Lineare Algebra II, SoSe 2016

(Prof. Dr. O. Bogopolski)

Beweise werden in diesem Skript nur in Einzelfällen aufgeschrieben.

1 Euklidische und unitäre Räume

Im Folgenden sei K gleich $\mathbb{R}$ oder $\mathbb{C}$.

1.1 Euklidische Räume

Definition 1.1.1 Sei V ein $\mathbb{R}$-Vektorraum. Ein *Skalarprodukt* in V ist eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$, welche die folgenden Eigenschaften für alle $u, v, w, z \in V$ und $\lambda \in \mathbb{R}$ besitzt:

(1) $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$,
$\langle u, w + z \rangle = \langle u, w \rangle + \langle u, z \rangle$,
$\langle \lambda v, u \rangle = \lambda \langle v, u \rangle$,
$\langle v, \lambda u \rangle = \lambda \langle v, u \rangle$, (Linearität)

(2) $\langle v, u \rangle = \langle u, v \rangle$, (Symmetrie)

(3) $\langle v, v \rangle > 0$ für alle $v \neq 0$. (positive Definitheit)

(Aus (1) folgt $\langle \mathbf{0}, \mathbf{0} \rangle = 0$).

Ein $\mathbb{R}$-Vektorraum V zusammen mit einem Skalarprodukt heißt *Euklidischer Raum.*

Beispiel 1.1.2.

(a) Standardskalarprodukt in $\mathbb{R}^n$.

Für die Vektoren $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ aus $\mathbb{R}^n$ wird das Standardskalarprodukt so definiert:

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i.$$

(b) Für $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ aus $\mathbb{R}^2$ sei

$$\langle x, y \rangle := x_1 y_1 + 5 x_1 y_2 + 5 x_2 y_1 + 27 x_2 y_2.$$

Dann ist $\langle \cdot, \cdot \rangle$ ein Skalarprodukt in $\mathbb{R}^2$. Das ist aber kein Standardskalarprodukt.

(c) Sei $[a,b] \subset \mathbb{R}$ ein Intervall und sei $C([a,b])$ die Menge aller stetigen Funktionen $f : [a,b] \to \mathbb{R}$. Für $f, g \in C([a,b])$ setzen wir

$$\langle f, g \rangle = \int_a^b f(x)g(x)\,\mathrm{dx}.$$

Dann ist $C([a,b])$ mit $\langle \cdot, \cdot \rangle$ ein Euklidischer Raum. Der Raum ist ∞-dimensional.

1.2 Unitäre Räume

Definition 1.2.1 Sei V ein $\mathbb{C}$-Vektorraum. Ein *Skalarprodukt* in V ist eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$, welche die folgenden Eigenschaften für alle $u, v, w, z \in V$ und $\lambda \in \mathbb{C}$ besitzt:

(1) $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$,
$\langle u, w + z \rangle = \langle u, w \rangle + \langle u, z \rangle$,
$\langle \lambda v, u \rangle = \lambda \langle v, u \rangle$,
$\langle v, \lambda u \rangle = \overline{\lambda} \langle v, u \rangle$, (Linearität im ersten Argument)

(2) $\langle v, u \rangle = \overline{\langle u, v \rangle}$, (Das Skalarprodukt ist hermetisch.)

(3) $\langle v, v \rangle > 0$ für alle $v \neq 0$. (positive Definitheit)
(Aus (1) folgt $\langle \mathbf{0}, \mathbf{0} \rangle = 0$).

Ein $\mathbb{C}$-Vektorraum V zusammen mit einem Skalarprodukt heißt *unitärer Raum.*

Beispiel 1.2.2 (a) Standardskalarprodukt in $\mathbb{C}^n$.

Für die Vektoren $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ aus $\mathbb{C}^n$ wird das Standardskalarprodukt so definiert:

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i \overline{y_i}.$$

(b) Für $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ aus $\mathbb{C}^2$ sei

$$\langle x, y \rangle := x_1\overline{y_1} + 5x_1\overline{y_2} + 5x_2\overline{y_1} + 27x_2\overline{y_2}.$$

Dann ist $\langle \cdot, \cdot \rangle$ ein Skalarprodukt in $\mathbb{R}^2$. Das ist aber kein Standardskalarprodukt.

(c) Sei $[a,b] \subset \mathbb{R}$ ein Intervall und sei $C([a,b], \mathbb{C})$ die Menge aller stetigen Funktionen $f : [a,b] \to \mathbb{C}$. Für $f, g \in C([a,b], \mathbb{C})$ setzen wir

$$\langle f, g \rangle = \int_a^b f(x)\overline{g(x)}\,\mathrm{dx}.$$

Dann ist $C([a,b], \mathbb{C})$ mit $\langle \cdot, \cdot \rangle$ ein unitärer Raum. Der Raum ist ∞-dimensional.

1.3 Die Cauchy-Schwarzsche Ungleichung

In diesem Abschnitt benutzen wir, dass $z\overline{z} = |z|^2$ für $z \in \mathbb{C}$ gilt.

Satz 1.3.1 (Cauchy-Schwarzsche-Ungleichung). Sei V ein Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Dann gilt für alle $x, y \in V$

$$|\langle x, y \rangle|^2 \leqslant \langle x, x \rangle \langle y, y \rangle.$$

Gleichheit tritt genau dann ein, wenn $\{x, y\}$ linear abhängig ist.

Beweis. Wir betrachten nur den Fall eines unitären Raumes. Für $y = 0$ ist die Aussage richtig. Sei $y \neq 0$. Für beliebiges $t \in \mathbb{C}$ gilt:

$$0 \leqslant \langle x - ty, x - ty \rangle = \langle x, x \rangle - t\overline{\langle x, y \rangle} - \overline{t}\langle x, y \rangle + |t|^2 \langle y, y \rangle.$$

Für $t := \frac{\langle x,y \rangle}{\langle y,y \rangle}$ bekommen wir

$$\begin{aligned} 0 &\leqslant \langle x - ty, x - ty \rangle \\ &= \langle x, x \rangle - \frac{\left|\langle x,y \rangle\right|^2}{\langle y,y \rangle} - \frac{\left|\langle x,y \rangle\right|^2}{\langle y,y \rangle} + \frac{\left|\langle x,y \rangle\right|^2}{\langle y,y \rangle} \\ &= \langle x, x \rangle - \frac{\left|\langle x,y \rangle\right|^2}{\langle y,y \rangle}. \end{aligned}$$

□

Folgerung 1.3.2 .

(a) Die Cauchy-Schwarzsche Ungleichung für das Standardskalarprodukt im $\mathbb{R}^n$: Für $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ aus $\mathbb{R}^n$ gilt

$$\left(\sum_{i=1}^n x_i y_i \right)^2 \leqslant \sum_{i=1}^n x_i^2 \cdot \sum_{i=1}^n y_i^2.$$

(b) Die Cauchy-Schwarzsche Ungleichung für das Standardskalarprodukt im $\mathbb{C}^n$: Für $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ aus $\mathbb{C}^n$ gilt

$$\left| \sum_{i=1}^n x_i \overline{y_i} \right|^2 \leqslant \sum_{i=1}^n |x_i|^2 \cdot \sum_{i=1}^n |y_i|^2.$$

(c) Im Falle von $C([a, b], \mathbb{C})$, den stetigen Funktionen von $[a, b]$ nach $\mathbb{C}$, lautet die Cauchy-Schwarzsche Ungleichung:

$$\left| \int_a^b f(x)\overline{g(x)} \, \mathrm{dx} \right|^2 \leqslant \int_a^b |f(x)|^2 \, \mathrm{dx} \cdot \int_a^b |g(x)|^2 \, \mathrm{dx}.$$

1.4 Norm eines Vektors

Definition 1.4.1 Sei V ein K-Vektorraum. Eine Funktion $||\cdot|| : V \to \mathbb{R}$ heißt *Norm*, wenn für alle $x, y \in V$ und $\lambda \in K$ die folgenden drei Bedingungen erfüllt sind:

(1) $||x|| \geqslant 0$.
Es gilt $||x|| = 0$ genau dann, wenn $x = 0$ ist, (Positive Definitheit)

(2) $||\lambda x|| = |\lambda| \cdot ||x||$, (Homogenität)

(3) $||x + y|| \leqslant ||x|| + ||y||$. (Dreiecksungleichung)

Ein Vektorraum mit Norm heißt *normierter Raum*.

Satz 1.4.2 Sei V ein K-Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$. Dann wird durch $||x|| := \sqrt{\langle x, x\rangle}$ auf V eine Norm definiert und es gilt

$$|\langle x, y\rangle| \leqslant ||x|| \cdot ||y||.$$

Beweis. Wir überprüfen nur (3). Dabei benutzen wir, dass für alle $z \in \mathbb{C}$ gilt:

$$z + \overline{z} = 2\mathrm{Re}\,(z) \text{ und } \mathrm{Re}\,(z) \leqslant |z|.$$

Für alle $x, y \in V$ gilt

$$\begin{aligned} ||x + y||^2 &= \langle x + y, x + y\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\ &= ||x||^2 + 2\,\mathrm{Re}\langle x, y\rangle + ||y||^2 \leqslant ||x||^2 + 2\,|\langle x, y\rangle| + ||y||^2 \\ &\leqslant ||x||^2 + 2||x|| \cdot ||y|| + ||y||^2 = (||x|| + ||y||)^2. \end{aligned}$$

□

Definition 1.4.3 Sei $(V, ||\cdot||)$ ein normierter Raum. Wir sagen, dass *diese Norm von einem Skalarprodukt auf V induziert ist*, wenn ein Skalarprodukt $\langle\cdot,\cdot\rangle$ auf V mit der Eigenschaft $||x|| = \sqrt{\langle x, x\rangle}$ existiert.

Satz 1.4.4 Sei $(V, ||\cdot||)$ ein normierter Raum. Wenn die Norm $||\cdot||$ von einem Skalarprodukt auf V induziert ist, dann gilt für alle $x, y \in V$ die Parallelogrammidentität:

$$||x - y||^2 + ||x + y||^2 = 2(||x||^2 + ||y||^2).$$

Beispiel. Die Abbildung $||\cdot|| : \mathbb{R}^2 \to \mathbb{R}$, $||(x_1, x_2)|| := |x_1| + |x_2|$ ist eine Norm, die von keinem Skalarprodukt auf $\mathbb{R}^2$ induziert ist.

1.5 Winkel und Orthogonalität

Definition 1.5.1

(a) Sei V ein Euklidischer Raum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und seien $a, c \in V \setminus \{0\}$. Dann heißt $\varphi \in [0, \pi]$ der *Winkel* zwischen a und c, wenn gilt

$$\cos \varphi = \frac{\langle a, c \rangle}{||a||\,||c||}.$$

(b) Sei V ein Vektorraum mit einem Skalarprodukt. Dann heißen $a, c \in V$ zueinander *orthogonal*, wenn $\langle a, c \rangle = 0$. Man schreibt in diesem Fall auch $a \perp c$.

Bemerkung. Die Definition der Winkel in (a) ist sinvoll, denn nach der Cauchy-Schwarzschen Ungleichung gilt für alle $a, c \neq 0$

$$-||a|| \cdot ||c|| \leqslant \langle a, c \rangle \leqslant ||a|| \cdot ||c||$$

und somit

$$-1 \leqslant \frac{\langle a, c \rangle}{||a||\,||c||} \leqslant 1.$$

Da die Kosinusfunktion das Intervall $[0, \pi]$ auf $[-1, 1]$ bijektiv und stetig abbildet, gibt es einen eindeutig bestimmten Winkel φ zwischen a und c.

Der Begriff der Orthogonalität, das heißt der Winkel ist $\pi/2$ und $\cos \varphi = 0$, hat auch in unitären Räumen Sinn und auch für Vektoren, die gleich Null sind.

Beispiel.

(a) Wir betrachten den Euklidischen Raum $\mathbb{R}^2$ mit dem Standardskalarprodukt. Dann ist der Winkel zwischen den Vektoren $a = (2, 0)$ und $c = (1, -\sqrt{3})$ gleich $\pi/3$.

(b) Wir betrachten den unitären Raum $V = C([0, 2\pi], \mathbb{C})$ der komplexwertigen, stetigen Funktionen auf $[0, 2\pi]$, mit dem Skalarprodukt $\langle f, g \rangle = \int_0^{2\pi} f(x)\overline{g(x)}\,\mathrm{d}x$. Die Vektoren cos und sin sind orthogonal in V.

Definition 1.5.2 Sei I eine nichtleere Menge. Für $i, j \in I$ wird das *Kronekersymbol* so definiert:

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

Definition 1.5.3 Sei V ein Vektorraum mit Skalarprodukt und sei $C = (c_i)_{i \in I}$ ein nichtleeres System von Vektoren von V.

(a) C heißt *Orthogonalsystem*, abgekürzt OS, wenn für alle $i, j \in I$ mit $i \neq j$ gilt $\langle c_i, c_j \rangle = 0$.

(b) C heißt *Orthonormalsystem*, abgekürzt ONS, wenn für alle $i, j \in I$ gilt $\langle c_i, c_j \rangle = \delta_{i,j}$.

(c) C heißt *Orthonormalbasis*, abgekürzt ONB, wenn C ein Orthonormalsystem ist und gleichzeitig eine Basis in V.

Bemerkung. Sei C eine Orthogonalsystem im Skalarproduktraum V, die nicht den Nullvektor enthält. Dann kann man C in eine Orthonormalsystem C' überführen, indem man die einzelnen Vektoren *normiert.* Man setzt

$$c_i' = \frac{c_i}{||c_i||}, \ \ i \in I.$$

Beispiel.

(a) Im $\mathbb{R}^3$ mit Standardprodukt sind $\begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}-1\\0\\1\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}$ orthogonal. Das zugehörige Orthonormalsystem ist $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\0\\1\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}$.

(b) Im $C([0, 2\pi])$ mit dem Skalarprodukt aus Beispiel 1.1.2 (c) ist das System

$$\{1, \cos x, \sin x, \cos 2x, \sin 2x, \dots\}$$

orthogonal, aber nicht orthonormal.

Lemma 1.5.4 (Pythagoras) Sei $\{c_1, \dots, c_n\}$ ein Orthogonalsystem in V. Dann gilt

$$||c_1 + \dots + c_n||^2 = ||c_1||^2 + \dots + ||c_n||^2.$$

Lemma 1.5.5 Jede orthogonale Menge, die den Nullvektor nicht enthält, ist linear unabhängig.

Satz 1.5.6 (Gram - Schmidt Verfahren) Sei V ein endlichdimensionaler Vektorraum mit Skalarprodukt und sei $C = \{c_1, \dots, c_n\}$ ein linear unabhängiges System in V. Dann existiert ein Orthonormalsystem $E = \{e_1, \dots, e_n\}$ in V mit

$$\mathcal{L}(\{c_1, \dots, c_k\}) = \mathcal{L}(\{e_1, \dots, e_k\})$$

für alle $k = 1, \dots, n$.

Beweis.
Schritt 1. Zuerst definierern wir Vektoren $f_1, \dots, f_n$ rekursiv:

$$\begin{aligned} f_1 : &= c_1, \\ f_{i+1} : &= c_{i+1} - \sum_{j=1}^{i} \tfrac{\langle c_{i+1}, f_j \rangle}{\langle f_j, f_j \rangle} \cdot f_j \end{aligned}$$

Die Vektoren $f_1, \dots, f_n$ sind zueinander orthogonal und für alle $k = 1, \dots, n$ gilt

$$\mathcal{L}(\{c_1, \dots, c_k\}) = \mathcal{L}(\{f_1, \dots, f_k\}).$$

Schritt 2. Wir normieren die Vektoren $f_1, \dots, f_n$:

$$e_i := \frac{f_i}{||f_i||}, \quad i = 1, \dots, n.$$

□

Satz 1.5.7 Jeder endlichdimensionale Vektorraum mit Skalarprodukt besitzt eine Orthonormalbasis.

Definition 1.5.8 Sei V ein Vektorraum mit Skalarprodukt $\langle , \rangle$. Für jeden Untervektorraum U von V heißt die Menge

$$U^{\perp} = \{x \in V \mid \langle u, x \rangle = 0 \text{ für alle } u \in U\}$$

orthogonales Komplement zu U.

Lemma 1.5.9 Sei V ein endlichdimensionaler Vektorraum mit Skalarprodukt $\langle , \rangle$ und sei U ein Untervektorraum von V. Dann gilt:

(a) $U^{\perp}$ ist ein Untervektorraum von V,

(b) $\left(U^{\perp}\right)^{\perp} = U$,

(c) $U \oplus U^{\perp} = V$.

2 Orthogonale und unitäre Endomorphismen und Matrizen

2.1 Orthogonale und unitäre Endomorphismen

Definition 2.1.1 Sei V ein Vektorraum über K.

(a) Eine lineare Abbildung $\varphi : V \to V$ heißt *Endomorphismus* von V. Die Menge aller Endomorphismen von V wird mit $\mathbf{End}(V)$ bezeichnet.

(b) Die Menge aller bijektiven Endomorphismen (mit anderen Worten Isomorphismen) von V wird mit $\mathbf{GL}(V)$ bezeichnet.

Definition 2.1.2 (a) Ein Endomorphismus $\varphi : V \to V$ eines Euklidischen Raums V heißt *orthogonal*, wenn

$$\langle \varphi(x), \varphi(y) \rangle = \langle x, y \rangle$$

für alle $x, y \in V$ gilt. Die Menge aller orthogonalen Endomorphismen eines Euklidischen Raums V wird mit $\mathbf{O}(V)$ bezeichnet.

(b) Ein Endomorphismus $\varphi : V \to V$ eines unitären Raums heißt *unitär*, wenn

$$\langle \varphi(x), \varphi(y) \rangle = \langle x, y \rangle$$

für alle $x, y \in V$ gilt. Die Menge aller unitären Endomorphismen eines unitären Raums V wird mit $\mathbf{U}(V)$ bezeichnet.

Satz 2.1.3 Sei V ein Euklidischer (ein unitärer) Raum. Sei $\varphi : V \to V$ ein orthogonaler (ein unitärer) Endomorphismus. Dann gilt für alle $x, y \in V$:

(a) $||\varphi(x)|| = ||x||$.

(b) $x \perp y$ impliziert $\varphi(x) \perp \varphi(y)$.

(c) $\mathbf{O}(V)$ (bzw. $\mathbf{U}(V)$) ist eine Untergruppe von $\mathbf{GL}(V)$, wenn V endlichdimensional ist.

(d) Ist $\lambda \in \mathbb{C}$ ein Eigenwert von φ, dann gilt $|\lambda| = 1$.

(e) Die Eigenvektoren, die zu verschiedenen Eigenwerten von φ gehören, sind orthogonal: d.h. wenn $\lambda_1 \neq \lambda_2$ ist und $\varphi(x_1) = \lambda_1 x_1$ und $\varphi(x_2) = \lambda_2 x_2$ ist, dann gilt $\langle x_1, x_2 \rangle = 0$.

Satz 2.1.4 Sei V ein endlichdimensionaler unitärer Raum und $\varphi : V \to V$ ein unitärer Endomorphismus. Dann existiert eine Orthonormalbasis e' von V, so dass $[\varphi]_{e'}^{e'}$ eine Diagonalmatrix mit den Eigenwerten von φ auf der Diagonale ist.

Wir brauchen die folgende Definition, um den Satz 2.1.6 zu beweisen.

Definition 2.1.5 Sei V ein Vektorraum über $\mathbb{R}$ mit einer Basis $\mathcal{B} = (b_i)_{i \in I}$.

(a) Der Vektorraum über $\mathbb{C}$ mit der Basis $\mathcal{B}$ wird mit $V_{\mathbb{C}}$ bezeichnet. Die Elemente von $V_{\mathbb{C}}$ sind also endliche lineare Kombinationen von Elementen aus $\mathcal{B}$ mit Koeffizienten aus $\mathbb{C}$. Der Vektorraum $V_{\mathbb{C}}$ heißt *Komplexifizierung* von V.

(b) Sei $\varphi : V \to V$ eine lineare Abbildung. Wir definieren $\widetilde{\varphi} : V_{\mathbb{C}} \to V_{\mathbb{C}}$ durch

$$\widetilde{\varphi}\Big(\sum_{i \in I_0} \alpha_i b_i\Big) := \sum_{i \in I_0} \alpha_i \varphi(b_i).$$

wobei I_0 eine (beliebige) endliche Teilmenge von I ist und α_i Koeffizienten aus $\mathbb{C}$ sind. Die Abbildung $\widetilde{\varphi}$ heißt *komplexe Erweiterung* von φ.

Bemerkung.

1) V ist eine Teilmenge, aber kein Untervektorraum von $V_{\mathbb{C}}$.

2) $(\widetilde{\varphi})_{|V} = \varphi$.

3) Es gilt $[\widetilde{\varphi}]_{\mathcal{B}}^{\mathcal{B}} = [\varphi]_{\mathcal{B}}^{\mathcal{B}}$.

Satz 2.1.6 Sei V ein endlichdimensionaler Euklidischer Raum und $\varphi : V \to V$ ein orthogonaler Endomorphismus. Dann existiert eine Orthonormalbasis e' von V, so dass $[\varphi]_{e'}^{e'}$ eine Block-Diagonalmatrix ist, wobei diese Blöcke die Größe 1×1 oder 2×2 haben.

Die Blöcke der Größe 1×1 sind (1) oder (-1).
Die Blöcke der Größe 2×2 haben die Form

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

2.2 Orthogonale und unitäre Matrizen

Bezeichnung. Für $M \in \mathrm{Mat}(n, m, K)$ bezeichnen wir mit M^t die Matrix, die zu M transponiert ist. Für $M \in \mathrm{Mat}(n, m, \mathbb{C})$ bezeichnen wir mit $\overline{M}$ die Matrix, die zu M komplex konjugiert ist.

Definition 2.2.1

(a) Eine Matrix $M \in \mathrm{Mat}(n, n, \mathbb{R})$ heißt *orthogonal*, wenn $M \cdot M^t = E$ gilt. Die Menge aller orthogonalen Matrizen aus $\mathrm{Mat}(n, n, \mathbb{R})$ wird mit $\mathbf{O}_n$ bezeichnet.

(b) Eine Matrix $M \in \mathrm{Mat}(n, n, \mathbb{C})$ heißt *unitär*, wenn $M \cdot \overline{M^t} = E$ gilt. Die Menge aller unitären Matrizen aus $\mathrm{Mat}(n, n, \mathbb{C})$ wird mit $\mathbf{U}_n$ bezeichnet.

Bemerkung. Ist $A \in \mathbf{O}_n$, dann gilt $\det(A) = \pm 1$.
Ist $A \in \mathbf{U}_n$, dann gilt $|\det(A)| = 1$.

Satz 2.2.2

(1) $\mathbf{O}_n$ ist eine Untergruppe von $\mathrm{GL}_n(\mathbb{R})$.
$\mathbf{U}_n$ ist eine Untergruppe von $\mathrm{GL}_n(\mathbb{C})$.

(2) Sei $A \in \mathbf{O}_n$. Dann ist die Abbildung $\varphi_A : \mathbb{R}^n \to \mathbb{R}^n$, $X \mapsto AX$, orthogonal.
Sei $A \in \mathbf{U}_n$. Dann ist die Abbildung $\varphi_A : \mathbb{C}^n \to \mathbb{C}^n$, $X \mapsto AX$, unitär.

(3) Sei $\varphi : V \to V$ eine orthogonale Abbildung eines endlichdimensionalen Euklidischen Vektorraums V. Ist $\mathcal{B}$ eine Orthonormalbasis von V, dann ist die Darstellungsmatrix $[\varphi]_{\mathcal{B}}^{\mathcal{B}}$ orthogonal.

Sei $\varphi : V \to V$ eine unitäre Abbildung eines endlichdimensionalen unitären Vektorraums V. Ist $\mathcal{B}$ eine Orthonormalbasis von V, dann ist die Darstellungsmatrix $[\varphi]_{\mathcal{B}}^{\mathcal{B}}$ unitär.

(4) Sei $A \in \mathbf{O}_n$ oder $A \in \mathbf{U}_n$. Dann gilt:

  (a) Ist λ ein Eigenwert von A, dann gilt $|\lambda| = 1$.

  (b) Die Eigenvektoren von A, die zu verschiedenen Eigenwerten von A gehören, sind orthogonal: d.h. wenn $\lambda_1 \neq \lambda_2$ ist und $AX_1 = \lambda_1 X_1$ und $AX_2 = \lambda_2 X_2$ ist, dann gilt $\langle X_1, X_2 \rangle = 0$.

Satz 2.2.3 Seien $\mathcal{B}_1$ und $\mathcal{B}_2$ zwei Orthonormalbasen eines endlichdimensionalen Euklidischen (unitären) Vektorraums V. Dann ist die Übergangsmatrix T von $\mathcal{B}_1$ zu $\mathcal{B}_2$ orthogonal (unitär).

Satz 2.2.4

(a) Sei A eine unitäre Matrix. Dann existiert eine unitäre Matrix C, so dass

$$B := \overline{C^t} A C$$

eine Diagonalmatrix mit den Eigenwerten von A auf der Diagonale ist.
(Diese B heißt *Standardform* von A.)

(b) Sei A eine Orthogonalmatrix. Dann existiert eine Orthogonalmatrix C, so dass

$$B := C^t AC$$

eine Block-Diagonalmatrix ist, wobei diese Blöcke die Größe 1×1 oder 2×2 haben und die folgende spezifische Form:

Die Blöcke der Größe 1×1 sind (1) oder (-1).

Die Blöcke der Größe 2×2 haben die Form

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

(Diese B heißt *Standardform* von A.)

Ein Algorithmus, der für eine gegebene Orthogonalmatrix A *die Standardform* B *und die Übergangsmatrix* C *wie im Satz 2.2.4 (b) sucht.*

Gegeben sei eine Orthogonalmatrix A der Größe $n \times n$.

(1) Wir finden alle Nullstellen des charakteristischen Polynoms $\chi_A(\lambda)$.

(2) Für jede reelle Nullstelle α suchen wir eine Orthonormalbasis $\mathcal{O}(\alpha) = \{w_1, \dots, w_k\}$ des Eigenraums $\mathrm{Eig}(A, \alpha)$.

Wir bereiten $\underline{k}$ kleine 1×1-Blöcke $B(\alpha) = \boxed{\alpha}$ für die zukünftige Matrix B vor.

(3) Die nicht reellen Nullstellen von $\chi_A(\lambda)$ erscheinen in Paaren $\{\alpha, \overline{\alpha}\}$. Aus jedem Paar von komplex-konjugierten Nullstellen $\{\alpha, \overline{\alpha}\}$ wählen wir einen Repräsentant, sagen wir $\alpha = a + ib$.

Für jeden Repräsentant α berechnen wir eine Orthogonalbasis $\{u_1 + iv_1, \dots, u_k + iv_k\}$ des unitären Vektorraums $\mathrm{Eig}(A, \alpha)$. Die reellen Vektoren $u_1, v_1, \dots, u_k, v_k$ werden automatisch zueinander orthogonal und es wird automatisch gelten $||u_1|| = ||v_1||$, $\dots$, $||u_k|| = ||v_k||$. Wir normieren diese Vektoren und bilden die Menge $\mathcal{O}(\alpha) := \{\frac{u_1}{||u_1||}, \frac{v_1}{||v_1||}, \dots, \frac{u_k}{||u_k||}, \frac{v_k}{||v_k||}\}$.

Wir bereiten $\underline{k}$ 2×2-Blöcke $B(\alpha) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ für die zukünftige Matrix B vor.

(4) Sei $\mathcal{O}$ die Vereinigung aller $\mathcal{O}(\alpha)$, die wir in (2) und (3) berechnet haben. Die Menge $\mathcal{O}$ wird automatisch eine Orthonormalbasis von $\mathbb{R}^n$.

- C ist die Matrix, deren Spalten die Vektoren aus $\mathcal{O}$ sind.
- B ist die Block-Diagomalmatrix, die aus Blöcken $B(\alpha)$ besteht, die in Schritten (2)-(3) gebildet wurden.

3 Adjungierte und selbstadjungierte Abbildungen

Um kurze Beweise im Abschnitt 3.2 herzustellen, brauchen wir gute Bezeichnungen und Rechnungsregeln für Skalarprodukte, Darstellungsmatrizen und Koordinaten.

3.1 Bezeichnungen

(1) Für zwei K-Vektorräume V, W sei $\mathrm{Hom}(V, W)$ die Menge aller linearen Abbildungen von V nach W. Des weiteren sei V, W endlichdimensional.

(2) Sei $\varphi \in \mathrm{Hom}(V, W)$. Sei $e = (e_1, \ldots, e_n)$ eine Basis in V und sei $e' = (e'_1, \ldots, e'_m)$ eine Basis in W. Angenommen

$$\begin{cases} \varphi(e_1) & = a_{11}e'_1 + \cdots + a_{1m}e'_m \\ \vdots & \\ \varphi(e_n) & = a_{n1}e'_1 + \cdots + a_{nm}e'_m \end{cases},$$

dann heißt die Matrix

$$[\varphi]_e^{e'} := \begin{pmatrix} a_{11} & \ldots & a_{n1} \\ \vdots & & \vdots \\ a_{1m} & \ldots & a_{nm} \end{pmatrix},$$

die *Darstellungsmatrix von φ bezüglich der Basen e und e'*. Es gilt

$$(\varphi(e_1), \ldots, \varphi(e_n)) = (e'_1, \ldots, e'_m) \cdot \begin{pmatrix} a_{11} & \ldots & a_{n1} \\ \vdots & & \vdots \\ a_{1m} & \ldots & a_{nm} \end{pmatrix}.$$

Kurz:

$$\varphi(e) = e' \cdot [\varphi]_e^{e'}.$$

(3) Schreiben wir $v \in V$ in der Form $v = v_1e_1 + \cdots + v_ne_n$, dann heißt

$$[v]_e := \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

Koordinatenvektor von v bezüglich e. Es gilt

$$v = (e_1, \ldots, e_n) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

Kurz:

$$v = e \cdot [v]_e.$$

Lemma 3.1.1 Sei e eine Orthonormalbasis von V und seien $u, v \in V$ zwei Vektoren mit Koordinaten

$$[u]_e = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}, \quad [v]_e := \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

Dann gilt

$$\langle u, v \rangle_V = (u_1, \ldots, u_n) \cdot \begin{pmatrix} \overline{v_1} \\ \vdots \\ \overline{v_n} \end{pmatrix} = \big([u]_e\big)^t \cdot \overline{[v]_e}.$$

Lemma 3.1.2 Sei $\varphi \in \mathrm{Hom}(V, W)$. Seien e und e' Orthonormalbasen von V und W. Seien $v \in V$ und $w \in W$ zwei Vektoren mit Koordinaten

$$[v]_e = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \quad [w]_{e'} := \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix}.$$

Dann gilt

$$\langle \varphi(v), w \rangle_W = (v_1, \ldots, v_n) \cdot \big([\varphi]_e^{e'}\big)^t \cdot \begin{pmatrix} \overline{w_1} \\ \vdots \\ \overline{w_m} \end{pmatrix} = ([v]_e)^t \cdot \big([\varphi]_e^{e'}\big)^t \cdot \overline{[w]_{e'}}.$$

3.2 Adjungierte Abbildungen

Sei $K = \mathbb{R}$ oder $\mathbb{C}$.

Satz 3.2.1 Seien U, V, W drei endlichdimensionale K-Vektorräume mit Skalarprodukten.

(a) Zu jedem $\varphi \in \mathrm{Hom}(V, W)$ gibt es genau ein $\varphi^* \in \mathrm{Hom}(W, V)$ mit

$$\langle \varphi(v), w \rangle = \langle v, \varphi^*(w) \rangle \quad \text{für alle } v \in V, w \in W.$$

(Wir nennen $\varphi^* : W \to V$ die *Adjungierte* zu $\varphi : V \to W$.)

(b) Sei $\varphi \in \mathrm{Hom}(V, W)$. Für je zwei Orthonormalbasen e und e' von V und W gilt

$$[\varphi^*]_{e'}^e = \overline{\big([\varphi]_e^{e'}\big)^t}.$$

(c) Es gilt $\varphi^{**} = \varphi$.

(d) Für alle $\varphi, \psi \in \mathrm{Hom}(V, W)$, alle $\theta \in \mathrm{Hom}(U, V)$ und alle $c \in K$ gelten

$$\begin{aligned}
&(\varphi + \psi)^* = \varphi^* + \varphi^*, \\
&(c\varphi)^* = \bar{c}\varphi^*, \\
&(\varphi \circ \theta)^* = \theta^* \circ \varphi^*.
\end{aligned}$$

(e) Es gelten

$$\operatorname{Kern}(\varphi^*) = (\operatorname{Im}(\varphi))^\perp \quad \text{und} \quad (\operatorname{Im}(\varphi^*)) = (\operatorname{Kern}(\varphi))^\perp.$$

Insbesondere haben φ und φ^* denselben Rang (d.h. $\dim(\operatorname{Im}(\varphi)) = \dim(\operatorname{Im}(\varphi^*))$).

(f) Ist φ injektiv (surjektiv), dann ist φ^* surjektiv (injektiv).
Ist φ ein Isomorphismus, so ist auch φ^*, und es gilt $(\varphi^*)^{-1} = (\varphi^{-1})^*$.

3.3 Normierte Algebren

Definition 3.3.1 Ein K-Vektorraum $\mathcal{A}$ heißt *K-Algebra*, falls gilt:

(1) Eine Abbildung $\cdot : \mathcal{A} \times \mathcal{A} \to$ gegeben ist, so dass $(\mathcal{A}, +, \cdot)$ ein Ring ist.
Also gelten in $\mathcal{A}$ das Assoziativgesetz und die beiden Distributivgesetze.
(Die Abbildung $\cdot$ heißt Multiplikation.)

(2) Für alle $a_1, a_2 \in \mathcal{A}$ und $k \in K$ gilt

$$k(a_1 a_2) = (k a_1) a_2 = a_1 (k a_2).$$

Beispiel. Sei V ein K-Vektorraum. Dann ist $\operatorname{End}_K(V)$ eine K-Algebra.

Definition 3.3.2 Sei $\mathcal{A}$ eine Algebra über $\mathbb{R}$ oder $\mathbb{C}$. Eine Vektorraumnorm $||\cdot||$ auf $\mathcal{A}$ (siehe Definition 1.4.1) heißt *Algebrennorm*, falls außer den Bedingungen in 1.4.1 auch noch

$$||ab|| \leqslant ||a||\,||b||$$

für alle $a, b \in \mathcal{A}$ gilt. Trägt die Algebra $\mathcal{A}$ eine Algebranorm $||\cdot||$, so heißt $\mathcal{A}$ *normierte Algebra*[1].

Satz 3.3.3 Sei $(V, ||\cdot||)$ ein endlichdimensionaler normierter K-Vektorraum ($K = \mathbb{R}$ oder $\mathbb{C}$). Dann ist $(\mathbf{End}(V), ||\cdot||')$ eine normierte K-Algebra, wobei $||\cdot||'$ so definiert ist:

$$||\varphi||' := \sup_{0 \neq v \in V} \frac{||\varphi(v)||}{||v||} \quad \text{für alle } \varphi \in \mathbf{End}(V). \tag{3.3.1}$$

Weiter gilt

$$||\varphi||' = \sup_{||v|| \leqslant 1} ||\varphi(v)|| = \sup_{||v||=1} ||\varphi(v)|| = \max_{||v||=1} ||\varphi(v)|| \tag{3.3.2}$$

und

$$||\varphi(v)|| \leqslant ||\varphi||' \cdot ||v|| \quad \text{für alle } v \in V. \tag{3.3.3}$$

[1]Ist $\mathcal{A}$ außerdem vollständig bezüglich der Norm $||\cdot||$, so heißt $\mathcal{A}$ *Banachalgebra.*

3.4 Norm der adjungierten Abbildung

Satz 3.4.1 Sei V ein endlichdimensionaler K-Vektorraum mit einem Skalarprodukt $\langle\cdot,\cdot\rangle$ und einer induzierten Norm $||\cdot||$. Wir betrachten die K-Algebra $\mathbf{End}(V)$ mit der Norm $||\cdot||'$ wie im Satz 3.3.3. Dann gelten für jedes $\varphi \in \mathbf{End}(V)$:

$$||\varphi^*||' = ||\varphi||',$$
$$||\varphi||'^2 = ||\varphi\varphi^*||' = ||\varphi^*\varphi||'.$$

3.5 Hermitesche (oder selbstadjungierte) Abbildungen und Matrizen

Definition 3.5.1

(a) Sei V ein endlichdimensionaler K-Vektorraum ($K \in \{\mathbb{R},\mathbb{C}\}$) mit Skalarprodukt $\langle\cdot,\cdot\rangle$. Weiter sei $\varphi \in \mathbf{End}(V)$. Dann heißt φ *hermitesch*, wenn $\varphi = \varphi^*$ gilt. (Ist φ hermitesch, so ist wegen Satz 3.2.1 (f) auch φ^{-1} hermitesch.)

(b) Eine Matrix $A \in \mathrm{Mat}(n,n,\mathbb{C})$ heißt *hermitesch*, wenn $A = \overline{A}^t$ gilt.

Behauptung 3.5.1′ Sei V ein endlichdimensionaler K-Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$. Ist $\varphi \in \mathbf{End}(V)$ hermitesch, dann ist $\langle\varphi(v), v\rangle$ reell für alle $v \in V$.

Satz 3.5.2 Sei V ein endlichdimensionaler K-Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$. Weiter sei $\varphi \in \mathbf{End}(V)$ hermitesch. Dann gilt:

(a) $||\varphi||' = \max\limits_{||v||\leqslant 1} |\langle\varphi(v), v\rangle|$.

(b) Ist $||\varphi||' = |\langle\varphi(v_0), v_0\rangle|$ mit $||v_0|| \leqslant 1$, so gilt

$$\varphi(v_0) = \langle\varphi(v_0), v_0\rangle \cdot v_0 = \pm||\varphi||' \cdot v_0$$

Lemma 3.5.3 Sei V ein endlichdimensionaler K-Vektorraum ($K \in \{\mathbb{R},\mathbb{C}\}$) mit Skalarprodukt $\langle\cdot,\cdot\rangle$ und $U \subseteq V$ ein Untervektorraum. Weiter sei $\varphi \in \mathbf{End}(V)$ mit $\varphi(U) \subseteq U$. Dann gilt $(\varphi\,|_U)^* = \varphi^*\,|_U$. Insbesondere folgt, dass $\varphi\,|_U$ hermitesch ist, falls φ hermitesch ist.

Satz 3.5.4 Sei V ein endlichdimensionaler K-Vektorraum ($K \in \{\mathbb{R},\mathbb{C}\}$) mit Skalarprodukt $\langle\cdot,\cdot\rangle$. Weiter sei $\varphi \in \mathbf{End}(V)$. Dann sind die folgenden Aussagen äquivalent:

(a) φ ist hermitesch.

(b) Es existiert eine Orthonormalbasis $v_1,\dots,v_n$ von V mit $\varphi(v_i) = a_i v_i$ und $a_i \in \mathbb{R}$.

Ist V ein $\mathbb{C}$-Vektorraum, so sind die Aussagen (a) und (b) äquivalent zu

(c) Für alle $v \in V$ ist $\langle\varphi(v), v\rangle$ reell.

Satz 3.5.5

(a) Sei $A \in \text{Mat}(n,n,\mathbb{C})$ hermitesch. Dann existiert eine unitäre matrix C, so dass $\overline{C^t}AC$ eine reelle Diagonalmatrix ist.

(b) Sei $A \in \text{Mat}(n,n,\mathbb{R})$ symmetrisch. Dann existiert eine orthogonale matrix C, so dass C^tAC eine reelle Diagonalmatrix ist.

3.6 Spektralnorm eines Endomorphismus (einer Matrix)

Definition 3.6.1 Sei $V = \mathbb{C}^n$ oder $\mathbb{R}^n$. Wir betrachten die Vektornorm $||\cdot||_2 : V \to \mathbb{R}$, die von Standardskalarprodukt induziert ist:

$$||x||_2 = \sqrt{|x_1|^2 + \ldots + |x_n|^2} \quad \text{für } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Sei $||\cdot||_2' : \mathbf{End}(V) \to \mathbb{R}$ die Operatorennorm, die von der Vektornorm $||\cdot||_2$ induziert ist, also gilt

$$||\varphi||_2' := \max_{||x||_2=1} ||\varphi(x)||_2.$$

Die Norm $||\cdot||_2'$ heißt *Spektralnorm* von φ. Die analoge Definition gilt für Matrizen.

Lemma 3.6.2 Sei V ein endlichdimensionaler K-Vektorraum ($K = \mathbb{R}$ oder $\mathbb{C}$) und sei $\varphi \in \mathbf{End}(V)$ ein beliebiger Endomorphismus. Dann ist der Endomorphismus $\psi := \varphi^*\varphi$ hermitesch und alle seine Eigenwqerte sind größer gleich 0.

Satz 3.6.3 Sei $V = \mathbb{C}^n$ oder $\mathbb{R}^n$. Für jeden Endomorphismus $\varphi \in \mathbf{End}(V)$ gilt

$$||\varphi||_2' = \sqrt{a_{\max}},$$

wobei $a_{\max}$ der maximale Eigenwert von $\varphi^*\varphi$ ist.

Folgerung 3.6.4 Für jede Matrix $A \in \text{Mat}(n,n,K)$ gilt

$$||A||_2' = \sqrt{a_{\max}},$$

wobei $a_{\max}$ der maximale Eigenwert von $\overline{A^t}A$ ist.

3.7 Positiv definite (positiv semidefinite) hermitesche Endomorphismen und Matrizen

Definition 3.7.1 Sei $K = \mathbb{R}$ oder $\mathbb{C}$.

(a) Sei V ein endlichdimensionaler K-Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$.
Sei $\varphi \in \mathbf{End}(V)$ ein beliebiger Endomorphismus. Wir schreiben:

- $\varphi \geqslant 0$, wenn $\langle\varphi(v), v\rangle \geqslant 0$ für alle $v \in V$ ist, (φ ist *positiv semidefinit*)
- $\varphi > 0$, wenn $\langle\varphi(v), v\rangle > 0$ für alle $v \in V \setminus \{0\}$ ist. (φ ist *positiv definit*)

(b) Wir betrachten den K-Vektorraum K^n mit Standardskalarprodukt $\langle \cdot, \cdot \rangle$.
Sei $A \in \mathbf{Mat}(n,n,K)$ eine beliebige Matrix. Wir schreiben:

- $A \geqslant 0$, wenn $\langle AX, X\rangle \geqslant 0$ für alle $X \in K^n$ ist, (A ist *positiv semidefinit*)
- $A > 0$, wenn $\langle AX, X\rangle > 0$ für alle $X \in K^n \setminus \{0\}$ ist. (A ist *positiv definit*)

(c) Analoge Definitionen gelten für die Begriffe *negativ semidefinit und negativ definit.* Merken wir an, dass A positiv definit ist genau dann, wenn $-A$ negativ definit ist.

(d) Eine Matrix $A \in \mathbf{Mat}(n,n,K)$ heißt *indefinit*, wenn $X, Y \in K^n$ mit $\langle AX, X\rangle > 0$ und $\langle AY, Y\rangle < 0$ existieren.

Lemma 3.7.2 Für eine beliebige Matrix $A \in \mathbf{Mat}(n,n,K)$ sind (a) und (b) äquivalent:

(a) $A \geqslant 0$ $\qquad$ $(A > 0)$

(b) $X^t A\overline{X} \geqslant 0$ für alle $X \in K^n$ $\qquad$ ($X^t A\overline{X} > 0$ für alle $X \in K^n \setminus \{0\}$)

Für eine hermitesche Matrix $A \in \mathbf{Mat}(n,n,K)$ sind (a) und (b) dem (c) äquivalent:

(c) Alle Eigenwerte von A sind größer gleich 0. (Alle Eigenwerte von A sind größer als 0.)

Lemma 3.7.3 Wir betrachten den Vektorraum K^n mit dem Standardskalarprodukt $\langle \cdot, \cdot \rangle$. Sei $A \in \mathbf{Mat}(n,n,K)$ eine hermitesche Matrix. Es gilt $A > 0$ genau dann, wenn die Abbildung

$$\begin{aligned} \langle \cdot, \cdot \rangle_A : \quad & K^n \times K^n \to K, \\ \langle X, Y\rangle_A : \quad & = \langle AX, Y\rangle \end{aligned}$$

auch ein Skalarprodukt auf K^n definiert.

Definition 3.7.4 Für jede Matrix

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \in \mathrm{Mat}(n,n,K).$$

und jedes $r = 1, \dots, n$ heißt die Zahl

$$\delta_r = \delta_r(A) = \det \begin{pmatrix} a_{11} & \dots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \dots & a_{rr} \end{pmatrix}$$

führender Hauptminor von A. Offenbar ist $\delta_n = \det(A)$.

Lemma 3.7.5 Sei $K = \mathbb{C}$ oder $\mathbb{R}$ und $A \in \mathbf{Mat}(n,n,K)$ hermitesch. Weiterhin gelte für alle führenden Hauptminoren $\delta_r \neq 0$ $(r = 1, \dots, n)$. Dann existiert eine *unipotente* Matrix $U \in \mathbf{Mat}(n,n,K)$ (d.h. eine obere Dreiecksmatrix mit Diagonale 1), so dass gilt:

$$\overline{U^t} A U = \begin{pmatrix} \delta_1 & \dots & 0 \\ \vdots & \frac{\delta_2}{\delta_1} & \vdots \\ 0 & \dots & \frac{\delta_n}{\delta_{n-1}} \end{pmatrix} = \mathbf{diag}\big(\delta_1, \frac{\delta_2}{\delta_1}, \dots, \frac{\delta_n}{\delta_{n-1}}\big).$$

Beweis. Sei
$$A = \begin{pmatrix} B & v \\ \overline{v^t} & k \end{pmatrix},$$
wobei $B \in \mathrm{Mat}(n-1, n-1, K)$ hermitesch, $v \in K^n$ und $k \in K$ ist. Wegen $\det(B) = \delta_{n-1} \neq 0$ ist die Matrix B invertierbar und es gilt $B^{-1} = (\overline{B}^{-1})^t$. Es gilt
$$A = \begin{pmatrix} E & 0 \\ \overline{(B^{-1}v)^t} & 1 \end{pmatrix} \cdot \begin{pmatrix} B & 0 \\ 0 & a \end{pmatrix} \cdot \begin{pmatrix} E & B^{-1}v \\ 0 & 1 \end{pmatrix}$$
für $a = k - \overline{v^t}B^{-1}v$. Also existiert eine unipotente Matrix X mit
$$A = \overline{X^t} \cdot \begin{pmatrix} B & 0 \\ 0 & a \end{pmatrix} \cdot X \tag{3.7.1}$$
Wir haben $\delta_n = \det(A) = \det(B) \cdot a$. Also gilt $a = \frac{\delta_n}{\delta_{n-1}}$. Per Induktion existiert eine unipotente Matrix Y mit
$$B = \overline{Y^t} \begin{pmatrix} \delta_1 & \dots & 0 \\ \vdots & \frac{\delta_2}{\delta_1} & \vdots \\ 0 & \dots & \frac{\delta_{n-1}}{\delta_{n-2}} \end{pmatrix} Y$$
Dann gilt
$$\begin{pmatrix} B & 0 \\ 0 & a \end{pmatrix} = \begin{pmatrix} \overline{Y^t} & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \delta_1 & \dots & 0 \\ \vdots & \frac{\delta_2}{\delta_1} & \vdots \\ 0 & \dots & \frac{\delta_n}{\delta_{n-1}} \end{pmatrix} \cdot \begin{pmatrix} Y & 0 \\ 0 & 1 \end{pmatrix} = \overline{Z^t}DZ \tag{3.7.2}$$
Setzen wir (3.7.2) in (3.7.1):
$$A = \overline{X^t Z^t} DZX.$$
□

Satz 3.7.6 (Sylvester-Kriterium) Sei $A \in \mathbf{Mat}(n, n, K)$ eine hermitesche Matrix. Die Matrix A ist positiv definit genau dann, wenn alle führenden Hauptminoren von A positiv sind.

Satz 3.7.7 (Cholesky-Zerlegung) Eine hermitesche Matrix A ist genau dann positiv definit, wenn es eine Zerlegung $A = \overline{G^t}G$ gibt, wobei G eine obere Dreiecksmatrix mit positiven Diagonalanträgen ist.

Beweis. 1) Sei $A = \overline{G^t}G$, wobei G eine obere Dreiecksmatrix mit positiven Diagonaleinträgen ist. Dann gilt
$$\langle AX, X\rangle = \langle \overline{G^t}GX, X\rangle = \langle GX, GX\rangle > 0$$
für alle $X \neq 0$, weil $\det(G) \neq 0$ ist.

2) Sei $A > 0$. Nach dem Sylvester-Kriterium sind alle führende Hauptminoren von A positiv. Nach Lemma 3.7.5 existieren eine unipotente Matrix V und eine positiv definite Diagonalmatrix D, so dass gilt:
$$A = \overline{V^t}DV.$$
Wir setzen $G = \sqrt{D}V$. Dann gilt $A = \overline{G^t}G$. □

4 Quadratische Formen

4.1 Erste Beispiele

Siehe die Bezeichnungen in Punkt 3.1.

Lemma 4.1.1 Sei V ein n-dimensionaler Vektorraum. Seien $e = (e_1, \ldots, e_n)$ und $e' = (e'_1, \ldots, e'_n)$ zwei Basen von V. Sei C die Übergangsmatrix von e zu e'. Dann gilt für jeden Vektor $v \in V$:

$$[v]_e = C \cdot [v]_{e'}$$

Beweis. Das folgt aus $e \cdot [v]_e = v = e' \cdot [v]_{e'} = e \cdot C \cdot [v]_{e'}$. □

Spezialfall einer quadratischen Form. Hauptachsen von Hyperflächen.
Sei $V = \mathbb{R}^n$. Sei $A \in \mathrm{Mat}(n, n, \mathbb{R})$ eine symmetrische Matrix (also gilt $A = A^t$). Wir definieren eine *quadratische Form* $q : V \to \mathbb{R}$ durch

$$\begin{aligned} q(X) &= X^t A X \\ &= \sum_{1 \leqslant i,j \leqslant n} a_{ij} x_i x_j = \sum_{i=1}^{n} a_{ii} x_i^2 + 2 \sum_{1 \leqslant i < j \leqslant n} a_{ij} x_i x_j \end{aligned}$$

für

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in V.$$

Für $d \in \mathbb{R}$ betrachten wir die *Hyperfläche*

$$\mathbf{F}_d := \{X \in V \mid q(X) = d\}.$$

(Setzen wir beispielsweise $A = E$, so ist $\mathbf{F}_1 = \{X \in V \mid \sum_{i=1}^{n} x_i^2 = 1\}$ die Einheitssphäre.)

Da A eine symmetrische reelle Matrix ist, existiert (siehe Satz 3.5.5) eine orthogonale Matrix $C \in \mathcal{O}_n$, so dass gilt:

$$C^t A C = D = \mathbf{diag}(d_1, \ldots, d_n).$$

Dann ist $A = CDC^t$, und es gilt

$$q(X) = X^t A X = X^t C \cdot D \cdot C^t X = Y^t D Y = \sum_{i=1}^{n} d_i y_i^2,$$

wobei

$$Y = C^t X$$

ist. Also gilt

$$X = CY$$

• Sei e die Standardbasis von V und e' die Basis von V, so dass die Übergangsmatrix von e zu e' gleich C ist. Die Vektoren der Basis e' heißen *Hauptachsen* der Hyperfläche $\mathbf{F}_d$.

Nach Lemma 4.3.1 gilt:

$\boxed{Y \text{ sind Koordinaten des Vektors } X \text{ in der Basis } e'.}$

$$\begin{aligned}\mathbf{F}_d &= \{X \in V \mid X^t A X = d\}\\ &= C \cdot \{Y \in V \mid Y^t D Y = \sum_{i=1}^{n} d_i y_i^2 = d\}.\end{aligned}$$

Die Spalten von C sind orthonormierte Eigenvektoren von A; die Diagonale von D besteht aus den Eigenwerten von A.

Klassifikation von $\mathbf{F}_d$ für $n = 2$.
Sei $A = \begin{pmatrix} a & c \\ c & b \end{pmatrix}$. Dann gilt

$$\begin{aligned}\mathbf{F}_d &= \{X \in \mathbb{R}^2 \mid a x_1^2 + 2c x_1 x_2 + b x_2^2 = d\}\\ &= C \cdot \{Y \in \mathbb{R}^2 \mid d_1 y_1^2 + d_2 y_2^2 = d\}.\end{aligned}$$

Fall 1. Sei $d \neq 0$.
Dann ist

$$\mathbf{F}_d = C \cdot \{Y \in \mathbb{R}^2 \mid \frac{d_1}{d} y_1^2 + \frac{d_2}{d} y_2^2 = 1\}.$$

Wir fügen folgende Bezeichnungen ein:

$$\begin{aligned}\alpha &= \frac{d_1 + d_2}{d} = \frac{\mathrm{Spur}(D)}{d} = \frac{\mathrm{Spur}(A)}{d},\\ \beta &= \frac{d_1 \cdot d_2}{d^2} = \frac{\det(D)}{d^2} = \frac{\det(A)}{d^2}.\end{aligned}$$

$\boxed{\textit{Fall 1.1.}}$ Seien $\beta > 0$, $\alpha \leqslant 0$.
Dann gilt $\frac{d_1}{d} < 0$ und $\frac{d_2}{d} < 0$. Deswegen ist $\mathbf{F}_d = \emptyset$.

$\boxed{\textit{Fall 1.2.}}$ Seien $\beta > 0$, $\alpha > 0$.
Dann gilt $\frac{d_1}{d} > 0$ und $\frac{d_2}{d} > 0$. Deswegen ist

$$\mathbf{F}_d = C \cdot \{Y \in \mathbb{R}^2 \mid \left(\sqrt{\frac{d_1}{d}} \cdot y_1\right)^2 + \left(\sqrt{\frac{d_2}{d}} \cdot y_2\right)^2 = 1\}$$

eine Ellipse.

$\boxed{\textit{Fall 1.3.}}$ Sei $\beta < 0$.
Dann haben $\frac{d_1}{d}$ und $\frac{d_2}{d}$ verschiedene Vorzeichen. O.B.d.A gilt $\frac{d_1}{d} > 0$ und $\frac{d_2}{d} < 0$. Dann ist

$$\mathbf{F}_d = C \cdot \{Y \in \mathbb{R}^2 \mid \left(\sqrt{\frac{d_1}{d}} \cdot y_1\right)^2 - \left(\sqrt{-\frac{d_2}{d}} \cdot y_2\right)^2 = 1\}$$

eine Hyperbel.

Fall 2. Sei $d = 0$.
In dem Fall ist $\mathbf{F}_d$ ein Punkt, eine Gerade, ein Paar von Geraden, oder $\mathbb{R}^2$.

4.2 Allgemeine Definitionen und Sätze

Definition 4.2.1 Sei K ein beliebiger Körper und sei V ein K-Vektorraum. Eine Abbildung $\beta : V \times V \to K$ heißt *bilineare Form* auf V, wenn für alle $u, v, w, z \in V$ und $\lambda \in K$ gelten:

(a) $\beta(u+v, w) = \beta(u, w) + \beta(v, w)$,

(b) $\beta(u, w+z) = \beta(u, w) + \beta(u, z)$,

(c) $\lambda\beta(v, u) = \beta(\lambda v, u) = \beta(v, \lambda u)$.

Die Bilinearform β heißt *symmetrisch*, wenn $\beta(u, v) = \beta(v, u)$ für alle $u, v \in V$ gilt.

Definition 4.2.2 Sei $\beta : V \times V \to K$ eine symmetrische Bilinearform auf einem K-Vektorraum V. Die Abbildung

$$q_\beta : V \to K, \quad q_\beta(X) = \beta(X, X)$$

heißt *mit β assoziierte quadratische Form* auf V.

Beispiel und die Bezeichnung q_A.
Sei $A \in \mathbf{Mat}(n, n, K)$ eine symmetrische Matrix. Die Abbildung $\beta : K^n \times K^n \to K$,

$$\beta(X, Y) = X^t AY,$$

ist eine symmetrische Bilinearform auf K^n. Man kann leicht beweisen, dass alle symmetrischen Bilinearformen auf K^n eine solche Gestalt haben. Die Abbildung $q_\beta : K^n \to K$,

$$q_\beta(X) = X^t AX,$$

ist mit β assoziierte quadratische Form auf K^n. Wir schreiben q_A statt q_β und sagen, dass q_A mit A assoziiert ist.

Definition 4.2.3 Zwei quadratische Formen $q_\beta : V \to K$ und $q_{\beta'} : V \to K$ heißen *äquivalent*, wenn ein Isomorphismus $\varphi : V \to V$ existiert, so dass $q_{\beta'}(X) = q_\beta(\varphi(X))$.

Man schreibt $q_\beta \sim q_{\beta'}$.

Lemma 4.2.4 Seien A und A' zwei symmetrische Matrizen aus $\mathbf{Mat}(n, n, K)$. Die folgenden zwei Bedingungen sind äquivalent:

(1) Die quadratischen Formen $q_{A'}$ und q_A sind äquivalent: $q_{A'} \sim q_A$.

(2) Es existiert eine Matrix $S \in \mathbf{GL}_n(K)$, so dass $q_{A'}(X) = q_A(SX)$ für alle $X \in K^n$ gilt.

Wenn $\mathrm{Char}(K) \neq 2$ ist (d.h. $1 + 1 \neq 0$ in K), dann sind diese Bedingungen der Bedingung (3) äquivalent:

(3) Es existiert eine Matrix $S \in \mathbf{GL}_n(K)$, so dass $A' = S^t AS$ gilt.

Satz 4.2.5 Sei K ein Körper und sei A eine symmetrische Matrix aus $\mathbf{Mat}(n,n,K)$. Dann existiert eine Diagonalmatrix $D \in \mathbf{Mat}(n,n,K)$, so dass $q_D \sim q_A$ gilt.

Beweis. Wenn $\mathrm{Char}(K) = 2$ ist, dann ist $q_A \sim q_D$ für $D = \mathbf{diag}(a_{11}, \ldots, a_{nn})$.

Sei $\mathrm{Char}(K) \neq 2$. Die Idee ist, die Matrix A mit Hilfe einiger Tranformationen der Form

$$A \to E_{ij}(\alpha) A E_{ji}(\alpha)$$

zu vereinfachen. Dabei werden i, j und α so gewählt, dass auf einer gewünschten nicht Diagonalstelle 0 erscheinen wird, oder auf einer gewünschten Diagonalstelle kein 0 erscheinen wird. Nach einigen Schritten bekommen wir eine Diagonalmatrix D. Wir illustrieren diese Idee mit folgendem Beispiel:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}$$

$$\downarrow$$

$$E_{21}(-2) A E_{12}(-2)$$

$$\downarrow$$

$$E_{31}(-3) E_{21}(-2) A E_{12}(-2) E_{13}(-3)$$

$$\downarrow$$

$$E_{23}(-1) E_{31}(-3) E_{21}(-2) A E_{12}(-2) E_{13}(-3) E_{32}(-1)$$

$$\downarrow$$

$$E_{32}(2) E_{23}(-1) E_{31}(-3) E_{21}(-2) A E_{12}(-2) E_{13}(-3) E_{32}(-1) E_{23}(2)$$
$$= \mathbf{diag}(1, -1, 1).$$

Dann ist

$$S = E_{12}(-2) E_{13}(-3) E_{32}(-1) E_{23}(2) = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 2 \\ 0 & -1 & -1 \end{pmatrix}.$$

$\square$

Definition 4.2.6 Sei A eine symmetrische Matrix aus $\mathbf{Mat}(n,n,\mathbb{K})$. Die Zahl $k = \mathrm{Rang}(A)$ heißt *Rang* der quadratischen Form q_A.

Lemma 4.2.7 Die Ränge von zwei äquivalenten quadratischen Formen q_A und q_B über einem Körper K mit $\mathrm{Char}(K) \neq 2$ sind gleich.

4.3 Äquivalenz von quadratischen Formen über $\mathbb{C}$

Satz 4.3.1 Sei A eine symmetrische Matrix aus $\mathbf{Mat}(n,n,\mathbb{C})$. Sei

$$\begin{aligned} q_A : \quad & \mathbb{C}^n \to \mathbb{C}, \\ & X \mapsto X^t A X \end{aligned}$$

die mit A assoziierte quadratische Form. Dann existiert eine Matrix $S \in \mathbf{GL}_n(\mathbb{C})$ und eine ganze Zahl $0 \leqslant k \leqslant n$, so dass für alle $Y \in \mathbb{C}^n$ gilt:

$$q_A(SY) = Y_1^2 + \cdots + Y_k^2.$$

Mit anderen Worten: Es gilt $q_D \sim q_A$ für eine Diagonalmatrix

$$D = \begin{pmatrix} E_k & \\ & O_{n-k} \end{pmatrix}.$$

Die Zahl k ist eindeutig, d.h. k hängt nicht von der Wahl von S ab. Es gilt $k = \mathrm{Rank}(A)$.

4.4 Äquivalenz von quadratischen Formen über $\mathbb{R}$

Lemma 4.4.1 Sei A eine symmetrische Matrix aus $\mathbf{Mat}(n,n,\mathbb{R})$. Sei

$$\begin{aligned} q_A : \quad & \mathbb{R}^n \to \mathbb{R}, \\ & X \mapsto X^t A X \end{aligned}$$

die mit A assoziierte quadratische Form. Dann existiert eine Unterraum-Zerlegung

$$\mathbb{R}^n = V_+ \oplus V_- \oplus V_0$$

mit den Eigenschaften:

(1) Ist $X \in V_+ \setminus \{0\}$, dann gilt $q_A(X) > 0$.

(2) Ist $X \in V_- \setminus \{0\}$, dann gilt $q_A(X) < 0$.

(3) Ist $X \in V_0$, dann gilt $q_A(X) = 0$.

(4) Liegen X und Y in zwei verschiedenen aus drei Untervektorräumen V_+, V_-, V_0, dann gilt $X^t A Y = 0$.

Satz 4.4.2 (Trägheitsgesetz von Sylvester)
Sei A eine symmetrische Matrix aus $\mathbf{Mat}(n,n,\mathbb{R})$. Sei

$$\begin{aligned} q_A : \quad & \mathbb{R}^n \to \mathbb{R}, \\ & X \mapsto X^t A X \end{aligned}$$

die mit A assoziierte quadratische Form. Dann existiert eine Matrix $S \in \mathrm{GL}_n(\mathbb{R})$ und es existieren einige Zahlen $k, \ell, m \in \mathbb{N} \cup \{0\}$, so dass $k + \ell + m = n$ ist und für alle $Y \in \mathbb{R}^n$ gilt:

$$q_A(SY) = Y_1^2 + \cdots + Y_k^2 - Y_{k+1}^2 - \cdots - Y_{k+\ell}^2.$$

Mit anderen Worten: Es gilt $q_D \sim q_A$ für eine Diagonalmatrix

$$D = \begin{pmatrix} E_k & & \\ & -E_\ell & \\ & & O_m \end{pmatrix}.$$

Die Zahlen k, l, m sind eindeutig, d.h. sie hängen nicht von der Wahl von S ab.

Definition 4.4.3 Sei A eine symmetrische Matrix aus $\mathbf{Mat}(n, n, \mathbb{R})$. Sei

$$\begin{aligned} q_A : \quad & \mathbb{R}^n \to \mathbb{R}, \\ & X \mapsto X^t A X \end{aligned}$$

die mit A assoziierte quadratische Form. Das Tripel (k, l, m) aus dem Satz 4.4.2 heißt *Signatur* der reellen quadratischen Form q_A.

5 Jordan-Formen

5.1 Kern-Zerlegung

Bezeichnungen.

(1) Für einen Ring K sei $K[x]$ der Ring aller Polynome von x mit Koeffizienten aus K.

(2) Für einen Vektorraum V sei $\mathbf{0} : V \to V$ die Abbildung, die alle Vektoren von V nach 0_V abbildet.

Den folgenden Satz kennen wir aus Lineare Algebra I:

Satz 5.1.1 (Cayley-Hamilton) Sei V ein endlichdimensionaler K-Vektorraum und sei $\varphi : V \to V$ eine lineare Abbildung. Dann gilt

$$\chi_\varphi(\varphi) = \mathbf{0}.$$

Bemerkung. Es existieren andere Polynome $p(x) \in K[x]$ mit der Eigenschaft $p(\varphi) = \mathbf{0}$. Jedes Polynom mit dieser Eigenschaft heißt *Annulator* von φ. Später werden wir alle Annulatoren von φ beschreiben.

Satz 5.1.2 (Kern-Zerlegung)
Sei V ein Vektorraum über einem Körper K und $\varphi : V \to V$ eine lineare Abbildung. Sei $p(x) \in K[x]$ ein Annulator von φ. Wenn $p(x) = p_1(x) \cdot p_2(x)$ gilt, wobei $p_1(x), p_2(x) \in K[x]$ zwei teilerfremde Polynome sind, dann gilt

$$V = \mathrm{Ker}(p_1(\varphi)) \oplus \mathrm{Ker}(p_2(\varphi)).$$

Folgender Satz wird später bewiesen:

Satz 5.1.3 Sei $p(x) \in \mathbb{C}[x]$ ein komplexes Polynom des Grades $n \geqslant 1$. Dann existieren komplexe Zahlen $\lambda_1, \dots, \lambda_s$ und natürliche Zahlen $k_1, \dots, k_s$, so dass $k_1 + \dots + k_s = n$ und

$$p(x) = (x - \lambda_1)^{k_1} \dots (x - \lambda_s)^{k_s}$$

gelten. Diese Zerlegung ist eindeutig.

Die Zahl k_i heißt *Vielfachheit* der Nullstelle λ_i, $i = 1, \dots, s$.

Folgerung 5.1.4 Sei V ein endlichdimensionaler Vektorraum über $\mathbb{C}$ und $\varphi : V \to V$ eine lineare Abbildung. Wir schreiben $\chi_\varphi(x) = (x - \lambda_1)^{k_1} \dots (x - \lambda_s)^{k_s}$, wobei $\lambda_1, \dots, \lambda_s$ paarweise verschiedene Nullstellen des charakteristischen Polynoms χ_φ sind. Dann gilt

$$V = \operatorname{Ker}\Big((\varphi - \lambda_1 \cdot \mathrm{id})^{k_1}\Big) \oplus \dots \oplus \operatorname{Ker}\Big((\varphi - \lambda_s \cdot \mathrm{id})^{k_s}\Big).$$

Definition 5.1.5 Sei U ein K-Vektorraum. Eine lineare Abbildung $\psi : U \to U$ heißt *nilpotent*, wenn ein $k \in \mathbb{N}$ mit der Eigenschaft $\psi^k = \mathbf{0}$ existiert.

Bemerkung 5.1.6 In der Situation der Folgerung 5.1.4 sei

$$U_i = \operatorname{Ker}\Big((\varphi - \lambda_i \cdot \mathrm{id})^{k_i}\Big),$$

$i = 1, \dots, s$. Dann gilt

$$V = U_1 \oplus \dots \oplus U_s.$$

Sei $\psi_i := (\varphi - \lambda_s \cdot \mathrm{id})|_{U_i}$. Dann ist

$$\psi_i : U_i \to U_i$$

nilpotent. Wir werden zeigen, dass jede nilpotente Abbildung eine einfache Darstellungsmatrix bezüglich einer passenden Basis hat. Danach werden wir die passenden Basen von allen U_i vereinigen und erhalten eine Basis e' von V. Wir werden zeigen, dass die Darstellungsmatrix von φ bezüglich e' einfach ist.

5.2 Minimalpolynom eines Endomorphismus (einer Matrix)

Bemerkung. Sei R ein kommutativer Ring mit 1 und sei $a \in R$. Dann ist die Menge

$$aR := \{a \cdot r \mid r \in R\}$$

ein Ideal in R.

Definition 5.2.1 Sei R ein kommutativer Ring mit 1. Ein Ideal I in R heißt *Hauptideal*, wenn $I = aR$ für ein $a \in R$ ist. In dem Fall heißt a *Erzeuger* von I.

Ein nichtnullsches Polynom $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_0$ heißt *normiert*, wenn $a_n = 1$ ist.

Lemma 5.2.2 Sei K ein Körper. Dann gilt:

(1) Jedes Ideal I in dem Ring $K[x]$ ist ein Hauptideal.

(2) Ist $I \neq \{0\}$, dann ist der normierte Erzeuger von I eindeutig bestimmt.

Satz 5.2.3 Sei K ein Körper und sei V ein endlichdimensionaler Raum über K. Sei $\varphi : V \to V$ ein Endomorphismus. Dann gelten:

(a) Die Menge $\operatorname{Ann}(\varphi) := \{f(x) \in K[x] \mid f(\varphi) = \mathbf{0}\}$ ist ein Ideal in $K[x]$.

(b) Es gibt ein eindeutig bestimmtes normiertes Polynom $m_\varphi(x) \neq 0$ mit der Eigenschaft

$$\operatorname{Ann}(\varphi) = m_\varphi(x) \cdot K[x].$$

(Wir nennen $m_\varphi(x)$ das *Minimalpolynom* von φ.)

(c) Das Minimalpolynom von φ ist ein Teiler des charakteristischen Polynoms von φ:

$$m_\varphi(x) \,\Big|\, \chi_\varphi(x).$$

(d) Ist a ein Eigenwert von φ, dann gilt $m_\varphi(a) = 0$.

Beispiel.

(1) Sei

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

Dann gilt $\chi_A(x) = (x-\alpha)^3$ und $m_A(x) = (x-\alpha)^2$.

(2) Sei A eine Block-Diagonalmatrix, die aus quadratischen Matrizen $A_1, \ldots, A_s$ gebaut ist:

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{pmatrix}$$

Dann gilt

$$m_A(x) = \mathbf{kgV}\,\{m_{A_1}(x), \ldots, m_{A_s}(x)\}.$$

Hier steht **kgV** für das kleinste gemeinsame Vielfache der Polynome.
Z.B ist $f(x) = (x-1)^3(x-2)^5$ und $g(x) = (x-1)^2(x-2)^7(x-5)$, so ist

$$\mathbf{kgV}\{f(x), g(x)\} = (x-1)^3(x-2)^7(x-5).$$

5.3 Normalformen nilpotenter Abbildungen

Satz 5.3.1 Sei K ein beliebiger Körper, sei V ein endlichdimensionaler K-Vektorraum und sei $\varphi : V \to V$ eine nilpotente lineare Abbildung. Dann existiert eine Basis $v = \{v_1, \ldots, v_n\}$ von V, so dass die Darstellungsmatrix von φ in der Basis v folgende Jordansche Normalform hat:

$$[\varphi]_v^v = \begin{pmatrix} J_1 & & & O \\ & J_2 & & \\ & & \ddots & \\ O & & & J_k \end{pmatrix},$$

wobei jeder Jordan-Block J_i folgende Form hat:

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

Der Beweis dieses Satzes wird klar, nachdem wir einige Beispiele, Definitionen und den Algorithmus 5.3.7 formulieren.

Beispiel 5.3.2 Sei v eine Basis, so dass

$$[\varphi]_v = \left(\begin{array}{ccc|cc|c} 0 & 1 & & & & \\ & 0 & 1 & & & \\ & & 0 & & & \\ \hline & & & 0 & 1 & \\ & & & & 0 & \\ \hline & & & & & 0 \end{array}\right)$$

gilt. Dann können wir die Basisvektoren $v_1, \ldots, v_6$ in die folgende Tabelle eintragen:

$$\begin{array}{cccc} & \overset{\varphi}{\curvearrowright} & \overset{\varphi}{\curvearrowright} & \overset{\varphi}{\curvearrowright} \\ \boxed{v_3} & \boxed{v_2} & \boxed{v_1} & 0 \\ & \boxed{v_5} & \boxed{v_4} & 0 \\ & & \boxed{v_6} & 0 \end{array}$$

Definition 5.3.3 Eine φ-Tabelle ist eine Tabelle der Form

v_s	v_{s-1}	v_{s-2}	$\cdots$	v_1
	v_t	v_{t-1}	$\cdots$	v_{s+1}
		$\cdots$	$\cdots$	$\cdots$
		v_m	$\cdots$	v_{l+1}

so dass
(1) die Tabelle rechtsbündig ist;
(2) ihre Elemente v_i Vektoren aus V sind;
(3) für jede Zeile $v_p, v_{p-1}, \dots, v_q$ dieser Tabelle $v_p \overset{\varphi}{\mapsto} v_{p-1} \overset{\varphi}{\mapsto} \dots \overset{\varphi}{\mapsto} v_q \overset{\varphi}{\mapsto} 0$ gilt.

Unser Ziel: Wir möchten für eine gegebene nilpotente Abbildung eine φ-Tabelle konstruieren, so dass folgendes gilt:

(A) die Vektoren $v_1, \dots, v_m$ aus der Tabelle sind linear unabhängig,

(B) $\mathrm{Span}(v_1, \dots, v_m) = V$.

Lemma 5.3.4 Wenn die Vektoren in der letzten Spalte der φ-Tabelle (das sind $v_1, v_{s+1}, \dots, v_{l+1}$) linear unabhängig sind, dann sind alle Vektoren der Tabelle linear unabhängig.

Beweis mit der Tabelle aus dem Beispiel 5.3.2. Unter der Voraussetzung, dass v_3, v_5, v_6 linear unabhängig sind, sei eine Linearkombination von $v_1, \dots, v_6$ gleich 0:

$$\begin{array}{llll} \alpha_1 v_1 + & \alpha_2 v_2 + & \alpha_3 v_3 & \\ & + \alpha_4 v_4 + & \alpha_5 v_5 & \\ & & + \alpha_6 v_6 & = 0 \end{array}$$

Indem wir φ^2 auf diese Gleichung anwenden, erhalten wir $\alpha_1 v_3 = 0$, also $\alpha_1 = 0$.
Durch Anwendung von φ bekommen wir $\alpha_2 v_3 + \alpha_4 v_5 = 0$, also $\alpha_2 = \alpha_4 = 0$.
Schließlich erhalten wir $\alpha_3 = \alpha_5 = \alpha_6 = 0$ aus der linearen Unabhängigkeit von v_3, v_5, v_6. $\square$

Definition 5.3.5 (Transformationen von φ-Tabellen)
(T1) Multiplikation einer Zeile der Tabelle mit einer Zahl $\alpha \neq 0$.
(T2) Addition eines Fragments der längeren Zeile zur kürzeren Zeile:

$\cdots$	u	$\varphi(u)$	$\cdots$	$\varphi^k(u)$
	v	$\varphi(v)$	$\cdots$	$\varphi^k(v)$

$\longrightarrow$

$\cdots$	u	$\varphi(u)$	$\cdots$	$\varphi^k(u)$
	$u+v$	$\varphi(u+v)$	$\cdots$	$\varphi^k(u+v)$

(T3) Wenn nach einer solchen Addition der letzte Vektor einer Zeile Null wird, löschen wir das Kästchen mit dem Vektor und schieben die Zeile einen Schritt nach rechts:

$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
	x	y	z	0

$\longrightarrow$

$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
		x	y	z

Lemma 5.3.6

(1) Die Transformationen (T1)–(T3) verändern den Span von Vektoren einer φ-Tabelle nicht.

(2) Mit Hilfe dieser Transformationen kann man die letzten Vektoren einer φ-Tabelle linear unabhängig machen.
(Nach Lemma 5.3.4 werden dann alle Vektoren der φ-Tabelle linear unabhängig.)

Algorithmus 5.3.7 Nehmen wir die Standardbasis $e_1, \ldots, e_n$ von V und bilden eine φ-Tabelle mit den ersten Vektoren $e_1, \ldots, e_n$.

e_1	$\varphi(e_1)$	$\varphi^2(e_1)$	$\cdots$	$\varphi^{k_1}(e_1)$
	e_2	$\varphi(e_2)$	$\cdots$	$\varphi^{k_2}(e_2)$
		$\cdots$	$\cdots$	$\cdots$
		e_n	$\cdots$	$\varphi^{k_n}(e_n)$

Klar, dass die Bedingung **(B)** erfüllt ist, obwohl es möglich ist, dass die Bedingung **(A)** nicht erfüllt ist. Mit Hilfe von Lemma 5.3.6 transformiere die φ-Tabelle in eine andere, für die die Bedingung **(A)** auch erfüllt wird. Die Vektoren der letzten Tabelle bilden eine Basis v, die im Satz 5.3.1 gesucht wurde.

Beispiel 5.3.8 Sei $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$, wobei $\varphi(X) = A \cdot X$ und

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

ist. Berechnen wir die Basis v aus dem Satz 5.3.1 für φ:

1 0 0	0 1 2
0 1 0	
0 0 1	

1 0 0	0 1 2
	0 1 0
	0 0 1

1 0 0	0 1 2
	0 0 −2
	0 0 1

1 0 0	0 1 2
	0 0 1

Also,

$$\begin{array}{lclcl} v_2 = (1,0,0) & \overset{\varphi}{\mapsto} & v_1 = (0,1,2) & \overset{\varphi}{\mapsto} & (0,0,0) \\ & & v_3 = (0,0,1) & \overset{\varphi}{\mapsto} & (0,0,0) \end{array}.$$

Jordansche Normalform von A ist

$$J = [\varphi]_v = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Übergangsmatrix von e zu v ist

$$T = \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 2 & 0 & 1 \end{pmatrix}.$$

Es gilt

$$[\varphi]_v = T^{-1}[\varphi]_e T,$$

also gilt

$$J = T^{-1}AT.$$

5.4 Jordansche Normalform der beliebigen linearen Abbildung

Definition 5.4.1′.

(1) Für $\alpha \in \mathbb{C}$ und $m \in \mathbb{N}$ heißt die $m \times m$-Matrix

$$J(\alpha, m) = \begin{pmatrix} \alpha & 1 & & & 0 \\ & \alpha & 1 & & \\ & & \ddots & \ddots & \\ & & & \alpha & 1 \\ 0 & & & & \alpha \end{pmatrix}$$

Jordan-Block der Größe m mit α auf der Diagonale.

(2) Eine blockdiagonale Matrix $J \in \mathbf{Mat}(n, n, \mathbb{C})$ heißt *Jordan-Matrix*, wenn ihre Blöcke Jordan-Blöcke sind:

$$J = \begin{pmatrix} J_1 & & & O \\ & J_2 & & \\ & & \ddots & \\ O & & & J_\ell \end{pmatrix},$$

d.h. wenn $J_i = J(\alpha_i, m_i)$ für einige $\alpha_i \in \mathbb{C}$ und $m_i \in \mathbb{N}$ ist, $i = 1, \ldots, \ell$.

Satz 5.4.1 Sei V ein Vektorraum über $\mathbb{C}$ von endlicher Dimension und $\varphi : V \to V$ eine lineare Abbildung. Dann existiert eine Basis $\mathcal{B}$ von V, so dass die Darstellungsmatrix von φ in der Basis $\mathcal{B}$ eine Jordan-Matrix ist: $[\varphi]_{\mathcal{B}} = J$.

Beweis. Sei $\chi_\varphi(\lambda) = (\lambda - \lambda_1)^{k_1} \ldots (\lambda - \lambda_s)^{k_s}$ das charakteristische Polynom von φ. Nach Folgerung 5.1.4 gilt $V = V_1 \oplus \cdots \oplus V_s$, wobei $V_i = \mathrm{Ker}(\varphi - \lambda_i \cdot \mathrm{id})^{k_i}$ ist. Für $i = 1, \ldots, s$ betrachten wir die Einschränkung

$$\psi_i := (\varphi - \lambda_i \cdot \mathrm{id})_{|V_i}$$

Dann ist $\psi_i : V_i \to V_i$ eine nilpotente Abbildung, für die $\psi_i^{k_i} = \mathbf{0}$ gilt. Nach Satz 5.3.1 existiert eine Basis $\mathcal{B}_i$ von V_i, so dass $[\psi_i]_{\mathcal{B}_i}$ die Jordansche Form hat. Sei $\mathcal{B}$ die Vereinigung der Basen $\mathcal{B}_i$, $i = 1, \ldots, s$. Dann ist $\mathcal{B}$ die gesuchte Basis von V und es gilt

$$[\varphi]_{\mathcal{B}} = \begin{pmatrix} [\varphi|_{V_1}]_{\mathcal{B}_1} & & & O \\ & [\varphi|_{V_2}]_{\mathcal{B}_2} & & \\ & & \ddots & \\ O & & & [\varphi|_{V_s}]_{\mathcal{B}_s} \end{pmatrix},$$

wobei

$$[\varphi|_{V_i}]_{\mathcal{B}_i} = [\psi_i]_{\mathcal{B}_i} + \lambda_i E_i$$

und E_i die $n_i \times n_i$-Einheitsmatrix mit $n_i = \dim(V_i)$ ist.
Merken wir an, dass $[\varphi|_{V_i}]_{\mathcal{B}_i}$ aus mehreren Jordan-Blöcken J_j bestehen kann. □

Satz 5.4.2 Für jede Matrix $A \in \mathbf{Mat}(n, n, \mathbb{C})$ existiert eine Matrix $T \in \mathbf{GL}_n(\mathbb{C})$ und eine Jordan-Matrix J, so dass

$$T^{-1}AT = J$$

gilt. Diese J ist eindeutig bis auf die Reihenfolge ihrer Jordan-Blöcke.
Für $\alpha \in \mathbb{C}$ und $k \in \mathbb{N}$ ist die Anzahl der Jordan-Blöcke $J(\alpha, k)$ in J gleich

$$\mathbf{rk}\big((A - \alpha E)^{k-1}\big) + \mathbf{rk}\big((A - \alpha E)^{k+1}\big) - 2 \cdot \mathbf{rk}\big((A - \alpha E)^{k}\big).$$

Definition 5.4.3 In der Situation des Satzes 5.4.2 heißt J *Jordanform* der Matrix A; die Matrix T heißt *Jordan-Konjugat* für A.

Beispiel 5.4.4 Sei

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

Wir werden die Jordanform J von A und eine Matrix $T \in \mathbf{GL}_n(\mathbb{C})$ finden, so dass $J = T^{-1}AT$ gilt.

Schritt 1. (Charakteristisches Polynom zerlegen)
Es gilt

$$\chi_A(\lambda) = (\lambda - 1)^2(\lambda + 1)^2.$$

Schritt 2. (Basen von Kern-Räumen finden)
Nach Schritt 1 und Satz 5.1.4 gilt

$$\mathbb{R}^4 = \mathrm{Ker}\big((A - E)^2\big) \oplus \mathrm{Ker}\big((A + E)^2\big).$$

(a) Wir finden eine Basis von $\mathrm{Ker}\big((A - E)^2\big)$:

$$(A - E)^2 = \begin{pmatrix} 2 & -4 & 0 & 2 \\ 4 & -6 & -2 & 4 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & 2 & -2 \end{pmatrix}^2 = \begin{pmatrix} -12 & 16 & 12 & -16 \\ -16 & 20 & 16 & -20 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Dann ist

$$\mathrm{Ker}\big((A - E)^2\big) = \{X \in \mathbb{R}^4 \mid ((A - E)^2 \cdot X = 0\} = \langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \rangle = \langle u_1, u_2 \rangle.$$

(b) Wir finden eine Basis von $\operatorname{Ker}\big((A+E)^2\big)$:

$$(A+E)^2 = \begin{pmatrix} 4 & -4 & 0 & 2 \\ 4 & -4 & -2 & 4 \\ 0 & 0 & 4 & -2 \\ 0 & 0 & 2 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 & 12 & -8 \\ 0 & 0 & 8 & -4 \\ 0 & 0 & 12 & -8 \\ 0 & 0 & 8 & -4 \end{pmatrix}.$$

Dann ist

$$\operatorname{Ker}\big((A+E)^2\big) = \{X \in \mathbb{R}^4 \mid ((A+E)^2 \cdot X = 0\} = \langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \rangle = \langle u_3, u_4 \rangle.$$

Schritt 3. (φ-Tabellen transformieren)

(a) Wir bilden die $(A-E)$-Tabelle mit Anfangsvektoren u_1, u_2 aus Schritt 2 (a):

$$\begin{array}{cccc|cccc|cccc}
 & & & \overset{A-E}{\curvearrowright} & & & & \overset{A-E}{\curvearrowright} & & & & \\
\hline
1 & 0 & 1 & 0 & 2 & 2 & 2 & 2 & 0 & 0 & 0 & 0 \\
\hline
0 & 1 & 0 & 1 & -2 & -2 & -2 & -2 & 0 & 0 & 0 & 0 \\
\hline
\end{array}$$

Wir transformieren diese $(A-E)$-Tabelle in eine andere $(A-E)$-Tabelle, deren Vektoren in der letzten Spalte linear unabhängig sind:

$$\begin{array}{|cccc|cccc|}
\hline
1 & 0 & 1 & 0 & 2 & 2 & 2 & 2 \\
\hline
0 & 1 & 0 & 1 & -2 & -2 & -2 & -2 \\
\hline
\end{array} \quad \overset{1}{\hookleftarrow}$$

$$\downarrow \text{(T2)}$$

$$\begin{array}{|cccc|cccc|}
\hline
1 & 0 & 1 & 0 & 2 & 2 & 2 & 2 \\
\hline
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
\hline
\end{array}$$

$$\downarrow \text{(T3)}$$

$$\begin{array}{|cccc|cccc|}
\hline
1 & 0 & 1 & 0 & 2 & 2 & 2 & 2 \\
\hline
 & & & & 1 & 1 & 1 & 1 \\
\hline
\end{array} \quad \overset{-1/2}{\hookleftarrow}$$

$$\downarrow \text{(T2)}$$

$$\begin{array}{|cccc|cccc|}
\hline
1 & 0 & 1 & 0 & 2 & 2 & 2 & 2 \\
\hline
\end{array}$$

Wir bezeichnen

$$b_1 = \begin{pmatrix} 2 \\ 2 \\ 2 \\ 2 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

(b) Wir bilden die $(A+E)$-Tabelle mit Anfangsvektoren u_3, u_4 aus Schritt 2 (b):

$$\begin{array}{cccc c cccc c cccc}
 & & & & \overset{A+E}{\curvearrowright} & & & & & \overset{A+E}{\curvearrowright} & & & & \\
1 & 0 & 0 & 0 & | & 4 & 4 & 0 & 0 & | & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & | & -4 & -4 & 0 & 0 & | & 0 & 0 & 0 & 0
\end{array}$$

Wir transformieren diese $(A+E)$-Tabelle in eine andere $(A+E)$-Tabelle, deren Vektoren in der letzten Spalte linear unabhängig sind:

$$\begin{array}{|cccc|cccc|}
\hline
1 & 0 & 0 & 0 & 4 & 4 & 0 & 0 \\
\hline
0 & 1 & 0 & 0 & -4 & -4 & 0 & 0 \\
\hline
\end{array} \quad 1$$

$$\downarrow \text{(T2)}$$

$$\begin{array}{|cccc|cccc|}
\hline
1 & 0 & 0 & 0 & 4 & 4 & 0 & 0 \\
\hline
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
\end{array}$$

$$\downarrow \text{(T3)}$$

$$\begin{array}{|cccc|cccc|}
\hline
1 & 0 & 0 & 0 & 4 & 4 & 0 & 0 \\
\hline
 & & & & 1 & 1 & 0 & 0 \\
\hline
\end{array} \quad -1/4$$

$$\downarrow \text{(T2)}$$

$$\begin{array}{|cccc|cccc|}
\hline
1 & 0 & 0 & 0 & 4 & 4 & 0 & 0 \\
\hline
\end{array}$$

Wir bezeichnen

$$b_3 = \begin{pmatrix} 4 \\ 4 \\ 0 \\ 0 \end{pmatrix}, \quad b_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Schritt 4. (Jordanform und Jordan-Konjugat aufschreiben)

$$J = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad T = \begin{pmatrix} b_1 & b_2 & b_3 & b_4 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 4 & 1 \\ 2 & 0 & 4 & 0 \\ 2 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{pmatrix}.$$

Bemerkung. Die Jordanform J einer Matrix A kann mehrere Jordan-Kästchen mit gleichen α auf der Diagonale besitzen:

$$J := \begin{pmatrix} \beta & 1 & & & & & \\ & \beta & & & & & \\ & & \alpha & 1 & & & \\ & & & \alpha & & & \\ & & & & \alpha & 1 & \\ & & & & & \alpha & 1 \\ & & & & & & \alpha \end{pmatrix}$$

Satz 5.4.5 Sei $A \in \mathbf{Mat}(n, n, \mathbb{C})$. Sei J die Jordan-Matrix von A und seien $\lambda_1, \dots, \lambda_s$ verschiedene Eigenwerte von A. Für $i = 1, \dots, s$ setzen wir

$$\ell_i := \max\{\ell \mid J \text{ bestitzt ein Kästchen } J(\lambda_i, \ell)\}.$$

Dann ist das Minimalpolynom von A gleich

$$m_A(\lambda) = (\lambda - \lambda_1)^{\ell_1} \cdot \ldots \cdot (\lambda - \lambda_s)^{\ell_s}.$$

5.5 Funktionen von Matrizen

Für die nächste Definition müssen wir uns an die Definition der Differenzierbarkeit einer Funktion $f : I \to \mathbb{C}$ mit $I \subseteq \mathbb{C}$ oder $I \subseteq \mathbb{R}$ erinnern.

Definition 5.5.1 Sei $f : I \to \mathbb{C}$ eine Funktion, wobei I eine Teilmenge von $\mathbb{C}$ oder $\mathbb{R}$ ist.

(a) Für einen Jordan-Block $J(\alpha, k)$ definieren wir

$$f(J(\alpha,k)) := \begin{pmatrix} f(\alpha) & \frac{f'(\alpha)}{1!} & \frac{f''(\alpha)}{2!} & \cdots & \frac{f^{(k-1)}(\alpha)}{(k-1)!} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \frac{f''(\alpha)}{2!} \\ & & & \ddots & \frac{f'(\alpha)}{1!} \\ & & & & f(\alpha) \end{pmatrix},$$

wenn $f(\alpha)$ und alle Ableitungen in der rechten Seite existieren.

(b) Für eine Jordanform $J = J(\alpha_1, n_1) \oplus \cdots \oplus J(\alpha_k, n_k)$ definieren wir

$$f(J) := f\big(J(\alpha_1, n_1)\big) \oplus \cdots \oplus f\big(J(\alpha_k, n_k)\big),$$

wenn alle Matrizen in der rechten Seite definiert sind.

(c) Sei $A \in \mathbf{Mat}_n(\mathbb{C})$. Nach Satz 5.4.2 existiert eine Jordan-Matrix J und eine Matrix $T \in \mathbf{GL}_n(\mathbb{C})$, so dass $J = T^{-1}AT$ gilt. Dann gilt $A = TJT^{-1}$. Wir definieren

$$f(A) := Tf(J)T^{-1},$$

wenn $f(J)$ existiert.

Bemerkung. Für ein Polynom $p(x) = a_m x^m + \cdots + a_0 \in \mathbb{C}[x]$ gibt es noch die <u>übliche</u> Definition von $p(A)$, nämlich:

$$p(A) := a_m A^m + \cdots + a_0 E.$$

Lemma 5.5.2 Sei $A \in \mathbf{Mat}(n, n, \mathbb{C})$ und sei $p(x) = a_m x^m + \cdots + a_0 \in \mathbb{C}[x]$ ein Polynom. Dann sind $p(A)$, berechnet nach Definition 5.5.1 und nach der üblichen Definition, gleich.

Lemma 5.5.3 Gegeben seien
- $\alpha_1, \ldots, \alpha_s \in \mathbb{C}$,
- $\ell_1, \ldots \ell_s \in \mathbb{N}$,
- $b_{ij} \in \mathbb{C}$ für $i = 1, \ldots, s$ und $j = 0, \ldots, \ell_i - 1$.

Dann existiert ein Polynom $p(x) \in \mathbb{C}[x]$ mit der Eigenschaft

$$p^{(j)}(\alpha_i) = b_{i,j}$$

für alle diese i und j. Hier ist $p^{(j)}$ die j-te Ableitung von $p(x)$.

Satz 5.5.4 Die Definition 5.5.1 (c) von $f(A)$ hängt nicht von der Wahl von T und von der Reihenfolge der Jordan-Kästschen in J ab.

5.6 Reihen von Matrizen

Der Raum $\mathbf{Mat}(n, n, \mathbb{C})$ ist ein n^2-dimensionaler normierter Raum (siehe Satz 3.3.3). Dadurch ist der Begriff *Konvergenz* für diesen Raum definiert.

Definition 5.6.1 Sei $A \in \mathbf{Mat}(n, n, \mathbb{C})$ eine Matrix. Sei $R(x) = \sum_{j=0}^{\infty} c_j x^j$ eine abstrakte Reihe mit Koeffizienten in $\mathbb{C}$ und sei $R_n(x) := \sum_{j=0}^{n} c_j x^j$ ihre n-te partielle Summe. Wenn $R_n(A)$ gegen eine Matrix konvergiert, dann bezeichnen wir diese Matrix mit $R(A)$ oder mit $\sum_{j=0}^{\infty} c_j A^j$.

Lemma 5.6.2 Sei $A \in \mathbf{Mat}(n, n, \mathbb{C})$ eine Matrix mit Minimalpolynom

$$m_A(x) = (x - \alpha_1)^{\ell_1} \cdot \ldots \cdot (x - \alpha_s)^{\ell_s}.$$

Sei I eine Teilmenge von $\mathbb{C}$ oder $\mathbb{R}$ und $f : I \to \mathbb{C}$ eine Funktion, für die die Werte

$$f(\alpha_i), f'(\alpha_i), \ldots, f^{(\ell_i - 1)}(\alpha_i)$$

für $i = 1, \ldots, s$ existieren. (Wir wissen, dass dann $f(A)$ existiert.)

Wir betrachten eine abstrakte Reihe $R(x) = \sum\limits_{j=0}^{\infty} c_j x^j$ mit Koeffizienten in $\mathbb{C}$; auch betrachten wir ihre partiellen Summen $R_n(x) = \sum\limits_{j=0}^{n} c_j x^j$. Angenommen, dass für $i = 1, \ldots, s$ folgendes gilt:

$$f(\alpha_i) = \lim_{n\to\infty} R_n(\alpha_i), \quad f'(\alpha_i) = \lim_{n\to\infty} R'_n(\alpha_i), \quad \ldots, \quad f^{(\ell_i-1)}(\alpha_i) = \lim_{n\to\infty} R_n^{(\ell_i-1)}(\alpha_i).$$

Dann gilt

$$f(A) = \sum_{j=0}^{\infty} c_j A^j.$$

Folgerung 5.6.3 Für jede Matrix $A \in \mathbf{Mat}(n, n, \mathbb{C})$ gilt

$$\exp(A) = \sum_{j=0}^{\infty} \frac{A^j}{j!}.$$

Beispiel 5.6.4 Sei

$$A = \begin{pmatrix} \pi - 1 & 1 \\ -1 & \pi + 1 \end{pmatrix}.$$

Dann ist

$$\cos(A) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sin(A) = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}.$$

Es gilt

$$(\cos(A))^2 + (\sin(A))^2 = E.$$

6 Faktorstrukturen

6.1 Faktorräume

Definition 6.1.1 Sei V ein K-Vekrorraum und W ein Untervektorraum von V. Für ein $v \in V$ definieren wir die *Klasse von v modulo W* wie folgt:

$$[v] := v + W := \{v + w \mid w \in W\}.$$

Merken wir an:

$$[v_1] = [v_2] \;\Leftrightarrow\; v_1 - v_2 \in W.$$

Der Faktorraum V/W ist die Menge

$$V/W := \{[v] \mid v \in V\}$$

mit der Addition und Skalarmultiplikation wie folgt:

$$\begin{aligned} [v] + [u] &= [v+u], \\ k \cdot [v] &= [kv], \quad k \in K, v, u \in V. \end{aligned}$$

Behauptung 6.1.2

(a) Die Addition und die Skalarmultiplikation sind auf dem Faktorraum V/W wohldefiniert.

(b) Der Faktorraum V/W ist ein K-Vektorraum.

(c) Die Abbildung $\theta : V \to V/W$, $v \mapsto [v]$ ist eine lineare Abbildung mit dem Kern W.

(d) Sei V endlich-dimensional. Wir erweitern eine Basis $b := \{b_1, \dots, b_k\}$ von W bis zu einer Basis $\mathcal{B} := \{b_1, \dots, b_n\}$ von V. Dann ist

$$\overline{\mathcal{B}} := \{[b_{k+1}], \dots [b_n]\}$$

eine Basis von V/W. Insbesondere gilt

$$\dim(V/W) := \dim(V) - \dim(W).$$

Definition 6.1.3 Sei V ein K-Vektorraum und W ein Untervektorraum von V. Sei $\varphi : V \to V$ ein Endomorphismus, für das W φ-invariant ist (also gilt $\varphi(W) \subseteq W$). Wir definieren eine Abbildung $\overline{\varphi} : V/W \to V/W$ durch

$$\overline{\varphi}\left([v]\right) := [\varphi(v)].$$

Behauptung 6.1.4

(a) Die Abbildung $\overline{\varphi} : V/W \to V/W$ ist wohldefiniert und ist ein Endomorphismus von V/W.

(b) Sei V endlichdimensional. Sei b eine Basis von W, $\mathcal{B}$ eine Basis von V und $\overline{\mathcal{B}}$ eine Basis von V/W wie in Behauptung 6.1.2. Dann gilt

$$[\varphi]_{\mathcal{B}} = \begin{pmatrix} [\varphi|_W]_b & * \\ O & [\overline{\varphi}]_{\overline{\mathcal{B}}} \end{pmatrix}.$$

6.2 Faktorringe

Definition 6.2.1 Sei R ein kommutativer Ring mit 1.

(1) Eine nicht-leere Teilmenge $I \subseteq R$ heißt Ideal, wenn folgendes gilt:

(a) $i_1, i_2 \in I$ impliziert $i_1 - i_2 \in I$,

(b) $i \in I$, $r \in R$ impliziert $i \cdot r \in I$.

(2) Ein Ideal I von R heißt *maximal*, wenn $I \neq R$ ist und kein Ideal I_1 von R mit

$$I \subsetneqq I_1 \subsetneqq R$$

existiert.

Beispiel.

(1) Ideale und maximale Ideale in $\mathbb{Z}$:

(a) Jedes Ideal $I \neq \{0\}$ in $\mathbb{Z}$ hat die Form $n\mathbb{Z}$ für ein $n \in \mathbb{N}$. Die Zahl n ist eindeutig durch I bestimmt.

(b) Ein Ideal $I \neq \{0\}$ in $\mathbb{Z}$ ist genau dann maximal, wenn I die Form $p\mathbb{Z}$ für eine Primzahl p hat.

(2) Ideale und maximale Ideale in $K[x]$, wobei K ein Körper ist:

(a) Jedes Ideal $I \neq 0$ in $K[x]$ hat die Form $f(x) \cdot K[X]$ für ein $f(x) \in K[x]$. Der Erzeuger $f(x)$ von I ist bis auf einen Faktor aus K eindeutig durch I bestimmt (siehe Lemma 5.2.2).

(b) Ein Ideal $I \neq 0$ in $K[x]$ ist genau dann maximal, wenn $I = f(x) \cdot K[x]$ für ein irreduzibles[2] Polynom $f(x)$ ist.

Definition 6.2.2 Sei R ein kommutativer Ring mit 1 und sei I ein Ideal in R. Für ein $r \in R$ definieren wir die *Klasse von r modulo I* wie folgt:

$$[r] := r + I := \{r + i \,|\, i \in I\}.$$

Merken wir an:

$$[r_1] = [r_2] \quad \Leftrightarrow \quad r_1 - r_2 \in I.$$

Der Faktorring R/I ist die Menge

$$R/I := \{[r] \,|\, r \in R\}$$

mit der Addition und Multiplikation wie folgt:

$$\begin{aligned} [a] + [b] &= [a + b], \\ [a] \cdot [b] &= [ab], \qquad a, b \in R. \end{aligned}$$

[2]Ein Polynom $f(x) \in K[x] \setminus K$ heißt *irreduzibel*, wenn keine Polynome $f_1(x), f_2(x) \in K[x]$ mit $f(x) = f_1(x) f_2(x)$ und $\mathrm{Grad}(f_i) < \mathrm{Grad}(f)$, $i = 1, 2$, existieren.

Behauptung 6.2.3 Sei R ein kommutativer Ring mit 1. Dann gilt:

(a) Die Addition und die Multiplikation sind auf dem Faktorring R/I wohldefiniert.

(b) Der Faktorring R/I ist ein Ring.

(c) Die Abbildung $\theta : R \to R/I$, $r \mapsto [r]$ ist ein Ringhomomorphismus mit dem Kern I.

(d) Die Abbildung $\theta : R \to R/I$ induziert eine Bijektion zwischen den Idealen von R, die I enthalten, und den Idealen von R/I:

$$\begin{array}{ccccc} I & \subseteq & J & \subseteq & R \\ \downarrow & & \downarrow & & \downarrow \\ \{[0]\} & \subseteq & \theta(J) & \subseteq & R/I \end{array}$$

6.3 Konstruktion von Körpern mit Hilfe von Faktorringen

Satz 6.3.1 Sei R ein kommutativer Ring mit 1. Sei I ein Ideal in R. Der Faktorring R/I ist ein Körper genau dann, wenn I ein maximales Ideal ist.

Bezeichnung. Sei K ein Körper. Wir wissen schon, dass alle Ideale in $K[x]$ Hauptideale sind, also gilt $I = f(x) \cdot K[x]$ für ein Polynom $f(x) \in K[x]$. Wir schreiben $I = \langle f(x) \rangle$.

Folgerung 6.3.2 Sei K ein Körper und $f(x) \in K[x]$. Der Faktorring $K[x]/\langle f(x) \rangle$ ist ein Körper genau dann, wenn das Polynom $f(x)$ irreduzibel ist.

Folgerung 6.3.3 Der Faktorring $\mathbb{R}[x]/\langle x^2+1 \rangle$ ist ein Körper. Dieser Körper ist isomorph zu $\mathbb{C}$.

Satz 6.3.4 Sei K ein Körper und $f(x) \in K[x]$ ein Polynom des Grades $n \geqslant 1$. Dann hat $f(x)$ höchstens n Nullstellen in K.

Lemma 6.3.5 Sei K ein Körper und $f(x) \in K[x]$ ein Polynom des Grades $n \geqslant 1$. Dann existiert ein Körper $\widetilde{K}$ mit folgenden Eigenschaften:

(1) $K \subseteq \widetilde{K}$,

(2) $f(x)$ hat eine Nullstelle in $\widetilde{K}$.

Satz 6.3.6 Sei K ein Körper und $f(x) \in K[x]$ ein Polynom des Grades $n \geqslant 1$. Dann existiert ein Körper $\widetilde{K}$ mit folgenden Eigenschaften:

(1) $K \subseteq \widetilde{K}$,

(2) Es existieren Elemente $\alpha_1, \dots, \alpha_s \in \widetilde{K}$ und natürliche Zahlen $k_1, \dots, k_s$, so dass $k_1 + \dots + k_s = n$ und

$$p(x) = (x - \alpha_1)^{k_1} \dots (x - \alpha_s)^{k_s}$$

gelten.

7 Fundamentalsatz der Algebra (1. Beweis)

In diesem Kapitel werden wir beweisen:

Satz 7.1 (Fundamentalsatz der Algebra) Jedes nichtkonstante Polynom $f(x) \in \mathbb{C}[x]$ besitzt eine Nullstelle in $\mathbb{C}$.

Dazu benötigen wir folgende Lemmata:

Lemma 7.2 (aus der Analysis) Sei D eine kompakte Teilmenge von $\mathbb{R}^n$ (d.h. abgeschlossen und beschränkt) und $f : D \to \mathbb{R}$ eine stetige Funktion. Dann nimmt f sein Infimum und Supremum auf D an.

Lemma 7.3 Sei $f(x) \in \mathbb{C}[x]$. Dann gilt $|f(z_0)| = \inf\limits_{x \in \mathbb{C}} |f(x)|$ für ein $z_0 \in \mathbb{C}$.
(Das Infimum ist also ein Minimum.)

Lemma 7.4 Sei $f(x) \in \mathbb{C}[x]$ ein nichtkonstantes Polynom. Dann gilt:

$$f(x_0) \neq 0 \quad \Rightarrow \quad |f(x_0)| \neq \min_{x \in \mathbb{C}} |f(x)|$$

8 Satz von Sturm

Behauptung A. Sei K ein Körper und $f(x) \in K[x]$. Wenn $\alpha \in K$ eine Nullstelle von $f(x)$ ist, dann gilt

$$f(x) = (x - \alpha)g(x)$$

für ein $g(x) \in K[x]$.

Definition B. Sei K ein Körper und $f(x) \in K[x]$. Wenn $\alpha \in K$ eine Nullstelle von $f(x)$ ist, dann kann man $f(x)$ in der Form

$$f(x) = (x - \alpha)^k g(x)$$

mit $g(x) \in K[x]$ und <u>maximalen</u> k aufschreiben. Diese k heißt *Vielfachheit* oder *Ordnung* von α in $f(x)$. Nach Behauptung A gilt dann $g(\alpha) \neq 0$.

Wenn die Vielfachheit von α gleich 1 ist, dann heißt α eine *einfache* Nullstelle. Wenn die Vielfachheit von α größer als 1 ist, dann heißt α eine *mehrfache* Nullstelle.

Definition 8.1

(1) Sei K ein Körper und seien $f(x), h(x) \in K[x]$ zwei Polynome. Wir sagen, dass $h(x)$ ein *Teiler* von $f(x)$ ist, wenn ein $q(x) \in K[x]$ mit der Eigenschaft $f(x) = h(x) \cdot q(x)$ existiert. In diesem Fall schreiben wir

$$h(x) \big| f(x).$$

(2) Der *größte gemeinsame Teiler* zweier Polynome $f(x), g(x) \in K[x] \setminus \{0\}$ ist das <u>normierte</u> Polynom $h(x)$ mit folgenden Eigenschaften:

(a) $h(x)$ ist ein Teiler von $f(x)$ und $g(x)$.

(b) Wenn $d(x)$ ein Teiler von $f(x)$ und $g(x)$ ist, dann ist $d(x)$ ein Teiler von $h(x)$.

Man kann zeigen, dass $h(x)$ eindeutig bestimmt ist. Wir schreiben

$$h(x) = \mathbf{ggT}(f(x), g(x)).$$

Behauptung C. Sei K ein Körper mit Charakteristik 0 (z.B. $K = \mathbb{R}$ oder $K = \mathbb{C}$). Sei $f(x) \in K[x] \setminus \{0\}$ und $\alpha \in K$ eine Nullstelle von $f(x)$. Folgende Aussagen sind äquivalent:

(1) α ist einfach.

(2) $f'(\alpha) \neq 0$.

Bemerkung 8.2 (Euklidischer Algorithmus für Polynome) Unter Benutzung der Polynomdivision mit Rest in $\mathbb{R}[x]$ lässt sich der Euklidische Algorithmus aus LA I eins zu eins übertragen. Das Ergebnispolynom ist i.A. jedoch noch nicht normiert.

Definition 8.3 Sei $f(x)$ ein reelles Polynom ohne mehrfache Nullstellen.

(1) Die *Sturmsche Kette* von $f(x)$ ist eine endliche Folge von Polynomen $p_i(x) \in \mathbb{R}[x]$, die auf folgende Weise definiert wird:

$$p_0(x) := f(x),\ p_1(x) := f'(x)$$

und $p_{i+2}(x)$ (für $i \geqslant 0$) ist implizit gegeben durch

$$p_i(x) = q_i(x)p_{i+1}(x) - p_{i+2}(x) \text{ für ein } q_i(x) \in \mathbb{R}[x],$$

wobei wir zusätzlich $\mathrm{Grad}(p_{i+2}(x)) < \mathrm{Grad}(p_{i+1})$ fordern. Die Folge endet mit einem konstanten Polynom $p_n(x)$.

(2) Sei $p_0(x), p_1(x), \ldots, p_k(x)$ die Sturmsche Kette von $f(x)$. Für ein $a \in \mathbb{R}$ bezeichnen wir mit $\sigma(a)$ die Anzahl der Vorzeichenwechsel in der Folge $p_0(a), p_1(a), \ldots, p_k(a)$. Die Nullen in der Folge werden ignoriert.

Satz 8.4 (Satz von Sturm) Sei $f(x)$ ein reelles Polynom ohne mehrfache Nullstellen. Dann ist die Anzahl der Nullstellen von $f(x)$ im halboffenen Intervall $(a, b]$ (für $a < b$) gleich $\sigma(a) - \sigma(b)$.

Satz 8.5 Sei $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ ein reelles Polynom mit $a_n \neq 0$. Dann liegen alle reellen Nullstellen von $f(x)$ in dem Intervall $[-N, N]$ mit

$$N = 1 + \frac{\max\{|a_{n-1}|, \ldots, |a_0|\}}{|a_n|}.$$

9 Symmetrische Polynome

Definition 9.1 Sei K ein Körper.

(a) Ein Polynom $f(x_1, \dots, x_n) \in K[x_1, \dots, x_n]$ heißt *symmetrisch*, wenn es invariant unter allen Permutationen von $x_1, \dots, x_n$ ist. Kurz geschrieben: wenn

$$f(x_1, \dots, x_n) = f(x_{\tau(1)}, \dots, x_{\tau(n)})$$

für alle $\tau \in S_n$ gilt.

(b) Folgende n Polynome heißen *elementar-symmetrisch* in $K[x_1, \dots, x_n]$:

$$\sigma_k(x_1, \dots, x_n) := \sum_{1 \leqslant i_1 < i_2 < \dots < i_k \leqslant i_n} x_{i_1} x_{i_2} \dots x_{i_k}, \qquad k = 1, \dots, n.$$

Etwas ausführlicher:

$$\begin{aligned}
\sigma_1(x_1, \dots, x_n) : &= x_1 + x_2 + \dots + x_n, \\
\sigma_2(x_1, \dots, x_n) : &= \sum_{1 \leqslant i < j \leqslant n} x_i x_j, \\
\vdots & \\
\sigma_n(x_1, \dots, x_n) : &= x_1 x_2 \dots x_n.
\end{aligned}$$

Satz 9.2 Sei $f(x) = x^n + a_{n-1}x^{n-1} + \dots + a_0$ ein Polynom aus $K[x]$. Seien $\alpha_1, \dots, \alpha_n$ seine Nullstellen in einem Körper $\widetilde{K}$, den K enthält (wir wiederholen jede Nullstelle so viele mal, wie ihre Vielfachheit ist). Also sei $f(x) = (x - \alpha_1) \cdot \ldots \cdot (x - \alpha_n)$. Dann gilt

$$\begin{aligned}
a_{n-1} &= -\sigma_1(\alpha_1, \dots, \alpha_n), \\
a_{n-2} &= \quad \sigma_2(\alpha_1, \dots, \alpha_n), \\
\dots & \\
a_{n-i} &= (-1)^i \sigma_i(\alpha_1, \dots, \alpha_n), \\
\dots & \\
a_0 &= (-1)^n \sigma_n(\alpha_1, \dots, \alpha_n).
\end{aligned}$$

Definition 9.3 Sei K ein Körper.

(1) Ein Ausdruck $ax_1^{k_1} x_2^{k_2} \dots x_n^{k_n}$ mit $a \in K \setminus \{0\}$ und $k_1, k_2, \dots, k_n \in \mathbb{N} \cup \{\infty\}$ heißt *Monom*. Es ist klar, dass jedes Polynom in $K[x_1, x_2, \dots, x_n] \setminus \{0\}$ eine Summe von Monomen ist.

(2) Wir schreiben

$$ax_1^{k_1} x_2^{k_2} \dots x_n^{k_n} \succ bx_1^{k'_1} x_2^{k'_2} \dots x_n^{k'_n}$$

genau dann, wenn ein $i \in \mathbb{N}$ mit $1 \leqslant i \leqslant n$ existiert, so dass gilt:

$$\begin{aligned}
k_1 &= k'_1, \\
\dots & \\
k_{i-1} &= k'_{i-1}, \\
k_i &> k'_i.
\end{aligned}$$

Wir schreiben

$$ax_1^{k_1}x_2^{k_2}\dots x_n^{k_n} \succcurlyeq bx_1^{k'_1}x_2^{k'_2}\dots x_n^{k'_n}$$

genau dann, wenn entweder

$$ax_1^{k_1}x_2^{k_2}\dots x_n^{k_n} \succ bx_1^{k'_1}x_2^{k'_2}\dots x_n^{k'_n},$$

oder

$$x_1^{k_1}x_2^{k_2}\dots x_n^{k_n} = x_1^{k'_1}x_2^{k'_2}\dots x_n^{k'_n}$$

ist.

(3) Jedes Polynom $f \in K[x_1,\dots,x_n] \setminus \{0\}$ kann in der Form

$$f = f_1 + f_2 + \dots + f_m$$

geschrieben werden, wobei $f_1, f_2, \dots, f_m$ Monome mit $f_1 \succ f_2 \succ \dots \succ f_m$ sind. Das Monom f_1 heißt *Hauptmonom* von f.

Satz 9.4 Für $f, g \in K[x_1,\dots,x_n]$ ist das Hauptmonom von fg gleich dem Produkt von Hauptmonomen von f und g.

Satz 9.5 Sei $f \in K[x_1,\dots,x_n]$ ein symmetrisches Polynom. Dann existiert ein Polynom $F \in K[x_1,\dots,x_n]$, so dass gilt:

$$f(x_1,\dots,x_n) = F\big(\sigma_1(x_1,\dots,x_n),\dots,\sigma_n(x_1,\dots,x_n)\big).$$

10 Hauptsatz der Algebra (2. Beweis)

Satz 10.1 (Hauptsatz der Algebra) Jedes Polynom $f(x) \in \mathbb{C}[x] \setminus \mathbb{C}$ hat eine Nullstelle in $\mathbb{C}$.

Beweis. Merken wir an: $f(x)\overline{f(x)} \in \mathbb{R}[x]$. Es ist also genügend zu zeigen, dass das relle Polynom $f(x)\overline{f(x)}$ eine Nullstelle in $\mathbb{C}$ hat.

Behauptung. Jedes Polynom $h(x) \in \mathbb{R}[x] \setminus \mathbb{R}$ hat eine Nullstelle in $\mathbb{C}$.

Beweis. Sei n der Grad von $h(x)$. Wir schreiben $n = 2^m \cdot \ell$ mit $m \in \mathbb{N} \cup \{0\}$ und ungeradem ℓ und führen die Induktion nach m.

IA. Sei $m = 0$. Dann ist der Grad von $h(x)$ ungerade. Nach dem Nullstellensatz von Bolzano hat $h(x)$ eine Nullstelle sogar in $\mathbb{R}$.

IS: $m \to m+1$**.** Sei $h(x) \in \mathbb{R}[x]$ ein Polynom mit Grad $n = 2^{m+1}\ell$.

Nach Satz 6.3.5 existiert ein Körper $\widetilde{K}$ mit folgenden Eigenschaften:

(1) $\mathbb{C} \subseteq \widetilde{K}$,

(2) es existieren $\alpha_1,\dots,\alpha_s \in \widetilde{K}$, so dass $h(x) = (x-\alpha_1)\dots(x-\alpha_n)$ gilt.

Wir nehmen eine beliebige Zahl $c \in \mathbb{R}$ und betrachten das Polynom

$$H_c(x) := \prod_{\substack{\{i,j\} \subseteq \{1,\dots,n\} \\ i \neq j}} (x - (\alpha_i + \alpha_j + c\alpha_i\alpha_j)).$$

Merken wir an:

(a) Der Grad von $H_c(x)$ ist $\frac{n\cdot(n-1)}{2} = 2^m \cdot \ell'$ mit $\ell' = \ell \cdot (2^{m+1}\ell - 1)$.

(b) $H_c(x) \in \mathbb{R}[x]$. Das folgt aus den folgenten Fakten:

- Die Koeffizienten von $H_c(x)$ sind symmetrische Polynome von $\alpha_1, \dots, \alpha_n$ über $\mathbb{R}$.
- Nach Satz 9.5 können diese symmetrischen Polynome als Polynome über $\mathbb{R}$ von elementar-symmetrischen Polynomen $\sigma_1(\alpha_1, \dots, \alpha_n), \dots, \sigma_1(\alpha_1, \dots, \alpha_n)$ dargestellt werden.
- Nach Satz 9.2 sind die Zahlen $\sigma_1(\alpha_1, \dots, \alpha_n), \dots, \sigma_1(\alpha_1, \dots, \alpha_n)$ bis auf Vorzeichen gleich den Koeffizienten von $f(x)$, also liegen in $\mathbb{R}$.

Nach Induktionsvoraussetzung hat $H_c(x)$ eine Nullstelle in $\mathbb{C}$. Also existieren i, j mit $\alpha_i + \alpha_j + c\alpha_i\alpha_j \in \mathbb{C}$. Diese i, j hängen von c ab.

Jetzt variieren wir $c \in \mathbb{R}$. Dann existieren zwei verschiedene Zahlen $c, \widetilde{c} \in \mathbb{R}$, für die das gleiche Paar (i, j) mit der Eigenschaft

$$\begin{aligned} \alpha_i + \alpha_j + c\alpha_i\alpha_j &\in \mathbb{C}, \\ \alpha_i + \alpha_j + \widetilde{c}\alpha_i\alpha_j &\in \mathbb{C} \end{aligned}$$

auftritt. Daraus folgt $\alpha_i + \alpha_j \in \mathbb{C}$, $\alpha_i\alpha_j \in \mathbb{C}$ und schließlich $\alpha_i, \alpha_j \in \mathbb{C}$. □ □

11 Resultante und Diskriminante

Satz 11.1 Sei K ein Körper und seien

$$\begin{aligned} f(x) &= a_k x^k + a_{k-1}x^{k-1} + \cdots + a_0, \\ g(x) &= b_\ell x^\ell + b_{\ell-1}x^{\ell-1} + \cdots + b_0. \end{aligned}$$

zwei Polynome aus $K[x]$. Wir erlauben auch $a_k = 0$ und $b_\ell = 0$.

(1) Die $(k+\ell) \times (k+\ell)$-Matrix

$$\mathrm{Syl}(f,g) := \begin{pmatrix} a_k & a_{k-1} & \dots & a_0 & & & \\ & a_k & a_{k-1} & \dots & a_0 & & \\ & & \ddots & & & \ddots & \\ & & & a_k & a_{k-1} & \dots & a_0 \\ b_\ell & b_{\ell-1} & \dots & b_0 & & & \\ & b_\ell & b_{\ell-1} & \dots & b_0 & & \\ & & \ddots & & & \ddots & \\ & & & b_\ell & b_{\ell-1} & \dots & b_0 \end{pmatrix}$$

heißt *Sylvester-Matrix* von f und g. Diese Matrix besteht aus ℓ Zeilen mit den Koeffizienten von f und k Zeilen mit den Koeffizienten von g. Alle in der obigen Matrix nicht beschrifteten Einträge sind Null.

(2) Die Zahl

$$\mathbf{Res}(f,g) := \det\bigl(\mathrm{Syl}(f,g)\bigr)$$

heißt *Resultant* von f und g.

Lemma. (Vandermonde-Determinante). Für $n \geqslant 2$ gilt

$$\det \begin{pmatrix} z_1^{n-1} & z_2^{n-1} & \dots & z_n^{n-1} \\ z_1^{n-2} & z_2^{n-2} & \dots & z_n^{n-2} \\ \dots & \dots & \dots & \dots \\ z_1 & z_2 & \dots & z_n \\ 1 & 1 & \dots & 1 \end{pmatrix} = \prod_{1 \leqslant i < j \leqslant n} (z_i - z_j).$$

Satz 11.2 Sei K ein Körper und seien

$$\begin{aligned} f(x) &= a_k x^k + a_{k-1} x^{k-1} + \dots + a_0, \\ g(x) &= b_\ell x^\ell + b_{\ell-1} x^{\ell-1} + \dots + b_0. \end{aligned}$$

zwei Polynome aus $K[x]$ mit $a_k \neq 0$ und $b_\ell \neq 0$. Dann (wie wir wissen) existiert eine Körpererweiterung $\widetilde{K}$ und Elemente $\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_\ell \in K$, so dass gilt:

$$\begin{aligned} f(x) &= a_k (x-\alpha_1)(x-\alpha_2)\dots(x-\alpha_k), \\ g(x) &= b_\ell (x-\beta_1)(x-\beta_2)\dots(x-\beta_\ell). \end{aligned}$$

Dann gilt

$$\mathbf{Res}(f,g) = a_k^\ell b_\ell^k \prod_{\substack{1 \leqslant i \leqslant k \\ 1 \leqslant j \leqslant \ell}} (\alpha_i - \beta_j).$$

Folgerung 11.3 Mit Voraussetzungen von Satz 11.2 gilt

$$\mathbf{Res}(f,g) = a_k^\ell \prod_{i=1}^{k} g(\alpha_i) = b_\ell^k \prod_{j=1}^{\ell} f(\beta_j).$$

Folgerung 11.4 Sei K ein Körper und seien

$$\begin{aligned} f(x) &= a_k x^k + a_{k-1} x^{k-1} + \cdots + a_0, \\ g(x) &= b_\ell x^\ell + b_{\ell-1} x^{\ell-1} + \cdots + b_0 \end{aligned}$$

zwei Polynome aus $K[x]$. Dann sind folgende Behauptungen äquivalent:

(1) $\mathbf{Res}(f, g) = 0$,

(2) f und g haben eine gemeinsame Nullstelle in einer Körpererweiterung $\widetilde{K}$ oder $a_k = b_\ell = 0$.

Algorithmus 11.5 (Eliminationsverfahren für Systeme von Polynomialgleichungen mit mehreren Unbekannten)
Gegeben sei ein System von zwei Polynomialgleichungen in zwei Unbekannten:

$$\begin{cases} f(x,y) := & a_k(y) x^k + a_{k-1}(y) x^{k-1} + \cdots + a_0(y) = 0, \\ g(x,y) := & b_\ell(y) x^\ell + b_{\ell-1}(y) x^{\ell-1} + \cdots + b_0(y) = 0. \end{cases} \tag{11.1}$$

Wir zeigen, dass alle Lösungen dieses Systems gefunden werden können, wenn man alle Lösungen polynomialer Gleichungen von einer Unbekannten finden kann.

Wir betrachten $f(x,y)$ und $g(x,y)$ als Polynome von x mit Koeffizienten $a_i(y)$ und $b_j(y)$. Nach Definition 11.1 ist die Resultante von $f(x,y)$ und $g(x,y)$ ein Polynom von diesen Koeffizienten. Also ist diese Resultante ein Polynom von y:

$$P(y) := \mathbf{Res}_x\big(f(x,y), g(x,y)\big).$$

Sei (α, β) eine Lösung des obigen Systems. Dann haben die Polynome

$$\begin{aligned} f(x,\beta) : &= a_k(\beta) x^k + a_{k-1}(\beta) x^{k-1} + \cdots + a_0(\beta), \\ g(x,\beta) : &= b_\ell(\beta) x^\ell + b_{\ell-1}(\beta) x^{\ell-1} + \cdots + b_0(\beta) \end{aligned}$$

eine gemeinsame Nullstelle, nämlich α. Dann gilt (nach Folgerung 11.4):

$$\mathbf{Res}\big(f(x,\beta), g(x,\beta)\big) = 0 \quad \text{oder} \quad a_k(\beta) = b_\ell(\beta) = 0.$$

Also erfüllt β die Gleichung $P(y) = 0$ oder das System

$$\begin{cases} a_k(y) &= 0, \\ b_\ell(y) &= 0. \end{cases}$$

Angenommen dass alle Nullstellen von Polynomialgleichungen in einer Unbekannten gefunden werden können. Dann können wir alle möglichen β finden. Danach substituieren wir jedes gefundene β in das System (11.1) und finden die zugehörige α. □

Definition 11.6 Sei K ein Körper und $f(x) = a_k x^k + a_{k-1}x^{k-1} + \cdots + a_0$ ein Polynom aus $K[x]$ mit $a_k \neq 0$. Seien $\alpha_1, \dots, \alpha_k$ seine Nullstellen in einer Körpererweiterung $\widetilde{K}$. Dann heißt das Element

$$\mathbf{Dis}(f) := a_k^{2k-2} \prod_{1 \leqslant i < j \leqslant k} (\alpha_i - \alpha_j)^2$$

Diskriminante von f.

Bemerkung. Es gilt $\mathbf{Dis}(f) \in K$.
Tatsächlich ist $\mathbf{Dis}(f)$ ein symmetrisches Polynom über $\mathbb{Z}$ von $\alpha_1, \dots, \alpha_k$, also ein Polynom über $\mathbb{Z}$ von elementar-symmetrischen Polynomen $\sigma_1(\alpha_1, \dots, \alpha_k), \dots, \sigma_k(\alpha_1, \dots, \alpha_k)$. Diese sind aber den Koeffizienten $a_{k-1}, \dots, a_0$ bis auf Vorzeichen gleich.

Satz 11.7 Sei K ein Körper. Für das Polynom $f(x) = a_k x^k + a_{k-1}x^{k-1} + \cdots + a_0$ mit $a_k \neq 0$ und $k \neq 0$ in K gilt

$$\mathbf{Dis}(f) = (-1)^{\frac{k(k-1)}{2}} \frac{1}{a_k} \mathbf{Res}(f, f').$$

wobei f' die Ableitung von f ist.

12 Tensoren

12.1 Tensorprodukt von Vektorräumen

Definition 12.1.1 Seien U, V, W drei Vektorräume über einem Körper K. Eine Abbildung $f : U \times V \to W$ heißt bilinear, wenn für alle $u, u_1 \in U$, $v, v_1 \in V$ und $\lambda \in K$ gelten:

(1) $f(u + u_1, v) = f(u, v) + f(u_1, v)$,

(2) $f(u, v + v_1) = f(u, v) + f(u, v_1)$,

(3) $\lambda \cdot f(u, v) = f(\lambda u, v) = f(u, \lambda v)$.

Bemerkung 12.1.2 Seien U, V, W, W' vier Vektorräume über einem Körper K. Sei $f : U \times V \to W$ eine bilineare Abbildung und $\psi : W \to W'$ eine lineare Abbildung. Dann ist die Abbildung $\psi \circ f : U \times V \to W'$ bilinear.

Satz 12.1.3 Seien U, V zwei Vektorräume über einem Körper K.

(1) Es existiert ein K-Vektorraum W und eine bilineare Abbildung $f : U \times V \to W$ mit folgender Eigenschaft:

Zu jedem K-Vektorraum X und zu jeder bilinearen Abbildung $\alpha : U \times V \to X$ gibt es genau eine lineare Abbildung $\varphi_\alpha : W \to X$, so dass das folgende Diagramm kommutativ ist:

$$\begin{tikzcd} U \times V \arrow[r, "f"] \arrow[dr, "\alpha"'] & W \arrow[d, dashed, "\varphi_\alpha"] \\ & X \end{tikzcd}$$

(2) Die bilineare Abbildung $f : U \times V \to W$ ist eindeutig im folgenden Sinne:

Für je zwei solche bilineare Abbildungen $f_i : U \times V \to W_i$, $i = 1, 2$ existiert ein Isomorphismus $\psi : W_1 \to W_2$, für den $f_2 = \psi \circ f_1$ gilt.

Beweis. (1) Wir definieren W als Faktorvektorraum $\mathcal{F}/\mathcal{F}_1$, wobei $\mathcal{F}$ und $\mathcal{F}_1$ wie folgt definiert werden:

- Sei $\mathcal{F}$ der K-Vektorraum mit der Basis

$$\mathcal{B} := \{(u \otimes v) \mid u \in U, v \in V\}.$$

Also sind Elemente von $\mathcal{F}$ alle endliche lineare Kombinationen von Elementen aus $\mathcal{B}$. Es ist klar, dass $\dim(\mathcal{F}) = |U| \cdot |V|$ ist.

- Sei $\mathcal{F}_1$ der Untervektorraum von $\mathcal{F}$, der von folgenden Elementen erzeugt ist (wobei $u, u_1 \in U$, $v, v_1 \in V$ und $\lambda \in K$ sind):

$$\begin{array}{l} (u+u_1)\otimes v - u\otimes v - u_1\otimes v, \\ u\otimes(v+v_1) - u\otimes v - u\otimes v_1, \\ \lambda(u\otimes v) - (\lambda u)\otimes v, \\ \lambda(u\otimes v) - u\otimes(\lambda v). \end{array}.$$

- Jetzt setzen wir $W := \mathcal{F}/\mathcal{F}_1$. Also ist

$$W := \{ \sum_{u\in U, v\in V} \lambda_{u,v}(u\otimes v) + \mathcal{F}_1 \,|\, \lambda_{u,v} \in K \text{ und nur endlich viele } \lambda_{u,v} \neq 0\}.$$

<u>**Verabredung.**</u> Einfachheitshalber werden die Elemente von W einfach als Summen

$$\sum_{u\in U, v\in V} \lambda_{u,v}(u\otimes v)$$

mit endlich vielen $\lambda_{u,v} \neq 0$ geschrieben. Dabei werden zwei solche Summen als gleiche in W betrachtet, wenn ihre Differenz in $\mathcal{F}_1$ liegt. Insbesondere werden die Elemente aus der linken und der rechten Seite folgender Tabelle gleich in W:

$(u+u_1)\otimes v$	$u\otimes v + u_1\otimes v$
$u\otimes(v+v_1)$	$u\otimes v + u\otimes v_1$
$\lambda(u\otimes v)$	$(\lambda u)\otimes v$
$\lambda(u\otimes v)$	$u\otimes(\lambda v)$

Merken wir an, dass jedes Element von W als eine endliche Summe von Summanden der Sorte $u\otimes v$ (mit $u\in U, v\in V$) geschrieben werden kann.

- Jetzt definieren wir $f : U\times V \to W$ nach der Regel

$$f\big((u,v)\big) := u\otimes v.$$

Diese Abbildung ist bilinear und erfüllt (1): Für eine bilineare Abbildung $\alpha : U\times V \to X$ definieren wir eine Abbildung $\varphi_\alpha : W \to X$ nach der Regel

$$\varphi_\alpha\Big(\sum_{u\in U, v\in V} \lambda_{u,v}(u\otimes v)\Big) := \sum_{u\in U, v\in V} \lambda_{u,v}\cdot\alpha\big((u,v)\big)$$

Dann ist φ_α wohldefiniert, linear und es gilt $\alpha = \varphi_\alpha \circ f$.

(2) kann formal bewiesen werden. □

Definition 12.1.4 In der Situation des Satzes 12.1.3 (1) heißt der Vektorraum W das *Tensorprodukt* von U und V. Bezeichnung: $U\underset{K}{\otimes}V$ oder $U\otimes V$, wenn es kein Missverständnis über K gibt.

Satz 12.1.5 Seien U und V zwei K-Vektorräume mit Basen $(u_i)_{i\in I}$ und $(v_j)_{j\in J}$ entsprechend. Dann ist $\{u_i \otimes v_j \mid i \in I, j \in J\}$ eine Basis von $U \otimes V$. Insbesondere gilt[3]

$$\dim_K(U \otimes V) = \dim_K(U) \cdot \dim_K(V).$$

Beispiel 12.1.6 Sei V ein $\mathbb{R}$-Vektorraum mit Basis $\mathcal{B}$. Wir betrachten $\mathbb{C}$ als $\mathbb{R}$-Vektorraum mit der Basis $\{1, i\}$. Dann ist $\mathbb{C} \underset{R}{\otimes} V$ ein $\mathbb{R}$-Vektorraum mit der Basis

$$\mathcal{A} = \{(1 \otimes b) \mid b \in \mathcal{B}\} \cup \{(i \otimes b) \mid b \in \mathcal{B}\}.$$

Insbesondere gilt

$$\dim_{\mathbb{R}}\big(\mathbb{C} \underset{\mathbb{R}}{\otimes} V\big) = 2 \cdot \dim_{\mathbb{R}}(V).$$

(1) Sei $\mathcal{B} = (b_1, b_2, b_3)$. Das Tensorprodukt $\mathbb{C} \underset{\mathbb{R}}{\otimes} V$ besitzt das Element

$$5((2+3i) \otimes (2b_1 - b_2) + 6((1-i) \otimes (b_2 - b_3)).$$

Fir finden seine Zerlegung in der Basis $\mathcal{A}$:

$$\begin{aligned}
&5((2+3i) \otimes (2b_1 - b_2) + 6((1-i) \otimes (b_2 - b_3)) = \\
&(10 \cdot 1 + 15i) \otimes (2b_1 - b_2)) + (6 \cdot 1 - 6i) \otimes (b_2 - b_3)) = \\
&20(1 \otimes b_1) - 10(1 \otimes b_2) + 30(i \otimes b_1) - 15(i \otimes b_2) + 6(1 \otimes b_2) - 6(1 \otimes b_3) - 6(i \otimes b_2) + 6(i \oplus b_3) = \\
&20(1 \otimes b_1) - 4(1 \otimes b_2) - 6(1 \otimes b_3) + 30(i \otimes b_1) - 21(i \otimes b_2) + 6(i \oplus b_3).
\end{aligned}$$

Wir können weiter schreiben:

$$= (1 \otimes (20b_1 - 4b_2 - 6b_3)) + (i \otimes (30b_1 - 21b_2 + 6b_3).$$

(2) In dem Vektorraum $\mathbb{C} \underset{\mathbb{R}}{\otimes} V$ gibt es zwei Untervektorräume

$$\{(1 \otimes v) \mid v \in V\} \text{ und } \{(i \otimes v) \mid v \in V\}\}.$$

Wir bezeichnen den ersten Raum wieder mit V und den zweiten mit iV. Dann gilt

$$\mathbb{C} \underset{\mathbb{R}}{\otimes} V = V \oplus iV.$$

(3) Wir können $\mathbb{C} \underset{\mathbb{R}}{\otimes} V$ als $\mathbb{C}$-Vektoraum betrachten, wenn wir die Skalarmultiplikation von $c \in \mathbb{C}$ mit einem Tensor so definieren:

$$c \cdot \Big(\sum_{u\in U, v\in V} \lambda_{u,v}(u \otimes v)\Big) := \sum_{u\in U, v\in V} \lambda_{u,v} \cdot (cu \otimes v)\big).$$

Der $\mathbb{C}$-Vektorraum $\mathbb{C} \underset{\mathbb{R}}{\otimes} V$ heißt *Komplexifizierung* des $\mathbb{R}$-Vektorraums V.

Eine Basis von $\mathbb{C} \underset{\mathbb{R}}{\otimes} V$ über $\mathbb{C}$ ist $\{(1 \otimes b) \mid b \in \mathcal{B}\}$. Insbesondere gilt

$$\dim_{\mathbb{C}}(\mathbb{C} \underset{\mathbb{R}}{\otimes} V) = \dim_{\mathbb{R}}(V).$$

[3]Zum Vergleich: Wenn U und V Unterräume eines Vektorraumes sind mit $U \cap V = \{0\}$, dann gilt $\dim_K(U \oplus V) = \dim_K(U) + \dim_K(V)$.

(4) Für jeden $c \in C$ betrachten wir die Abbildung

$$\begin{aligned} \Psi_c : \quad \mathbb{C} \underset{\mathbb{R}}{\otimes} V &\to \mathbb{C} \underset{\mathbb{R}}{\otimes} V, \\ w &\mapsto cw, \end{aligned}$$

wobei das Produkt cw in (3) definiert wurde.

(a) Zuerst betrachten wir $\mathbb{C} \underset{\mathbb{R}}{\otimes} V$ als $\mathbb{R}$-Vektorraum. Dann ist Ψ_c $\mathbb{R}$-linear. Sei $n = \dim(V)$. Die Darstellungsmatrix von $\Psi_{\mathbf{i}}$ in der Basis

$$\mathcal{A} = \{(1 \otimes b) \,|\, b \in \mathcal{B}\} \cup \{(i \otimes b) \,|\, b \in \mathcal{B}\}$$

ist

$$\begin{pmatrix} O_n & -E_n \\ E_n & O_n \end{pmatrix},$$

wobei O_n die $(n \times n)$-Nullmatrix und E_n die $(n \times n)$-Einheitsmatrix sind.

(b) Jetzt betrachten wir $\mathbb{C} \underset{\mathbb{R}}{\otimes} V$ als $\mathbb{C}$-Vektorraum. Dann ist Ψ_c $\mathbb{C}$-linear. Sei $n = \dim(V)$. Die Darstellungsmatrix von Ψ_c in der Basis $\{(1 \otimes b) \,|\, b \in \mathcal{B}\}$ ist

$$cE_n.$$

Lemma 12.1.7 Sei V ein K-Vektorraum mit einer Basis $\mathcal{B}$, sei $\Phi : V \to V$ eine K-lineare Abbildung und sei L ein Körper, der K enthält. Dann gilt:

(1) Das Tensorprodukt $L \underset{K}{\otimes} V$ wird zu einem L-Vektorraum, wenn wir die Skalarmultiplikation von $\lambda \in L$ mit einem Tensor so definieren:

$$\lambda \cdot \Big(\sum_{\ell \in L, v \in V} \lambda_{\ell,v} (\ell \otimes v) \Big) := \sum_{\ell \in U, v \in V} \lambda_{\ell,v} \cdot (\lambda\ell \otimes v) \Big).$$

(2) Der L-Vektorraum $L \underset{K}{\otimes} V$ hat die Basis

$$1 \otimes \mathcal{B} := \{(1 \otimes b) \,|\, b \in \mathcal{B}\}.$$

(3) Wir betrachten die Abbildung

$$\begin{aligned} \widetilde{\Phi} : L \underset{K}{\otimes} V &\to L \underset{K}{\otimes} V, \\ \sum_{\ell \in L, v \in V} \lambda_{\ell,v} (\ell \otimes v) &\mapsto \sum_{\ell \in L, v \in V} \lambda_{\ell,v} (\ell \otimes \Phi(v)). \end{aligned}$$

Dann ist diese Abbildung L-linear und für die Darstellungsmatrizen der Abbildungen Φ und $\widetilde{\Phi}$ gilt:

$$[\widetilde{\Phi}]_{1 \otimes \mathcal{B}} = [\Phi]_{\mathcal{B}}.$$

Satz 12.1.8 Seien U, V, W drei K-Vektorräume. Dann gilt:

(1) $(U \oplus V) \otimes W = (U \otimes W) \oplus (V \otimes W)$,

(2) $V \otimes W \cong W \otimes V$,

(3) $K \otimes V \cong V$,

(4) $(U \otimes V) \otimes W \cong U \otimes (V \otimes W)$.

12.2 Dualraum

Definition 12.2.1 Sei V ein K-Vektorraum. Jede lineare Abbildung $f : V \to K$ heißt *Funktional.* Für je zwei Funktional $f : V \to K$ und $f_1 : V \to K$ wird ihre Summe $(f + f_1) : V \to K$ üblicherweise definiert:

$$(f + f_1)(v) := f(v) + f_1(v), \;\; v \in V.$$

Für ein Element $\lambda \in K$ und ein Funktional $f : V \to K$ wird ihr Produkt $\lambda f : V \to K$ so definiert:

$$(\lambda f)(v) := \lambda \cdot f(v), \;\; v \in V.$$

Bezüglich dieser Addition Skalarmultiplikation wird die Menge

$$V^* := \{f : V \to K \,|\, f \text{ ist linear}\}$$

zu einem K-Vektorraum. Dieser heißt *Dualraum* zu V.

Satz 12.2.2 Sei V ein K-Vektorraum mit Basis $\mathcal{B} := (v_i)_{i\in I}$. Für $i \in I$ definieren wir ein Funktional $v_i^* : V \to K$ nach der Regel

$$v_i^*(v_j) = \delta_{ij}, \;\; j \in I.$$

(1) Ist I endlich, dann ist $\mathcal{B}^* := \{v_i^* \,|\, i \in I\}$ eine Basis des Dualraums V^*. In diesem Fall gilt $\dim(V^*) = \dim(V)$.

(Die Basis $\mathcal{B}^*$ heißt *Dualbasis* zu $\mathcal{B}$.)

(2) Ist I unendlich, dann ist $\mathcal{B}^* := \{v_i^* \,|\, i \in I\}$ keine Basis von V^*.

Bemerkung. Ist $|K| \leqslant |I| = \infty$, dann gilt sogar $\dim(V^*) > \dim(V)$.

Definition 12.2.3 Seien V und W zwei K-Vektorräume. Wir definieren eine Abbildung:

$$\psi : V^* \otimes W^* \to (V \otimes W)^*$$

zuerst auf elementaren Tensoren: $f \otimes g$ (mit $f \in V^*$ und $g \in W^*$) durch die Formel

$$f \otimes g \mapsto \big(v \otimes w \mapsto f(v) \cdot g(w)\big),$$

danach setzen wir diese Abbildung linear auf $V^* \otimes W^*$ fort.

Satz 12.2.4 Sind V und W endlichdimensional, dann ist $\psi : V^* \otimes W^* \to (V \otimes W)^*$ ein Isomorphismus.

Definition 12.2.5 Seien $A \in \mathbf{Mat}(n, m, K)$ und $B \in \mathbf{Mat}(n', m', K)$ zwei Matrizen. Wir definieren eine Matrix $A \otimes B \in \mathbf{Mat}(nn', mm', K)$ wie folgt:

$$A \otimes B = \begin{pmatrix} a_{11}B & \dots & a_{11}B \\ \vdots & & \vdots \\ a_{n1}B & \dots & a_{nm}B \end{pmatrix}.$$

Diese Matrix heißt *Kroneker-Produkt* (*oder Tensorprodukt*) von A und B.

Definition 12.2.6 Seien $\alpha : V \to V'$ und $\beta : W \to W'$ zwei lineare Abbildungen. Wir definieren eine Abbildung

$$\alpha \otimes \beta : V \otimes W \to V' \otimes W'$$

zuerst auf elementaren Tensoren: $v \otimes w$ (mit $v \in V$ und $w \in W$) durch die Formel

$$(\alpha \otimes \beta)(u \otimes w) = (\alpha(u) \otimes \beta(w)),$$

danach setzen wir diese Abbildung linear auf $V \otimes W$ fort.

Bezeichnung. Seien V und W zwei endlichdimensionale K-Vektorräume mit Basen $e = (e_1, e_2, \dots, e_n)$ und $f = (f_1, \dots, f_m)$. Die Basis

$$\begin{array}{llll} e_1 \otimes f_1, & e_1 \otimes f_2, & \dots, & e_1 \otimes f_m, \\ e_2 \otimes f_1, & e_2 \otimes f_2, & \dots, & e_2 \otimes f_m, \\ \dots & & & \\ e_n \otimes f_1, & e_n \otimes f_2, & \dots, & e_n \otimes f_m, \end{array}$$

von $V \otimes W$ wird mit $e \otimes f$ bezeichnet.

Satz 12.2.7 Seien V, V', W, W' vier endlichdimensionaler K-Vektorräume mit Basen e, e', f, f'. Seien $\alpha : V \to V'$ und $\beta : W \to W'$ zwei lineare Abbildungen. Wir betrachten die Abbildung

$$\alpha \otimes \beta : V \otimes W \to V' \otimes W'.$$

Dann sind die Darstellungsmatrizen von α, β und $\alpha \otimes \beta$ bezüglich bestimmten Basen wie folgt verbunden:

$$[\alpha \otimes \beta]_{e \otimes f}^{e' \otimes f'} = [\alpha]_e^{e'} \otimes [\beta]_f^{f'}.$$

(In der rechten Seite der Formel steht das Kroneker-Produkt von zwei Matrizen.)

13 Äußere Formen

13.1 k-Formen

Definition 13.1.1 Sei V ein K-Vektorraum.
Eine *k-Form* (oder *Multilinearform* vom Grad k) auf V ist eine multilineare Abbildung

$$\alpha : \underbrace{V \times \cdots \times V}_{k} \to K.$$

Das bedeutet, dass α die folgenden Eigenschaften für alle $i = 1, \dots, k$ erfüllt:

(1)
$$\begin{aligned} \alpha(v_1, \dots, v_{i-1}, v_i' + v_i'', v_{i+1}, \dots, v_k) &= \alpha(v_1, \dots, v_{i-1}, v_i', v_{i+1}, \dots, v_k) \\ &\quad + \alpha(v_1, \dots, v_{i-1}, v_i'', v_{i+1}, \dots, v_k), \end{aligned}$$

(2) $\alpha(v_1, \dots, v_{i-1}, \lambda v_i, v_{i+1}, \dots, v_k) = \lambda \cdot \alpha(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_k), \quad \lambda \in K.$

Bezeichnung. Mit $Q^k(V)$ bezeichnen wir den K-Vektorraum aller k-Formen auf V.

Bemerkung. Wie in Satz 12.1.3 existiert für jede k-Form

$$\alpha : \underbrace{V \times \cdots \times V}_{k} \to K$$

eine eindeutige lineare Abbildung

$$\varphi_\alpha : \underbrace{V \otimes \cdots \otimes V}_{k} \to K,$$

so dass das folgende Diagramm kommutativ ist:

$$\begin{tikzcd}
\underbrace{V \times \cdots \times V}_{k} \arrow[r, "f"] \arrow[dr, "\alpha"'] & \underbrace{V \otimes \cdots \otimes V}_{k} \arrow[d, dashed, "\varphi_\alpha"] \\
& K
\end{tikzcd}$$

Dabei ist f eine k-lineare Abbildung, die durch $f(v_1, \ldots, v_k) = v_1 \otimes \cdots \otimes v_k$ gegeben ist.

Satz 13.1.2 Die Abbildung

$$\begin{aligned} b_k : Q^k(V) &\to (\underbrace{V \otimes \cdots \otimes V}_{k})^* \\ \alpha &\mapsto \varphi_\alpha \end{aligned}$$

ist ein Isomorphismus von Vektorräumen. Insbesondere gilt

$$Q^k(V) \cong \underbrace{V^* \otimes \cdots \otimes V^*}_{k}.$$

Definition 13.1.3 Wir definieren die Multiplikation von k-und ℓ-Formen:

$$\begin{aligned} Q^k(V) \times Q^\ell(V) &\to Q^{k+\ell}(V), \\ (\alpha, \beta) &\mapsto \alpha \bullet \beta, \end{aligned}$$

wobei

$$(\alpha \bullet \beta)(v_1, \ldots, v_k, v_{k+1}, \ldots, v_{k+\ell}) := \alpha(v_1, \ldots, v_k) \cdot \beta(v_{k+1}, \ldots, v_{k+\ell})$$

ist.

Definition 13.1.4 Eine Gruppe G operiert auf einer Menge M, wenn zu jedem Paar $(g, m) \in G \times M$ ein Element aus M zugeordnet ist (dieses wird mit gm bezeichnet) und für alle $g_1, g_2 \in G$ und $m \in M$ gilt:

(1) $em = m$, (e ist das neutrale Element von G)

(2) $(g_1g_2)m = g_1(g_2m)$.

Beispiele.

(i) S_n operiert auf $M := \{1, 2, \ldots, n\}$ durch $\sigma i := \sigma(i)$.

(ii) $\mathrm{GL}_n(K)$ operiert auf K^n durch $Az := A \cdot z$.

(iii) $\mathrm{GL}_2(\mathbb{R})$ operiert auf dem Poincaré-Modell der hyperbolischen Ebene

$$\mathbb{H} := \{z \in \mathbb{C} \,|\, \mathrm{Im}(z) > 0\}$$

durch

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} z = \frac{a \cdot z + b}{c \cdot z + d}.$$

Satz 13.1.5 Für eine Permutation $\sigma \in S_n$ und eine n-Form $\alpha \in Q^n(V)$ definieren wir eine n-Form $\sigma\alpha \in Q^n(V)$ wie folgt:

$$(\sigma\alpha)(v_1, \ldots, v_n) := \mathrm{sign}(\sigma) \cdot \alpha(v_{\sigma(1)}, \ldots, v_{\sigma(n)}).$$

Diese Definition gibt eine Operierung von S_n auf $Q^n(V)$.

Definition 13.1.6 Eine n-Form $\alpha \in Q^n(V)$ heißt *alternierend* (oder *äußere* n-Form, oder *schief-symmetrische* n-Form), wenn eine von zwei äquivalenten Definitionen gilt:

(1) $\sigma\alpha = \alpha$ für alle $\sigma \in S_n$.

(2) $\alpha(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -\alpha(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n)$ für alle $1 \leqslant i < j \leqslant n$.
Wir definieren die Menge

$$\mathrm{Alt}^n(V) := \{\alpha \in Q^n(V) \,|\, \alpha \text{ ist alternierend}\}.$$

Satz 13.1.7 Sei V ein endlichdimensionaler K-Vektorraum und sei $n \in \mathbb{N}$.

(1) $\mathrm{Alt}^n(V)$ ist ein Untervektorraum des Vektorraums $Q^n(V)$.

(2) Ist $\alpha \in \mathrm{Alt}^n(V)$ und $\mathrm{char}(K) \neq 2$, dann gilt $\alpha(v_1, \ldots, v_n) = 0$, falls $v_1, \ldots, v_n$ linear abhängig sind. Insbesondere gilt $\alpha(v_1, \ldots, v_n) = 0$, wenn $v_i = v_j$ für einige $i \neq j$ ist.

Beispiel. Sei V ein K-Vektorraum der endlichen Dimension n. Sei $e_1, \ldots, e_n$ eine Basis von V. Wir definieren eine Abbildung

$$\mathbf{det} : V^n \to K$$
$$(v_1, \ldots, v_n) \mapsto \det(a_{ij}),$$

wobei

$$\begin{aligned} v_1 &= a_{11}e_1 + \cdots + a_{1n}e_n, \\ &\vdots \\ v_n &= a_{n1}e_1 + \cdots + a_{nn}e_n. \end{aligned}$$

ist. Dann gilt $\mathbf{det} \in \mathrm{Alt}^n(V)$.

Verabredung. Des Weiteren sei $\mathrm{char}(K) = 0$. Das ist erfüllt für $K = \mathbb{Q}, \mathbb{R}, \mathbb{C}$.

Definition 13.1.8 Für eine k-Form $\alpha \in Q^k(V)$ definieren wir eine andere k-Form $\mathrm{Pr}(\alpha) \in Q^k(V)$ durch

$$\mathrm{Pr}(\alpha) := \frac{1}{k!} \sum_{\sigma \in S_k} \sigma\alpha.$$

Bemerkung. Für alle $\alpha \in Q^k(V)$ gilt $\mathrm{Pr}(\alpha) \in \mathrm{Alt}^k(V)$. Deswegen heißt $\mathrm{Pr}(\alpha)$ *Alternator* von α.

Satz 13.1.9 Die Abbildung

$$\begin{aligned} \mathrm{Pr} : Q^k(V) &\to \mathrm{Alt}^k(V), \\ \alpha &\mapsto \mathrm{Pr}(\alpha) \end{aligned}$$

ist eine Projektion, d.h. linear und identisch auf $\mathrm{Alt}^k(V)$.

Definition 13.1.10 Für zwei alternierende Formen $\alpha \in \mathrm{Alt}^k(V)$ und $\beta \in \mathrm{Alt}^\ell(V)$ definieren wir ihr *äußeres Produkt* (oder das *Dachprodukt*) $\alpha \wedge \beta \in \mathrm{Alt}^{k+\ell}(V)$ wie folgt:

$$\alpha \wedge \beta := \frac{(k+\ell)!}{k!\,\ell!} \mathrm{Pr}(\alpha \bullet \beta).$$

(Zur Erinnerung: $\alpha \bullet \beta$ wurde in Definition 13.1.3 gegeben.)

Bemerkung. Bis zu diesem Moment haben wir zwei Abbildungen definiert:

$$\bullet : Q^k \times Q^\ell \to Q^{k+l} \quad \text{und} \quad \wedge : \mathrm{Alt}^k(V) \times \mathrm{Alt}^\ell(V) \to \mathrm{Alt}^{k+\ell}(V).$$

Notation.

(1) Für eine Untergruppe H einer Gruppe G bezeichnen wir mit $\mathbf{Rep}(H, G)$ die Menge der Repräsentanten der linken Nebenklassen von H in G. Wenn also $\mathbf{Rep}(H, G) = \{g_1, \dots g_n\}$ ist, dann gilt
$$G = g_1 H \sqcup \cdots \sqcup g_n H.$$

(2) Sei $H_{k,\ell}$ die Untergruppe von $S_{k+\ell}$, die die Menge $\{1, \dots, k\}$ auf sich abbildet (dann bildet sie auch die Menge $\{k+1, \dots, k+\ell\}$ auf sich ab). Da die Gruppen $H_{k,\ell}$ und $S_k \times S_\ell$ isomorph sind, schreiben wir $S_k \times S_\ell$ statt $H_{k,\ell}$.

Lemma 13.1.11 Sei $\alpha \in \mathrm{Alt}^k(V)$ und $\beta \in \mathrm{Alt}^\ell(V)$. Dann gilt

$$\alpha \wedge \beta = \sum_{\mathbf{Rep}(S_k \times S_\ell, S_{k+\ell})} \sigma(\alpha \bullet \beta).$$

Satz 13.1.12 Das Dachprodukt $\wedge$ ist

(1) bilinear, also linear in jedem Faktor bei festem anderen,

(2) graduiert antikommutativ: für alle $\alpha \in \mathrm{Alt}^k(V)$ und $\beta \in \mathrm{Alt}^\ell(V)$ gilt

$$\alpha \wedge \beta = (-1)^{k \cdot \ell} \cdot \beta \wedge \alpha,$$

(3) assoziativ: für alle $\alpha \in \mathrm{Alt}^k(V)$, $\beta \in \mathrm{Alt}^\ell(V)$ und $\gamma \in \mathrm{Alt}^r(V)$ gilt

$$(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma) = \frac{(k+\ell+r)!}{k!\,\ell!\,r!} \,\mathrm{Pr}(\alpha \bullet \beta \bullet \gamma).$$

Bemerkung. Das Assoziativitätsgesetz berechtigt uns, $\alpha_1 \wedge \cdots \wedge \alpha_\ell$ ohne Klammern zu schreiben. Per Induktion erhalten wir

$$\alpha_1 \wedge \cdots \wedge \alpha_\ell = \frac{(k_1 + \cdots + k_\ell)!}{k_1! \dots k_\ell!} \,\mathrm{Pr}(\alpha_1 \bullet \cdots \bullet \alpha_\ell),$$

wobei k_i der Grad von α_i ist.

Bemerkung.
(1) Es wird gesetzt $\mathrm{Alt}^0(V) := K$.
(2) Es gilt $\mathrm{Alt}^1(V) = Q^1(V) = V^*$.
(3) Ist $n = \dim(V)$ endlich, dann gilt $\mathrm{Alt}^k(V) = 0$ für alle $k > n$.

Definition 13.1.13 Sei $n = \dim(V)$. Die direkte Summe

$$\mathrm{Alt}(V) := \bigoplus_{k=0}^{n} \mathrm{Alt}^k(V)$$

heißt *äußere* Algebra von V. Die Multiplikation in dieser Algebra ist durch das Dachprodukt induziert. Die Summa $\alpha + \beta$ für α und β aus verschiedenen Komponenten wird als formale Kombination definiert.[4]

Satz 13.1.14 Seien $\alpha_1, \dots, \alpha_\ell$ aus $\mathrm{Alt}^1(V)$. Dann gilt

$$(\alpha_1 \wedge \cdots \wedge \alpha_\ell)(v_1, \dots, v_\ell) = \det(\alpha_i(v_j)).$$

[4] Diese Summa wird schon keine Form sein. Deswegen kann sie nicht auf einem Tupel von Vektoren aus V bewertet werden.

Beweis.

$$\begin{aligned}
(\alpha_1 \wedge \cdots \wedge \alpha_\ell)(v_1, \ldots, v_\ell) &= \ell! \cdot \mathrm{Pr}(\alpha_1 \bullet \cdots \bullet \alpha_\ell)(v_1, \ldots, v_\ell) \\
&= \ell! \cdot \frac{1}{\ell!} \sum_{\sigma \in S_\ell} \sigma(\alpha_1 \bullet \cdots \bullet \alpha_\ell)(v_1, \ldots, v_\ell) \\
&= \sum_{\sigma \in S_\ell} \mathrm{sign}(\sigma)(\alpha_1 \bullet \cdots \bullet \alpha_\ell)(v_{\sigma(1)}, \ldots, v_{\sigma(\ell)}) \\
&= \sum_{\sigma \in S_\ell} \mathrm{sign}(\sigma) \cdot \alpha_1(v_{\sigma(1)}) \cdot \ldots \cdot \alpha_\ell(v_{\sigma(\ell)}) \\
&= \det(\alpha_i(v_j)).
\end{aligned}$$

□

Folgerung 13.1.15 Sei $e = (e_1, \ldots, e_n)$ eine Basis von V und sei $e^* = (e_1^*, \ldots, e_n^*)$ die Dualbasis in V^*. Seien

$$\begin{aligned}
v_1 &= a_{11}e_1 + \cdots + a_{1n}e_n, \\
&\vdots \\
v_n &= a_{n1}e_1 + \cdots + a_{nn}e_n.
\end{aligned}$$

Vektoren aus V. Dann gilt

$$(e_1^* \wedge \cdots \wedge e_n^*)(v_1, \ldots, v_n) = \det(a_{ij})$$

Insbesondere gilt

$$(e_1^* \wedge \cdots \wedge e_n^*)(e_1, \ldots, e_n) = 1.$$

Deswegen kann die alternierende Form $e_1^* \wedge \cdots \wedge e_n^*$ mit der Determinante $\mathbf{det} : V^n \to K$ identifiziert werden.

Satz 13.1.16 Sei $\dim(V) = n$ endlich. Die Elemente

$$e_{i_1}^* \wedge \cdots \wedge e_{i_k}^*, \quad 1 \leqslant i_1 < \cdots < i_k \leqslant n,$$

bilden eine Basis von $\mathrm{Alt}^k(V)$. Insbesondere gilt

$$\dim(\mathrm{Alt}^k(V)) = \binom{n}{k}$$

und

$$\dim(\mathrm{Alt}(V)) = 2^n.$$

Beweis. Beachte:

- Eine alternierende k-Form auf V ist bestimmt duch ihre Wirkung auf allen k-Tupel der Form $(e_{j_1}, \ldots e_{j_k})$ mit $1 \leqslant j_1 < \cdots < j_k \leqslant n$.

- Wenn $1 \leqslant i_1 < \cdots < i_k \leqslant n$ und $1 \leqslant j_1 < \cdots < j_k \leqslant n$ ist, dann gilt:

$$(i_1, \ldots, i_k) \neq (j_1, \ldots, j_k) \Leftrightarrow i_1 \neq j_1,\ i_2 \neq j_2, \ldots, i_k \neq j_k.$$

Dann ist

$$(e_{i_1}^* \wedge \cdots \wedge e_{i_k}^*)(e_{j_1}, \ldots, e_{j_k}) = \det(e_{i_p}^*(e_{j_q}))$$

gleich 0, falls $(i_1, \ldots, i_k) \neq (j_1, \ldots, j_k)$, und 1 sonst. Daraus folgt, dass die angegebenen Elemente linear unabhängig sind, und dass der von ihnen erzeugte Raum auf den $(e_{j_1}, \ldots, e_{j_k})$ alle Werte annimmt. □

14 Appendix

14.1 Gruppentheorie

Lemma 14.1.1 Sei $A \leqslant B \leqslant G$ eine Kette von Gruppen. Seien
$\{b_1, \ldots, b_m\}$ Repräsentanten von linken Nebenklassen von A in B,
$\{g_1, \ldots, g_n\}$ Repräsentanten von linken Nebenklassen von B in G.
Dann sind $\{g_i b_j \mid i = 1, \ldots, n,\ j = 1, \ldots, m\}$ Repräsentanten von linken Nebenklassen von A in G. Kurz:

$$\mathbf{Rep}(B, G) \cdot \mathbf{Rep}(A, B) = \mathbf{Rep}(A, G).$$

Lemma 14.1.2 Seien $H \leqslant G$ Gruppen und sei A eine weitere Gruppe. Sind $\{g_1, \ldots, g_n\}$ Repräsentanten von linken Nebenklassen von H in G, dann sind $\{(g_1, e_A), \ldots, (g_n, e_A)\}$ Repräsentanten [5] von linken Klassen von $H \times A$ in $G \times A$.

[5] Hierbei ist e_A das neutrale Element von A.