim V_2$ und $\mathrm{rg}\ \psi_1 = \mathrm{rg}\ \psi_2$ (Beweis: selbst!).

§17 Bilinearformen über $\mathbb{R}$

Definition 17.1 Eine symmetrische Bilinearform $\psi : V \times V \to \mathbb{R}$ auf einem $\mathbb{R}$-Vektorraum V heißt

(a) positiv definit, wenn $\psi(v,v) > 0$ für alle $v \in V \smallsetminus \{0\}$,

(b) negativ definit, wenn $\psi(v,v) < 0$ für alle $v \in V \smallsetminus \{0\}$,

(c) indefinit sonst.

Entsprechend heißt eine symmetrische Matrix $A \in M_n(\mathbb{R})$ positiv definit oder negativ definit oder indefinit, wenn dies für die zugehörige Bilinearform $(x,y) \mapsto x^t A y$ auf $\mathbb{R}^n$ gilt.

Dies sind alles Eigenschaften der zugehörigen quadratischen Formen, also $v \mapsto q_\psi(v) = \psi(v,v)$ bzw. $x \mapsto Q_A(x) = x^t A x$, und entsprechend heißen diese Formen auch positiv, negativ, oder indefinit

Beispiele 17.2 (a) Die Standard-Bilinearform

$$(x,y) \mapsto x^t y$$

auf $\mathbb{R}^n$ ist symmetrisch und positiv definit, denn für $x \neq 0$ ist $x^t x = \sum\limits_{i=1}^{n} x_i^2 > 0$.

(b) Auf dem Raum $C[a,b]$ der stetigen Funktionen auf dem Intervall $[a,b]$ ist die Bilinearform

$$(f,g) \mapsto \int\limits_a^b f(t)g(t)dt$$

symmetrisch und positiv definit, denn für stetiges $f(t) \not\equiv 0$ ist $\int\limits_a^b f(t)^2 dt > 0$.

(c) In der Analysis zeigt man: Sei $U \subseteq \mathbb{R}^n$ offen und

$$f : U \to \mathbb{R}$$

zweimal stetig differenzierbar, und sei $x_0 \in U$. Ist $(\mathrm{grad}\ f)(x_0) = 0$ und ist die (symmetrische!) Hesse-Matrix

$$\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0)\right)_{i,j=1,\dots,n}$$

positiv (bzw. negativ) definit, so hat f bei x_0 ein lokales Minimum (bzw. Maximum). Ist die Hesse-Matrix indefinit, so kann ein lokales Minimum oder ein lokales Maximum oder ein Sattelpunkt (kein lokales Extremum) vorliegen.

Satz 17.3 (Trägheitssatz von Sylvester)

(a) Sei V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum, und sei $\psi : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform. Dann gibt es eine Basis b von V so, dass die Fundamentalmatrix von

ψ bezüglich b die Form

$$
(17.3.1) \qquad \begin{array}{c} r\left\{\begin{array}{c} \\ \\ \\ \end{array}\right. \\ s\left\{\begin{array}{c} \\ \\ \\ \end{array}\right. \\ \\ \\ \\ \end{array}
\begin{pmatrix}
1 & & & & & & & & \\
& \ddots & & & & & & & \\
& & 1 & & & & 0 & & \\
& & & -1 & & & & & \\
& & & & \ddots & & & & \\
& & & & & -1 & & & \\
& 0 & & & & & 0 & & \\
& & & & & & & \ddots & \\
& & & & & & & & 0
\end{pmatrix}
$$

hat (Eine solche Basis nennen wir eine Sylvester-Basis für ψ).

(b) (Matrixversion) Ist $B \in M_n(\mathbb{R})$ symmetrisch, so gibt es ein $T \in GL_n(\mathbb{R})$, so dass

$$T^t BT$$

die Gestalt (17.3.1) hat.

(c) Die Anzahl r der Diagonalelemente gleich 1 und die Anzahl s der Diagonalelemente gleich -1 (und damit auch die Anzahl der Nullen auf der Diagonalen) in (17.3.1) ist eindeutig durch φ (bzw. durch B) bestimmt.

Die Zahl $r+s$ ist der Rang von φ (bzw. B, bzw. der assoz. quadratischen Form q und die Zahl $r-s$ heißt die **Signatur** von φ (bzw. B, bzw. q).

Beweis Offenbar sind (a) und (b) äquivalent (16.6).

(b): Nach 16.19 gibt es ein $T_1 \in Gl_n(\mathbb{R})$ mit

$$
B' = T_1^t B T_1 = \begin{pmatrix}
\alpha_1 & & & & \\
& \ddots & & 0 & \\
& & \ddots & & \\
& 0 & & \ddots & \\
& & & & \alpha_n
\end{pmatrix}
$$

mit $\alpha_1, \ldots, \alpha_r \in \mathbb{R}$. Sei

$$
\beta_i = \begin{cases} \sqrt{|a_i|} \quad \text{(positive Wurzel)} & , \quad \alpha_i \neq 0 \\ 1 & , \quad a_i = 0\,. \end{cases}
$$

Dann ist die Matrix

$$
T_2 = \begin{pmatrix}
\beta_1^{-1} & & & & \\
& \ddots & & 0 & \\
& & \ddots & & \\
& 0 & & \ddots & \\
& & & & \beta_n^{-1}
\end{pmatrix} \in Gl_n(\mathbb{R})\,,
$$

und es ist

$$B'' = T_2^t B' T_2 = (T_1T_2)^t B(T_1T_2) = \begin{pmatrix} \gamma_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \ddots \\ & & & & \gamma_n \end{pmatrix}$$

mit

$$\gamma_i = \begin{cases} 1 & \alpha_i > 0 \\ -1 & \alpha_i < 0 \\ 0 & \alpha_i = 0\,. \end{cases}$$

Durch Umordnung der Basis erhält man also die gewünschte Form (17.3.1). Diese Umordnung bedeutet den Übergang zu $T_3^{-1}B''T_3 = T_3^t B'' T_3 = (T_1T_2T_3)^t B(T_1T_2T_3)$ mit einer Permutationsmatrix T_3; eine solche ist bekanntlich orthogonal.

(c) Es ist rg ψ = rg B = Rang der Matrix (17.3.1) $= r + s$.

Weiter behaupten wir

Behauptung: r ist die maximale Dimension eines Untervektorraums $U \subseteq V$, auf dem ψ positiv-definit ist.

Beweis: Sei $b_1, \ldots, b_n$ eine Sylvester-Basis bezüglich ψ, und sei $V_1 =< b_1, \ldots, b_r >_{\mathbb{R}}$ (r wie in (17.3.1)) und $V_2 =< b_{r+1}, \ldots, b_n >_{\mathbb{R}}$. Dann ist ψ positiv definit auf V_1. Angenommen, es gibt ein $U \subseteq V$ mit $\dim U > r$ und ψ positiv definit auf U. Mit der Dimensionsformel

$$\dim U + \dim V_2 = \dim(U + V_2) + \dim U \cap V_2$$

folgt

$$\dim U \cap V_2 = \dim U + \dim V_2 - \dim(U + V_2) > r + (n - r) - n = 0\,,$$

da $\dim U > r$, $\dim V_2 = n - r$ und $\dim(U - V_2) \leq \dim V = n$. Also ist $U \cap V_2 \neq 0$. Aber auf V_2 ist ψ negativ semi-definit, d.h., $\psi(v, v) \leq 0$ für alle $v \in V_2$. Widerspruch zur Positiv-Definitheit von ψ auf U!

Daher sind r und s durch φ bestimmt.

(Es gilt auch noch: s = maximale Dimension eines Unterraums $U' \subseteq V$ auf dem ψ negativ definit ist).

§18 Quadriken

Definition 18.1 Sei K ein Körper der Charakteristik $\neq 2$, und sei $n \in \mathbb{N}$.

(a) Eine **quadratische Funktion** $f : K^n \to K$ ist eine Abbildung der Form

$$f(x) = \sum_{i,j=1}^{n} g_{ij} x_i x_j + \sum_{i=1}^{n} a_i x_i + b = x^t G x + a^t x + b\,,$$

mit $G = (g_{ij}) \in M_n(K)$, $a = (a_i) \in K^n$ und $b \in K$, wobei ohne Einschränkung G symmetrisch sei.

(b) Die Menge

$$M = \{x \in K^n \mid f(x) = 0\}$$

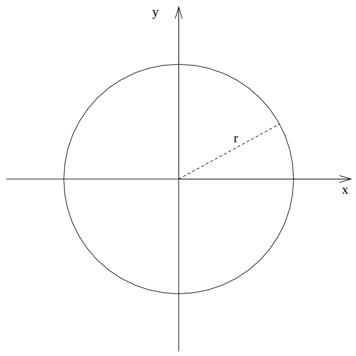
heißt die durch f bestimmte **affine Quadrik** oder **Hyperfläche zweiter Ordnung** (auch zweiten Grades).

Beispiele 18.2 (a) Ist $a = 0$ und $b = 0$, so haben wir eine quadratische Form (vergl. 12.8).

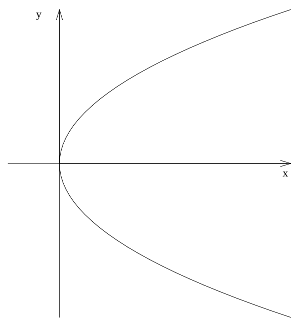
(b) Ist $G = 0$, so haben wir eine affine Hyperfläche in K^n.

(c) Sei $K = \mathbb{R}$

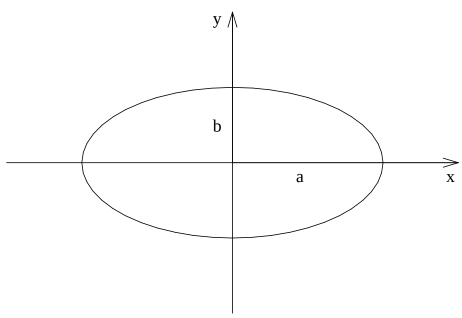
1) $x^2 + y^2 = r^2$ beschreibt den Kreis um 0 mit Radius r



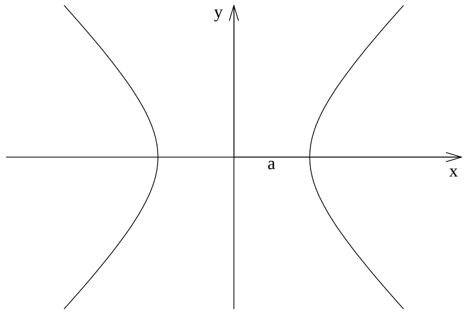
2) $y^2 = px$ ist eine Parabel



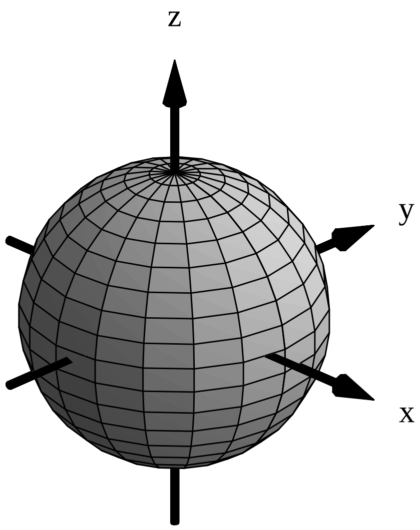
3) $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ ist eine Ellipse



4) $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ ist eine Hyperbel



5) $x^2 + y^2 + z^2 = r^2$ ist die Sphäre mit Radius r im $\mathbb{R}^3$.



Sei nun $K = \mathbb{R}$. Wie sieht eine Quadrik bis auf eine Bewegung in $\mathbb{R}^n$ aus? Hierbei definieren wir

Definition 18.3 Eine **Bewegung** im $\mathbb{R}^n$ ist eine Abbildung

$$f : \mathbb{R}^n \to \mathbb{R}^n$$

der Form $x \mapsto Ax + v$ mit $A \in O(n)$ und $v \in \mathbb{R}^n$.

Dies sind gerade die Abbildungen, die Längen und Winkel erhalten und entsprechen in $\mathbb{R}^2$ und $\mathbb{R}^3$ den Abbildungen, bei denen ein Körper "seine Form behält". Wir wollen sie benutzen, um $f(x)$ auf einfachere Gestalt zu bringen.

Sei $f(x) = x^t G x + ax + b$ eine quadratische Form, mit symmetrischem G. Wir führen erst eine orthogonale Transformation $x \mapsto T_1 x$ durch, $T_1 \in O(n)$, die G auf Diagonalgestalt bringt, d.h., es ist dann

$$fT_1(x) = \sum_{i=1}^{n} \lambda_i x_i^2 + 2\bar{a}^t x + b\,,$$

mit $T_1^t G T_1 = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ und $\bar{a}^t = \frac{1}{2} a^t T_1$. Jetzt wenden wir eine Transformation $T_2 x = x + v$ an und erhalten

$$fT_1T_2(x) = \sum_{i=1}^{n} \lambda_i x_i^2 + 2 \sum_{i=1}^{n} \tilde{a}_i x_i + \tilde{b}$$

mit $\tilde{b} = b + \sum_{i=1}^{n} \lambda_i v_i^2 + 2 \sum_{i=1}^{n} \tilde{a}_i v_i$ und $\tilde{a}_i = \bar{a}_i + \lambda_i v_i$.

Durch Umordnen der Basis knnen wir erreichen: Es gibt ein k, $1 \le k \le n$ mit $\lambda_i \ne 0$ für alle $i = 1, \ldots, k$ und $\lambda_i = 0$ für $i > k$. Dann können wir $v_1, \ldots, v_k$ so bestimmen, dass

$$\tilde{a}_i = \bar{a}_i + \lambda_i v_i = 0 \quad \text{für} \quad i = 1, \ldots, k\,.$$

1. Fall $\tilde{a}_{k+1} = \ldots = \tilde{a}_n = 0$

Dann können wir f noch mit einer Konstanten $\ne 0$ multiplizieren (wodurch sich die Lösungsmenge nicht ändert) und erhalten eine Gleichung der Form

$$\sum_{i=1}^{k} \lambda_i x_i^2 = 0 \quad \text{oder} \quad \sum_{i=1}^{k} \tilde{\lambda}_i x_i^2 = 1\,.$$

2. Fall Andernfalls wählen wir im Raum $< e_{k+1}, \ldots, e_n >$ der letzten $n-k$ Komponenten eine Orthonormalbasis, deren erster Vektor gleich $w/\,||w||$ ist, wobei

$$w = \begin{pmatrix} \tilde{a}_{k+1} \\ \vdots \\ \tilde{a}_n \end{pmatrix}.$$

Übergang zu dieser Basis – was wieder ein orthogonale Transformation T_3 ist – liefert eine Gleichung

$$fT_1T_2T_3(x) = \sum_{i=1}^{k} \lambda_i x_i^2 + 2c\, x_{k+1} + d$$

mit $c, d \in \mathbb{R}$. Ersetzen wir x_{k+1} durch $x_{k+1} - \frac{d}{2c}$ (Transformation T_4 mit $T_4 x = x - \frac{d}{2c} e_{k+1}$) und multiplizieren wir die Gleichung mit $-\frac{1}{c}$, so erhalten wir die Form

$$\sum_{i=1}^{k} \tilde{\lambda}_i x_i^2 = 2x_{k+1}\,.$$

Zusammengefasst erhalten wir

Satz 18.4 Durch eine Bewegung des $\mathbb{R}^n$ läßt sich die quadratische Funktion f bis auf Multiplikation mit einem Faktor $\ne 0$ in eine der Formen

(i) $\sum_{i=1}^{k} \lambda_i x_i^2 = 0$

(ii) $\sum_{i=1}^{k} \lambda_i x_i^2 = 1$

(iii) $\sum_{i=1}^{k} \lambda_i x_i^2 = 2x_{k+1}$

transformieren. Insbesondere wird die Quadrik $M = \{x \in \mathbb{R}^n \mid f(x) = 0\}$ durch eine Bewegung in eine Quadrik vom Typ (i), (ii) oder (iii) transformiert.

Definition 18.5 Die obigen Formen heißen die **euklidischen Normalformen** der Quadriken, und der angegebene Prozess heißt **Hauptachsentransformation** für die Quadrik.

Wir machen uns nun ein Bild von den Quadriken mit den Normalformen.

18.6 Die Quadriken hängen nicht von den Koordinaten ab, die in den Gleichungen 18.4 (i)-(iii) nicht mehr vorkommen; diese letzten Koordinaten sind also beliebig. Es ist also z.B. im Fall (ii)

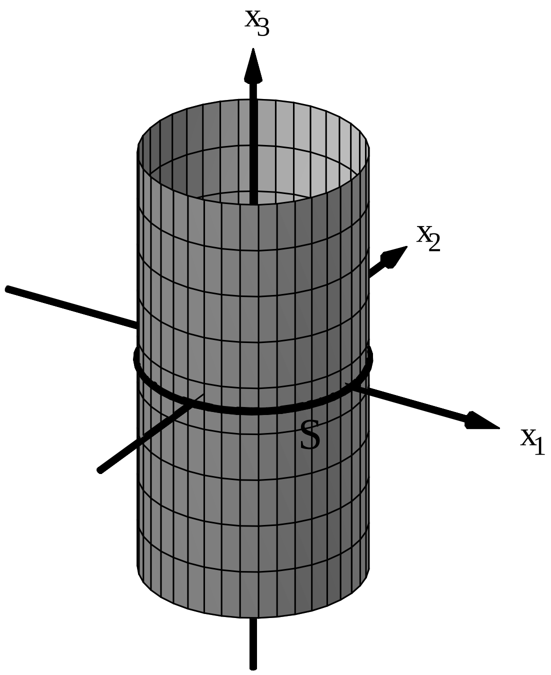
$$M = \{x \in \mathbb{R}^n \mid \sum_{i=1}^{k} \lambda_i x_i^2 = 1\} = \{x \in \mathbb{R}^k \mid \sum_{i=1}^{k} \lambda_i x_i^2 = 1\} \times \mathbb{R}^{n-k},$$

d.h., es genügt also, den Fall $k = n$ zu betrachten; der Fall $n < k$ ergibt sich durch das Produkt mit $\mathbb{R}^{n-k}$. Dasselbe gilt auch im Fall (ii), während wir im Fall (iii) nur die Situation $n = k + 1$ betrachten müssen und dann das Produkt mit $\mathbb{R}^{n-k-1}$ bilden.

Beispiel: $n = 3, k = 2$; dann ist z.B.

$$M = \{x \in \mathbb{R}^3 \mid x_1^2 + x_2^2 = 1\} = \{x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\} \times \mathbb{R},$$

wobei $\{x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\}$ der Kreis in $\mathbb{R}^2$ um 0 mit Radius 1 ist. Also ist M ein Zylinder:



Entsprechendes gilt in den Fällen (i) und (iii).

18.7 Weiter gibt es die degenerierten Fälle:

(a) (i) und alle $\lambda_i > 0$ oder alle $\lambda_i < 0 \quad \Rightarrow \quad M = \{0\}$ ist ein Punkt. Mit 18.6 und Bewegungen erhält man alle affinen Räume.

(b) (ii) und alle $\lambda_i < 0 \quad \Rightarrow \quad M = \emptyset$ ist leer.

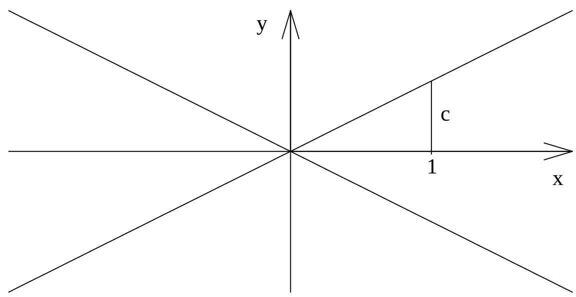
(c) (ii) und $n = 1$: $\quad x^2 = \frac{1}{\lambda} > 0 \Rightarrow M = \pm\frac{1}{\sqrt{\lambda}}$ zwei Punkte

18.8 Wir betrachten nun den Fall $\mathbb{R}^2$, mit maximalen k

(i): $c^2x^2 - y^2 = 0 \quad \text{mit}\, c \in \mathbb{R}$: Dann ist

$$M = \{(x, y) \in \mathbb{R}^2 \mid y = cx \cup y = -cx\},$$

die Vereinigung der beiden Geraden durch 0 mit Steigung $\pm c$



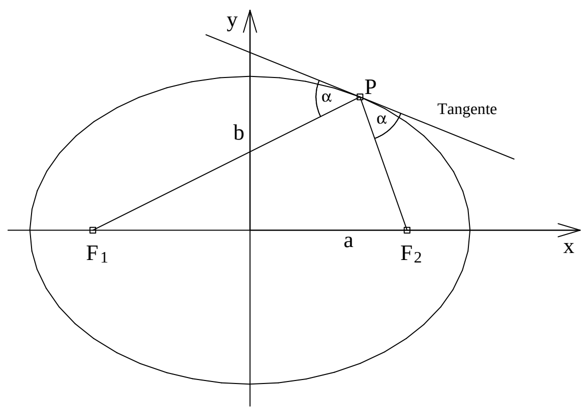
(ii)1): Seien $\lambda_1, \lambda_2 > 0$: Setze

$$a = \frac{1}{\sqrt{\lambda_1}}, \quad b = \frac{1}{\sqrt{\lambda_2}} \quad \in \mathbb{R}_{>0},$$

dann erhalten wir die Gleichung

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

M ist eine Ellipse mit den **Hauptachsenabschnitten** a und b.



Gilt $a \geq b$, wie hier, so ist $2a$ die Länge der **großen Achse**.

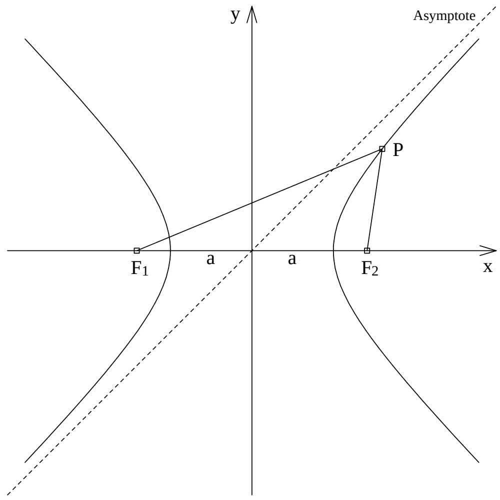
$2e = 2\sqrt{a^2 - b^2}$ heißt die **Exzentrizität**, und die Punkte $F_1 = (-e, 0)$ und $F_2(e, 0)$ sind die **Brennpunkte** der Ellipse: Für $P = (x, y) \in M$ ist $d(F_1, P) + d(F_2, P)$ konstant gleich $2a$.

Im Spezialfall $a = b = r$ ergibt sich der Kreis mit Radius r.

(ii) 2): Seien $\lambda_1 > 0, \lambda_2 < 0$: Setze

$$a = \frac{1}{\sqrt{\lambda_1}} \quad , \quad b = \frac{1}{\sqrt{(\lambda_2)}} \quad \in \mathbb{R}_{>0}$$

Damit erhalten wir die Normalform einer **Hyperbelgleichung**.



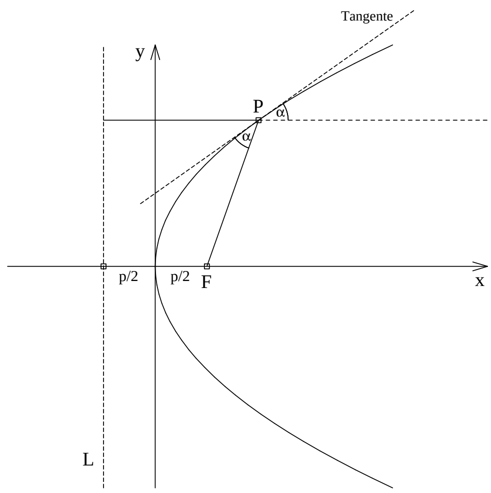
Die Exzentrizität ist $2e = 2\sqrt{a^2 + b^2}$, die **Brennpunkte** sind wieder $F_1 = (-e, 0)$ und $F_2 = (e, 0)$, und M ist die Menge der Punkte $P \in \mathbb{R}^3$, für die

$$|d(P, F_1) - d(P, F_2)| = 2a\,.$$

(iii) Hier erhalten wir die **Parabel**. Indem wir geeignet nummerieren und teilen, erhalten wir die Normalform

$$y^2 = 2px$$

Für $p > 0$ ist das Bild



Der **Brennpunkt** ist $F = (\frac{p}{2}, 0)$, die **Leitlinie** L ist die Geraden $x = -\frac{p}{2}$. Die Parabel ist die Menge der Punkte P, die den gleichen Abstand von L und von F haben.

18.9 Im Höherdimensionalen ist es am besten, eine Streckung

$$x_i \mapsto \frac{x_i}{\sqrt{|\lambda_i|}} \quad (i = 1, \dots, k)$$

der Koordinaten vorzunehmen. Dann erhält man als einzige Möglichkeit die Gleichungen

(i) $\| x \|^2 - \| y \|^2 = 0$,

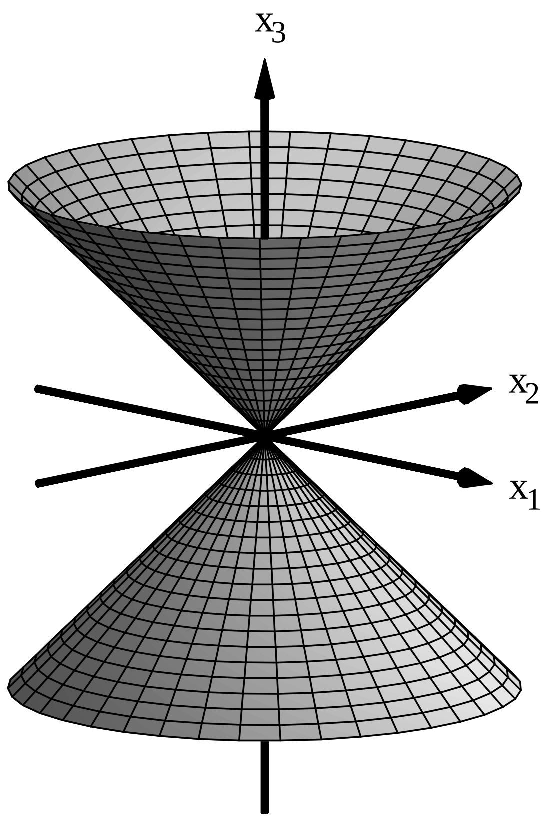
(ii) $\| x \|^2 - \| y \|^2 = 1$,

(iii) $\| x \|^2 - \| y \|^2 = 2z$,

wobei $x = (x_1, \dots, x_r)$ die Koordinaten mit positiven λ_i und $y = (y_1, \dots, y_s)$ die Koordinaten mit negativem λ_i zusammenfasst (Umnummerierung = orthogonale Abbildung). Es gelten also in den Fällen (i) und (ii) analoge Gleichungen wie in $\mathbb{R}^2$, aber für die Normen, und es entstehen oft Rotationskörper.

Wir illustrieren dies in $\mathbb{R}^3$

(i) $x_1^2 + x_2^2 = x_3^2$: (Doppel-)Kegel um die x_3-Achse

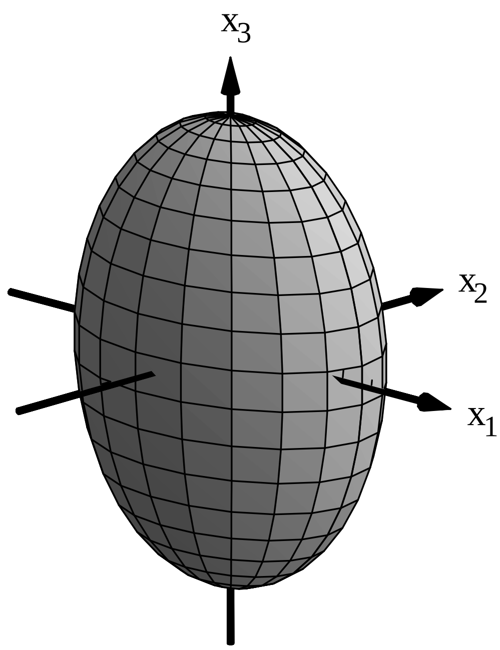


(Rotation des Geradenpaares $x_1^2 - x_3^2 = 0$ um die x_3-Achse)

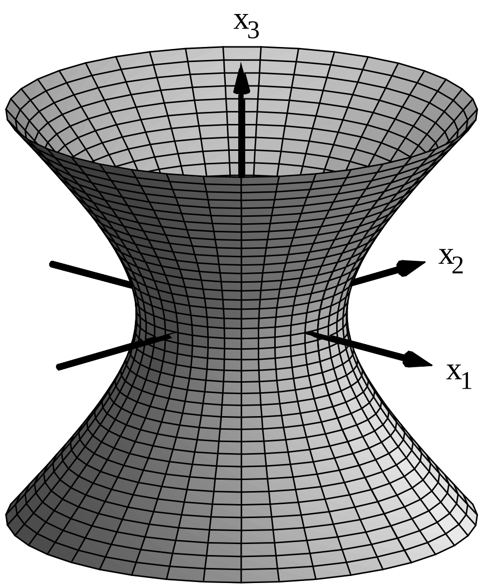
(ii) 1) $x_1^2 + x_2^2 + x_3^2 = 1$: Kugel mit Radius 1. Ohne Streckungen erhalten wir

$$\frac{x_1^2}{a_2} + \frac{x_2^2}{b^2} + \frac{x_3^2}{c^2} = 1\,,$$

ein Ellipsoid:

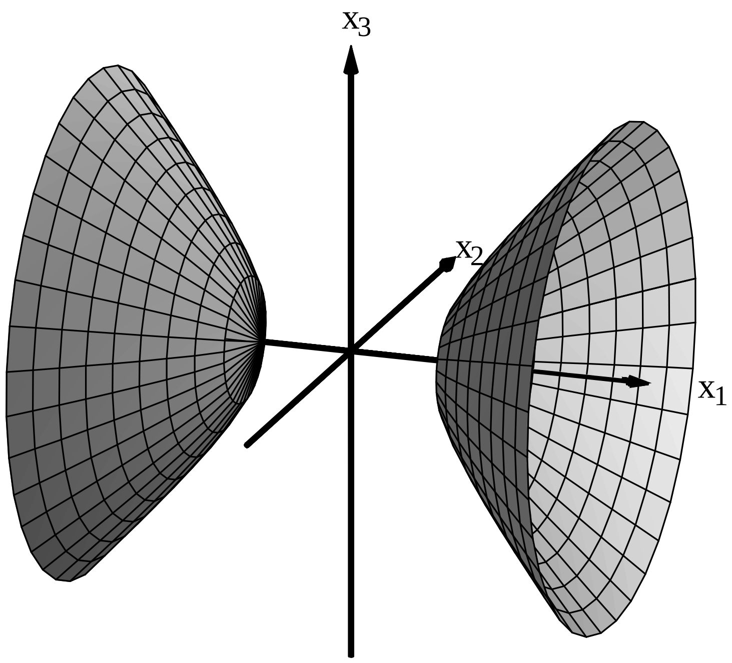


(ii) 2) $x_1^2 + x_2^2 - x_3^2 = 1$ bzw. $\| x \|^2 - y^2 = 1$



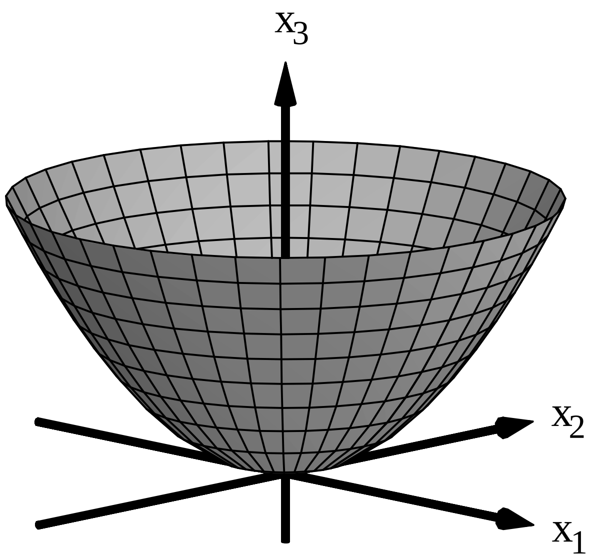
Einschaliges Hyperboloid (Rotation der Hyperbel $x_1^2 - x_3^2 = 1$ um die x_3-Achse)

(ii) 3) $x_1^2 - x_2^2 - x_3^2 = 1$ bzw. $x^2 - \parallel y \parallel^2 = 1$:



Zweischaliges Hyperboloid (Rotation der Hyperbel $x_1^2 - x_3^2 = 1$ um die x_1-Achse)

Für (iii) ergeben sich zwei Fälle:
(iii) 1) $x_1^2 + x_2^2 = 2x_3$ bzw. $\parallel x \parallel^2 = 2y$



Rotationsparaboloid (Rotation der Parabel $x_1^2 = 2x_3$ um die x_3-Achse), also wieder ein Rotationskörper.

(iii) 2) Als **neue** Figur im $\mathbb{R}^3$ ergibt sich

$$x_1^2 - x_2^2 = 2x_3,$$

die **Sattelfläche**