Lineare Algebra II

Burkhard Külshammer

SS 05
Universität Jena

Inhaltsverzeichnis

17 Die Jordansche Normalform

K Körper

Kurze Wiederholung und Motivation

Seien V ein endlich-dimensionaler K-Vektorraum und $f \in \mathrm{End}(V)$, d.h. $f : V \longrightarrow V$ linear. Man nennt f ***diagonalisierbar***, falls eine Basis $b_1, \ldots, b_n$ von V existiert mit der Eigenschaft, dass die Matrix A von f bzgl. $b_1, \ldots, b_n$ eine Diagonalmatrix ist:

$$A = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}.$$

Das bedeutet: $f(b_i) = a_i b_i$ $(i = 1, \ldots, n)$.
Für das ***charakteristische Polynom*** $p = \det(X \cdot \mathrm{id}_V - f) = \det(X \cdot 1_n - A)$ von f gilt also:

$$p = (X - a_1) \cdots (X - a_n);$$

insbesondere sind $a_1, \ldots, a_n$ die ***Eigenwerte*** von f.
Nicht jedes $f \in \mathrm{End}(V)$ ist diagonalisierbar. Betrachte etwa $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2, (x, y) \longmapsto (0, x)$. Die Matrix von f bzgl. der Standardbasis $(1, 0), (0, 1)$ ist

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

Das charakteristische Polynom von f ist also $p = X^2$. Daher ist 0 der einzige Eigenwert von f. Wäre f diagonalisierbar, so gäbe es eine Basis b_1, b_2 von $\mathbb{R}^2$, bzgl. der die Matrix von f die folgende Form hat:

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

Dann wäre aber $f(v) = 0$ für alle $v \in \mathbb{R}^2$, was nicht der Fall ist. Also ist f nicht diagonalisierbar.
Wir werden im folgenden die Frage untersuchen, ob man auch für nicht diagonalisierbare Abbildungen eine Basis finden kann, bzgl. der die Matrix "möglichst einfach" ist.
Wir werden sehen, dass dies zumindest dann gilt, wenn K algebraisch abgeschlossen (z.B. $K = \mathbb{C}$) ist. Dann kann man nämlich eine Basis von V so wählen, dass die entsprechende Matrix folgende Form hat:

$$\begin{pmatrix} \star & & & & & 0 \\ \star & \star & & & & \\ & \star & \star & & & \\ & & \star & \star & & \\ & & & \star & \star & \\ 0 & & & & \star & \star \end{pmatrix} \qquad \text{(Jordansche Normalform)}$$

Dabei stehen unterhalb der Hauptdiagonalen nur Nullen und Einsen. Wichtige Anwendungen liefert die Theorie linearer Differentialgleichungen (Differenzengleichungen).

Definition 17.1. Seien V ein endlich-dimensionaler K-Vektorraum und $f \in \mathrm{End}(V)$. Ein Untervektorraum U von V mit $f(U) \subseteq U$ heißt f-***invariant***.

Bemerkung 17.1. (i) Ggf. ist die Einschränkung $f|U : U \longrightarrow U, u \longmapsto f(u)$ wieder linear. Wählt man eine Basis $b_1, \ldots, b_m$ von U und ergänzt man diese zu einer Basis $b_1, \ldots, b_m, b_{m+1}, \ldots, b_n$ von V, so hat die Matrix von f bzgl. $b_1, \ldots, b_n$ die Form

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

dabei ist $B \in K^{m \times m}$ die Matrix von $f|U$ bzgl. $b_1, \ldots, b_m$.

(ii) Ist $V = U_1 \oplus U_2$ mit f-invarianten Untervektorräumen U_1, U_2 und wählt man Basen $b_1, \ldots, b_m$ von U_1 und $b_{m+1}, \ldots, b_n$ von U_2, so ist $b_1, \ldots, b_m, b_{m+1}, \ldots, b_n$ eine Basis von V, bzgl. der die Matrix von f die folgende Form hat:

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

dabei sind B und D die Matrizen von $f|U_1$ bzw. $f|U_2$ bzgl. der Basen $b_1, \ldots, b_m$ bzw. $b_{m+1}, \ldots, b_n$. Fragen über f lassen sich häufig auf Fragen über $f|U_1$ und $f|U_2$ zurückführen. Da U_1, U_2 i. Allg. kleinere Dimensionen haben, lassen sich diese leichter und schneller beantworten.

Satz 17.1. *(Fitting)*
Seien V ein endlich-dimensionaler K-Vektorraum und $f \in \mathrm{End}(V)$. wir setzen

$$f^0 := \mathrm{id}_V, f^1 = f, f^2 = f \circ f, f^3 = f \circ f \circ f, \ldots$$

Dann existiert ein $k \in \mathbb{N}_0$ mit folgenden Eigenschaften:

(i) $V = \mathrm{Bld}(f^0) \supset \mathrm{Bld}(f^1) \supset \mathrm{Bld}(f^2) \supset \cdots \supset \mathrm{Bld}(f^k) = \mathrm{Bld}(f^{k+1}) = \ldots$

(ii) $0 = \mathrm{Ker}(f^0) \subset \mathrm{Ker}(f^1) \subset \mathrm{Ker}(f^2) \subset \cdots \subset \mathrm{Ker}(f^k) = \mathrm{Ker}(f^{k+1}) = \ldots$

(iii) $V = \mathrm{Ker}(f^k) \oplus \mathrm{Bld}(f^k)$ *(Fitting-Zerlegung).*

Dabei sind alle auftretenden Untervektorräume f-invariant.

Beweis. Wegen $V \supseteq f(V)$ ist

$$f(V) \supseteq f(f(V)) = f^2(V),$$
$$f^2(V) = f(f(V)) \supseteq f(f^2(V)) = f^3(V), \ldots$$

Daher ist $V = \mathrm{Bld}\,(f^0) \supseteq \mathrm{Bld}\,(f^1) \supseteq \mathrm{Bld}\,(f^2) \supseteq \ldots$. Wegen $\dim V < \infty$ existiert ein $k \in \mathbb{N}_0$ mit $\mathrm{Bld}\,(f^k) = \mathrm{Bld}\,(f^{k+1})$. Dann ist aber

$$\mathrm{Bld}\,(f^{k+1}) = f(f^k(V)) = f(f^{k+1}(V)) = \mathrm{Bld}\,(f^{k+2}).$$

Es existiert also ein $k \in \mathbb{N}_0$, das (i) erfüllt.
Für $v \in \mathrm{Ker}\,(f)$ ist $f(v) = 0$, also auch $f^2(v) = f(f(v)) = f(0) = 0$, d.h. $v \in \mathrm{Ker}\,(f^2)$. Daher ist $\mathrm{Ker}\,(f) \subseteq \mathrm{Ker}\,(f^2)$. Analog zeigt man $\mathrm{Ker}\,(f^2) \subseteq \mathrm{Ker}\,(f^3), \ldots$. Also ist

$$0 = \mathrm{Ker}\,(f^0) \subseteq \mathrm{Ker}\,(f^1) \subseteq \mathrm{Ker}\,(f^2) \subseteq \ldots$$

Wegen $\dim V < \infty$ existiert ein $l \in \mathbb{N}_0$ mit $\mathrm{Ker}\,(f^l) = \mathrm{Ker}\,(f^{l+1})$. Für $v \in \mathrm{Ker}\,(f^{l+2})$ ist dann $0 = f^{l+1}(f(v))$, d.h. $f(v) \in \mathrm{Ker}\,(f^{l+1}) = \mathrm{Ker}\,(f^l)$. Daher ist auch $0 = f^l(f(v)) = f^{l+1}(v)$, d.h. $v \in \mathrm{Ker}\,(f^{l+1})$. Damit ist gezeigt: $\mathrm{Ker}\,(f^{l+1}) = \mathrm{Ker}\,(f^{l+2})$. Es existiert also ein $l \in \mathbb{N}_0$ mit

$$0 = \mathrm{Ker}\,(f^0) \subset \mathrm{Ker}\,(f^1) \subset \mathrm{Ker}\,(f^2) \subset \cdots \subset \mathrm{Ker}\,(f^l) = \mathrm{Ker}\,(f^{l+1}) = \ldots$$

Wegen $\mathrm{Bld}\,(f^k) = \mathrm{Bld}\,(f^{k+1})$ ist

$$\begin{aligned}\dim \mathrm{Ker}\,(f^k) &= \dim V - \dim \mathrm{Bld}\,(f^k) \\ &= \dim V - \dim \mathrm{Bld}\,(f^{k+1}) = \dim \mathrm{Ker}\,(f^{k+1}),\end{aligned}$$

also $k \geq l$. Analog ist $k \leq l$, also $k = l$. Damit existiert ein $k \in \mathbb{N}_0$, das (i) und (ii) erfüllt. Wir zeigen, dass dann auch (iii) gilt.
Sei $v \in \mathrm{Ker}\,(f^k) \cap \mathrm{Bld}\,(f^k)$. Schreibe $v = f^k(u)$ mit $u \in V$. Dann ist $0 = f^k(v) = f^k(f^k(u)) = f^{2k}(u)$, d.h. $u \in \mathrm{Ker}\,(f^{2k}) = \mathrm{Ker}\,(f^k)$. Daher ist $0 = f^k(u) = v$. Damit ist gezeigt: $\mathrm{Ker}\,(f^k) \cap \mathrm{Bld}\,(f^k) = \{0\}$. Folglich ist

$$\dim\left(\mathrm{Ker}\,(f^k) + \mathrm{Bld}\,(f^k)\right) = \dim \mathrm{Ker}\,(f^k) + \dim \mathrm{Bld}\,(f^k) = \dim V,$$

d.h. $V = \mathrm{Ker}\,(f^k) + \mathrm{Bld}\,(f^k) = \mathrm{Ker}\,(f^k) \oplus \mathrm{Bld}\,(f^k)$. Damit ist (iii) bewiesen.
Für $i \in \mathbb{N}_0$ ist $f(f^i(V)) = f^{i+1}(V) \subseteq f^i(V)$, d.h. $f^i(V)$ ist f-invariant.
Für $i \in \mathbb{N}_0$ ist auch $f(\mathrm{Ker}\,(f^i)) \subseteq \mathrm{Ker}\,(f^i)$; denn aus $f^i(v) = 0$ folgt $f^{i+1}(v) = 0$, d.h. $f(v) \in \mathrm{Ker}\,(f^i)$. Daher ist auch $\mathrm{Ker}\,(f^i)$ f-invariant. □

Beispiel 17.1. Betrachte $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3, (x, y, z) \longmapsto (-x + y, -y + z, x - z)$.
Die Matrix von f bzgl. der Standardbasis ist

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

Man berechnet: $\dim \mathrm{Bld}\,(f) = \mathrm{rg}\,(A) = \cdots = 2$.

$$A^2 = \begin{pmatrix} 1 & -2 & 1 \\ 1 & 1 & -2 \\ -2 & 1 & 1 \end{pmatrix}, \qquad \dim \mathrm{Bld}\,(f^2) = \mathrm{rg}\,(A^2) = 2.$$

Nach Fitting ist also $\mathbb{R}^3 = \operatorname{Ker}(f) \oplus \operatorname{Bld}(f)$.
Basis von $\operatorname{Ker}(f)$: $(1,1,1)$
Basis von $\operatorname{Bld}(f)$: $(-1,0,1), (0,1,-1)$
Also bilden $(1,1,1), (-1,0,1), (0,1,-1)$ eine Basis von $\mathbb{R}^3$.
$f(-1,0,1) = (1,1,-2) = -1(-1,0,1) + 1(0,1,-1)$
$f(0,1,-1) = (1,-2,1) = -1(-1,0,1) - 2(0,1,-1)$
Matrix von f bzgl. der Basis $(1,1,1), (-1,0,1), (0,1,-1)$:

$$B = \left(\begin{array}{c|cc} 0 & 0 & 0 \\ \hline 0 & -1 & -1 \\ 0 & 1 & -2 \end{array}\right).$$

Bemerkung 17.2. Die Bezeichnungen seien wie in 17.1. Dann gilt:

(i) $\operatorname{Bld}(f^k) = V \Leftrightarrow f^k$ surjektiv $\Leftrightarrow f^k$ bijektiv $\Leftrightarrow f$ bijektiv

(ii) $\operatorname{Ker}(f^k) = V \Leftrightarrow 0 = f^k$.
Ein $g \in \operatorname{End}(V)$ mit $g^m = 0$ für ein $m \in \mathbb{N}$ nennt man ***nilpotent***.

(iii) Ist f weder bijektiv noch nilpotent, so sind $\operatorname{Bld}(f^k)$ und $\operatorname{Ker}(f^k)$ echte Untervektorräume von V. Ferner ist $f|\operatorname{Bld}(f^k)$ bijektiv, und $f|\operatorname{Ker}(f^k)$ ist nilpotent. Wir haben also f einen bijektiven und einen nilpotenten Teil zerlegt.

Definition 17.2. Für $m \in \mathbb{N}$ sei

$$J_m := \begin{pmatrix} 0 & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & 0 \end{pmatrix} \in K^{m \times m}$$

Satz 17.2. *Seien V ein endlich-dimensionaler K-Vektorraum und $f \in \operatorname{End}(V)$ nilpotent. Dann existiert eine Basis von V, bzgl. der die Matrix von f folgende Form hat:*

$$A = \begin{pmatrix} J_{k_1} & & & 0 \\ & J_{k_2} & & \\ & & \ddots & \\ 0 & & & J_{k_l} \end{pmatrix}.$$

Beweis. Sei $m \in \mathbb{N}$ mit $f^m = 0$. Nach 17.1 ist dann

$$0 \subseteq \operatorname{Ker}(f) \subseteq \operatorname{Ker}(f^2) \subseteq \cdots \subseteq \operatorname{Ker}(f^{m-1}) \subseteq \operatorname{Ker}(f^m) = V$$

Für $x \in \operatorname{Ker}(f^i)$ ist $f^{i-1}(f(x)) = f^i(x) = 0$, d.h. $f(x) \in \operatorname{Ker}(f^{i-1})$. Also ist $f(\operatorname{Ker}(f^i)) \subseteq \operatorname{Ker}(f^{i-1})$ für alle i. Wir wählen Untervektorräume $U_1, U_2, \ldots, U_m$ mit

$$\begin{aligned} V = \operatorname{Ker}(f^m) &= \operatorname{Ker}(f^{m-1}) \oplus U_1, \\ \operatorname{Ker}(f^{m-1}) &= [\operatorname{Ker}(f^{m-2}) + f(U_1)] \oplus U_2, \\ \operatorname{Ker}(f^{m-2}) &= [\operatorname{Ker}(f^{m-3}) + f^2(U_1) + f(U_2)] \oplus U_3, \\ &\ldots \\ \operatorname{Ker}(f) &= [0 + f^{m-1}(U_1) + f^{m-2}(U_2) + \cdots + f(U_{m-1})] \oplus U_m. \end{aligned}$$

Wir zeigen zunächst induktiv, dass für $i = 0, 1, \ldots, m-1$ gilt:

$$(\star) \qquad \operatorname{Ker}(f^{m-i}) = \operatorname{Ker}(f^{m-i-1}) \oplus f^i(U_1) \oplus f^{i-1}(U_2) \oplus \cdots \oplus f(U_i) \oplus U_{i+1}.$$

Im Fall $i = 0$ folgt dies aus der Wahl von U_1. Sei also $i > 0$ und bereits gezeigt, dass gilt:

$$\operatorname{Ker}(f^{m-i+1}) = \operatorname{Ker}(f^{m-i}) \oplus f^{i-1}(U_1) \oplus f^{i-2}(U_2) \oplus \cdots \oplus f(U_{i-1}) \oplus U_i.$$

Ferner sei

$$0 = x + f^i(u_1) + f^{i-1}(u_2) + \cdots + f(u_i) + u_{i+1}$$

mit $x \in \operatorname{Ker}(f^{m-i-1}), u_1 \in U_1, \ldots, u_{i+1} \in U_{i+1}$. Nach Wahl von U_{i+1} ist dann $u_{i+1} = 0$. Daher ist

$$\begin{aligned} 0 = f^{m-i-1}(0) &= 0 + f^{m-1}(u_1) + f^{m-2}(u_2) + \cdots + f^{m-i}(u_i) \\ &= f^{m-i}(f^{i-1}(u_1) + f^{i-2}(u_2) + \cdots + u_i). \end{aligned}$$

Folglich ist $f^{i-1}(u_1) + f^{i-2}(u_2) + \cdots + u_i \in \operatorname{Ker}(f^{m-i})$. Aus der Induktionsvoraussetzung folgt also $f^{i-1}(u_1) = f^{i-2}(u_2) = \cdots = u_i = 0$. Daher ist auch

$$f^i(u_1) = f^{i-1}(u_2) = \cdots = f(u_i) = 0.$$

Damit ist auch $x = 0$ und $(\star)$ bewiesen. Folglich gilt:

$$\begin{aligned} V = U_1 &\oplus f(U_1) \oplus f^2(U_1) \oplus \cdots \oplus f^{m-1}(U_1) \\ &\oplus U_2 \oplus f(U_2) \oplus f^2(U_2) \oplus \cdots \oplus f^{m-2}(U_2) \\ &\oplus U_3 \oplus f(U_3) \oplus \cdots \oplus f^{m-3}(U_3) \\ &\oplus \cdots \oplus \\ &\oplus U_m. \end{aligned}$$

Wir wählen jetzt Basen $b_1, \ldots, b_{r_1}$ von $U_1, b_{r_1+1}, \ldots, b_{r_2}$ von $U_2, \ldots, b_{r_{m-1}+1}, \ldots, b_{r_m}$ von U_m

und behaupten, dass die folgenden Elemente eine Basis von V bilden:

$$
\begin{aligned}
&b_1, f(b_1), f^2(b_1), \ldots, f^{m-1}(b_1),\\
&\vdots\\
&b_{r_1}, f(b_{r_1}), f^2(b_{r_1}), \ldots, f^{m-1}(b_{r_1}),\\
&b_{r_1+1}, f(b_{r_1+1}), f^2(b_{r_1+1}), \ldots, f^{m-2}(b_{r_1+1}),\\
&\vdots\\
&b_{r_2}, f(b_{r_2}), f^2(b_{r_2}), \ldots, f^{m-2}(b_{r_2}),\\
&\vdots\\
&b_{r_{m-1}+1},\\
&\vdots\\
&b_{r_m}.
\end{aligned}
$$

Offenbar spannen diese Elemente V auf. Daher genügt es zu zeigen, dass für $i = 1, \ldots, m$ und $j = 1, \ldots, m - i$ die Einschränkung $f^j : U_i \longrightarrow f^j(U_i)$ injektiv ist. Dies folgt wegen $U_i \cap \operatorname{Ker}(f^j) \subseteq U_i \cap \operatorname{Ker}(f^{m-i}) = 0$.
Bzgl. der obigen Basis hat also die Matrix von f die gewünschte Form:

$$
\begin{pmatrix}
J_m & & & & & & 0\\
& \ddots & & & & & \\
& & J_m & & & & \\
& & & J_{m-1} & & & \\
& & & & \ddots & & \\
& & & & & J_{m-1} & \\
0 & & & & & & \ddots
\end{pmatrix},
$$

wobei es r_1 Blöcke J_m, r_2 Blöcke J_{m-1}, usw. gibt. □

Beispiel 17.2. Seien $V := \mathbb{R}^5$ und $f \in \operatorname{End}(V)$ mit Matrix

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

bzgl. der Standardbasis. Dann hat f^2 Matrix A^2, f^3 Matrix $A^3, \ldots$ bzgl. der Standardbasis.

Ferner ist

$$A^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & -2 & 0 & 0 & 1 \\ -1 & 2 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \text{ und } A^3 = 0, \text{ d.h. } f^3 = 0.$$

Wie im Beweis des Satzes berechnen wir zunächst $\operatorname{Ker}(f^2)$:

$$(x_1, \ldots, x_5) \in \operatorname{Ker}(f^2) \Leftrightarrow A^2 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = 0 \Leftrightarrow x_1 = 2x_2 - x_5$$

Die folgenden Elemente bilden also eine Basis von $\operatorname{Ker}(f^2)$:

$$(-1, 0, 0, 0, 1),\ (0, 0, 0, 1, 0),\ (0, 0, 1, 0, 0),\ (2, 1, 0, 0, 0).$$

Diese ergänzen wir zu einer Basis von V, z.B. durch

$$b_1 := (1, 0, 0, 0, 0).$$

Es ist also $U_1 = \mathbb{R}b_1$. Ferner ist

$$f(b_1) = (2, 1, 0, 1, 0), f^2(b_1) = (0, 0, 1, -1, 0).$$

Als nächstes berechnen wir $\operatorname{Ker}(f)$:

$$(x_1, \ldots, x_5) \in \operatorname{Ker}(f) \Leftrightarrow A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = 0 \Leftrightarrow \cdots \Leftrightarrow \begin{cases} x_3 = -x_4 \\ x_2 = 0 \\ x_1 = -x_5 \end{cases}$$

Eine Basis von $\operatorname{Ker}(f)$ wird also gegeben durch:

$$(-1, 0, 0, 0, 1),\ (0, 0, -1, 1, 0).$$

Nach obigem Beweis ist $\operatorname{Ker}(f) + f(U_1) = \operatorname{Ker}(f) \oplus f(U_1)$, und eine Basis dieses Untervektorraums wird gegeben durch

$$(-1, 0, 0, 0, 1),\ (0, 0, -1, 1, 0),\ (2, 1, 0, 1, 0).$$

Wir ergänzen diese zu einer Basis von $\operatorname{Ker}(f^2)$, z.B. durch

$$b_2 := (2, 1, 0, 0, 0).$$

Es ist also $U_2 = \mathbb{R}b_2$. Ferner ist $f(b_2) = (1, 0, 2, -2, -1)$. Wir erhalten also folgende Basis von V:

$$b_1,\ f(b_1),\ f^2(b_1),\ b_2,\ f(b_2).$$

Bzgl. dieser Basis hat die Matrix von f folgende Form:

$$\left(\begin{array}{ccc|cc} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right) = \begin{pmatrix} J_3 & 0 \\ 0 & J_2 \end{pmatrix}.$$

Schema:

Untervektorräume:	V	$=$	$\mathrm{Ker}\,(f^3)$	$\supseteq$	$\mathrm{Ker}\,(f^2)$	$\supseteq$	$\mathrm{Ker}\,(f)$	$\supseteq$	$\mathrm{Ker}\,(f^0) = 0$
Dimensionen:			5		4		2		0
Basis:			b_1		$f(b_1)$		$f^2(b_1)$		
					b_2		$f(b_2)$		

Matrix:

$$\begin{pmatrix} J_3 & 0 \\ 0 & J_2 \end{pmatrix} = \left(\begin{array}{ccc|cc} 0 & 0 & 0 & & \\ 1 & 0 & 0 & & 0 \\ 0 & 1 & 0 & & \\ \hline & & & 0 & 0 \\ & 0 & & 1 & 0 \end{array}\right)$$

Größeres Beispiel:

Untervektorräume:	$V =$	$\mathrm{Ker}\,(f^3)$	$\supseteq$	$\mathrm{Ker}\,(f^2)$	$\supseteq$	$\mathrm{Ker}\,(f)$	$\supseteq$	$\mathrm{Ker}\,(f^0)$	$= 0$
Dimensionen:		15		13		8			
Basis:		b_1		$f(b_1)$		$f^2(b_1)$			
		b_2		$f(b_2)$		$f^2(b_2)$			
				b_3		$f(b_3)$			
				b_4		$f(b_4)$			
				b_5		$f(b_5)$			
						b_6			
						b_7			
						b_8			

Matrix:

$$\begin{pmatrix} J_3 & & & & & & 0 \\ & J_3 & & & & & \\ & & J_2 & & & & \\ & & & J_2 & & & \\ & & & & J_2 & & \\ & & & & & J_1 & \\ & & & & & & J_1 \\ 0 & & & & & & J_1 \end{pmatrix}$$

Satz 17.3. *Sei $A \in K^{n\times n}$* **nilpotent** *(d.h. $A^m = 0$ für ein $m \in \mathbb{N}$).*
Dann existieren $k_1, \dots, k_l \in \mathbb{N}$ mit der Eigenschaft, dass A zu einer Matrix der folgenden Form ähnlich ist:

$$B = \begin{pmatrix} J_{k_1} & & 0 \\ & \ddots & \\ 0 & & J_{k_l} \end{pmatrix}$$

(d.h. $B = S^{-1}AS$ für ein $S \in \mathrm{GL}(n, K)$).

Beweis. Sei $f \in \mathrm{End}(K^n)$ mit Matrix A bzgl. der Standardbasis von K^n. Mit A ist auch f nilpotent. Nach Satz 17.2 existiert eine Basis von K^n, bzgl. der die Matrix B von f obige Form hat. Nach Bemerkung 11.9 (ii) sind A und B ähnlich. □

Bemerkung 17.3. Wir wollen zeigen, dass die in 17.2 und 17.3 auftretenden Zahlen $k_1, \dots, k_l$ bis auf die Reihenfolge eindeutig bestimmt sind. Dazu bezeichnen wir für $i = 1, \dots, n$ mit m_i die Anzahl der $j \in \{1, \dots, l\}$ mit $k_j = i$. Die Matrix B enthält also m_1 Matrizen J_1, m_2 Matrizen J_2, usw.

Satz 17.4. *In der obigen Situation gilt für $i = 1, \dots, n$:*

$$m_i = \mathrm{rg}(A^{i+1}) + \mathrm{rg}(A^{i-1}) - 2\,\mathrm{rg}(A^i).$$

Beweis. Für $k \in \mathbb{N}_0$ gilt:

$$J_k^0 = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}, J_k^1 = \begin{pmatrix} 0 & & 0 & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & 0 \end{pmatrix}, J_k^2 = \begin{pmatrix} 0 & & & & 0 \\ 0 & \ddots & & & \\ 1 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ 0 & & 1 & 0 & 0 \end{pmatrix},$$

$$\dots, J_k^{k-1} = \begin{pmatrix} 0 & & & \\ \vdots & & 0 & \\ 0 & & & \\ 1 & 0 & \dots & 0 \end{pmatrix}, J_k^k = 0 = J_k^{k+1} = \dots$$

Daher ist

$$\mathrm{rg}(J_k^0) = k, \mathrm{rg}(J_k^1) = k-1, \mathrm{rg}(J_k^2) = k-2, \dots,$$
$$\mathrm{rg}(J_k^{k-1}) = 1, \mathrm{rg}(J_k^k) = 0 = \mathrm{rg}(J_k^{k+1}) = \dots$$

Daraus folgt:

$$\begin{aligned} \mathrm{rg}(B^0) &= n \\ \mathrm{rg}(B^1) &= n - (m_1 + m_2 + \dots + m_n) \\ \mathrm{rg}(B^2) &= n - (m_1 + 2m_2 + \dots + 2m_n) \\ \mathrm{rg}(B^3) &= n - (m_1 + 2m_2 + 3m_3 + \dots + 3m_n) \\ &\dots \end{aligned}$$

Also ist $\operatorname{rg}(B^{i+1}) + \operatorname{rg}(B^{i-1}) - 2\operatorname{rg}(B^i) = m_i$ für alle i.
Ferner sind jeweils A^i und B^i ähnlich, d.h. $\operatorname{rg}(A^i) = \operatorname{rg}(B^i)$. □

Satz 17.5. *Seien V ein endlich-dimensionaler K-Vektorraum und $f \in \operatorname{End}(V)$. Dann existiert eine Basis von V, bzgl. der die Matrix von f die folgende Form hat:*

$$A = \begin{pmatrix} J_{k_1} & & 0 \\ & \ddots & \\ & & J_{k_l} \\ 0 & & B \end{pmatrix}, \qquad B \text{ invertierbar}$$

Beweis. 17.1 und 17.2 □

Bemerkung 17.5. Wir werden zeigen, dass man auch B noch speziell wählen kann.

Definition 17.6. Für $k \in \mathbb{N}$ und $r \in K$ sei

$$J_k(r) := \begin{pmatrix} r & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & r \end{pmatrix} \in K^{k\times k}.$$

Man nennt $J_k(r)$ einen ***Jordan-Block***.

Satz 17.6. *(Jordan)*
Seien V ein endlich-dimensionaler K-Vektorraum und $f \in \operatorname{End}(V)$. Dann existiert eine Basis von V, bzgl. der die Matrix von f die folgende Form hat:

$$J = \begin{pmatrix} J_{k_1}(r_1) & & & 0 \\ & \ddots & & \\ & & J_{k_t}(r_t) & \\ 0 & & & B \end{pmatrix};$$

dabei sind $k_1, \ldots, k_t \in \mathbb{N}, r_1, \ldots, r_t \in K$, und B ist eine quadratische Matrix ohne Eigenwerte.

Beweis. (Induktion nach $\dim V$)
Im Fall $\dim V = 1$ ist nichts zu tun. Sei also $\dim V > 1$. Besitzt f keinen Eigenwert, so nehmen wir eine beliebige Basis und erhalten eine Matrix der Form

$$J = (B),$$

wobei B keine Eigenwerte hat. Wir können also annehmen, dass f einen Eigenwert $r \in K$ hat. Wir wenden die vorigen Sätze auf den Endomorphismus $f - r \cdot \mathrm{id}_V$ von V an. Die Fitting-Zerlegung liefert ein $m \in \mathbb{N}_0$ mit

$$V = \underbrace{\operatorname{Ker}\left((f - r\cdot\mathrm{id}_V)^m\right)}_{=:U} \oplus \underbrace{\operatorname{Bld}\left((f - r\cdot\mathrm{id}_V)^m\right)}_{=:W}.$$

Dabei sind U und W invariant unter $f - r \cdot \mathrm{id}_V$ und f. [Denn für $u \in U$ ist $f(u) = (f - r \cdot \mathrm{id}_V)(u) + ru \in U$; analog für W.]
Da die Einschränkung von $f - r \cdot \mathrm{id}_V$ auf U nilpotent ist, existiert nach 17.2 eine Basis von U, bzgl. der die Matrix von $(f - r \cdot \mathrm{id}_V)|U$ folgende Form hat:

$$I = \begin{pmatrix} J_{k_1} & & 0 \\ & \ddots & \\ 0 & & J_{k_2} \end{pmatrix}$$

Die Matrix $f|U$ hat also die Form

$$I + r1_k = \begin{pmatrix} J_{k_1}(r) & & 0 \\ & \ddots & \\ 0 & & J_{k_s}(r) \end{pmatrix}.$$

Sei $v \in V$ ein Eigenvektor von f zum Eigenwert von r. Dann ist $v \in U$, d.h. $U \neq 0$ und $\dim W < \dim V$. Nach Induktion existiert eine Basis von W, bzgl. der die Matrix von $f|W$ die folgende Form hat:

$$\begin{pmatrix} J_{k_{s+1}}(r_{s+1}) & & & 0 \\ & \ddots & & \\ & & J_{k_t}(r_t) & \\ 0 & & & B \end{pmatrix};$$

dabei ist B eine quadratische Matrix ohne Eigenwerte. Nach Bemerkung 17.1 existiert also eine Basis von V, bzgl. der die Matrix von f die gewünschte Form hat. □

Beispiel 17.6. Sei $V := \mathbb{R}^3$ und $f \in \mathrm{End}(V)$ mit Matrix

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

bzgl. der Standardbasis.
Charakteristisches Polynom: $|r1_3 - A| = \cdots = (r-1)^2(r-2)$.
Eigenwerte: $1, 2$
Betrachte zunächst den Eigenwert 1:

$$A - 1_3 = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix} \longrightarrow \ldots \text{ (Gauß) } \ldots \longrightarrow \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Basis von $\mathrm{Ker}\,(f - \mathrm{id}_V)$: $(1, 1, 1)$

$$(A - 1_3)^2 = \begin{pmatrix} 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{pmatrix} \longrightarrow \ldots \longrightarrow \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Basis von $\operatorname{Ker}((f-\mathrm{id}_V)^2) : (0,1,1),(1,0,0)$
$(A-1_3)^3 = \begin{pmatrix} 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{pmatrix} = (A-1_3)^2$
Fitting-Zerlegung: $V = \operatorname{Ker}((f-\mathrm{id}_V)^2) \oplus \operatorname{Bld}((f-\mathrm{id}_V)^2)$
Basis von $\operatorname{Ker}((f-\mathrm{id}_V)^2) : b_1 = (1,0,0), b_2 = (f-\mathrm{id}_V)(b_1) = (1,1,1)$
Basis von $\operatorname{Bld}((f-\mathrm{id}_V)^2) : (1,0,1), f(1,0,1) = (2,0,2) = 2(1,0,1)$.
Basis von V: $b_1 = (1,0,0), b_2 = (1,1,1), b_3 = (1,0,1)$
Matrix von f bzgl. b_1, b_2, b_3:

$$\left(\begin{array}{cc|c} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \hline 0 & 0 & 2 \end{array}\right) = \begin{pmatrix} J_2(1) & 0 \\ 0 & J_1(2) \end{pmatrix}.$$

Satz 17.7. *Jede Matrix $A \in K^{n\times n}$ ist zu einer Matrix der folgenden Form ähnlich:*

$$J = \begin{pmatrix} J_{k_1}(r_1) & & & 0 \\ & \ddots & & \\ & & J_{k_t}(r_t) & \\ 0 & & & B \end{pmatrix};$$

dabei ist B eine quadratische Matrix ohne Eigenwerte.

Beweis. Analog zu 17.3. □

Bemerkung 17.7. Man kann sich analog zu 17.4 überlegen, dass die Paare $(k_1, r_1), \ldots, (k_t, r_t)$ durch A bis auf die Reihenfolge eindeutig bestimmt sind. Genauer gilt für die Anzahl $m_i(r)$ aller $j \in \{1, \ldots, t\}$ mit $k_j = i$ und $r_j = r$:

$$m_i(r) = \operatorname{rg}((A - r1_n)^{i+1}) + \operatorname{rg}((A - r1_n)^{i-1}) - 2\operatorname{rg}((A - r1_n)^i).$$

Die Matrix B ist i.Allg. nicht eindeutig bestimmt. Oft tritt sie aber gar nicht auf.

Satz 17.8. *Seien K algebraisch abgeschlossen (z.B. $K = \mathbb{C}$), V ein endlich-dimensionaler K-Vektorraum und $f \in \operatorname{End}(V)$. Dann existiert eine Basis von V, bzgl. der die Matrix von f die folgende Form hat:*

$$J = \begin{pmatrix} J_{k_1}(r_1) & & 0 \\ & \ddots & \\ 0 & & J_{k_t}(r_t) \end{pmatrix}$$

Beweis. Satz 17.6. □

Bemerkung 17.8. (i) Das charakteristische Polynom von f und J ist gleich

$$(X - r_1)^{k_1} \cdots (X - r_t)^{k_t}.$$

Daher sind $r_1, \ldots, r_t$ genau die Eigenwerte von f. Ferner ist

$$k_1 + \cdots + k_t = n := \dim V.$$

(ii) Für $r \in K$ sei $d_r := \dim \operatorname{Ker}(f - r \cdot \mathrm{id}_V)$, d.h. d_r ist die Dimension des Eigenraums $E_r(f) = \operatorname{Ker}(f - r \cdot \mathrm{id}_V)$ von f zum Eigenwert r. Dann ist

$$d_r = |\{i : 1 \le i \le t, r_i = r\}|,$$

d.h. J enthält genau d_r Jordan-Blöcke $J_{k_i}(r_i)$ mit Eigenwert $r_i = r$. Es gilt nämlich jeweils:

$$J_{k_i}(r_i) - r1_{k_i} = \begin{pmatrix} r_i - r & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & r_i - r \end{pmatrix},$$

d.h.

$$\operatorname{rg}(J_{k_i}(r_i) - r1_{k_i}) = \begin{cases} k_i - 1 & \text{falls } r_i = r \\ k_i & \text{sonst.} \end{cases}$$

Kennt man also die Dimensionen der Eigenräume von f, so kennt man die Anzahl der Jordan-Blöcke in J.

(iii) Für $r \in K$ nennt man

$$H_r(f) := \operatorname{Ker}(f - r \cdot \mathrm{id}_V)^n \qquad (n = \dim V)$$

den ***Hauptraum*** von f zum Eigenwert r. Dann ist $E_r(f) \subseteq H_r(f)$, und $h_r := \dim H_r(f)$ ist die Anzahl der Einträge r auf der Hauptdiagonalen von J; denn für alle i mit $r_i = r$ ist $(J_{k_i}(r_i) - r1_{k_i})^n = 0$ und für alle i mit $r_i \neq r$ ist $J_{k_i}(r_i) - r1_{k_i}$ invertierbar.
Das charakteristische Polynom von f hat also die Form $(X - r)^{k_r} g$ mit $g(r) \neq 0$.

Satz 17.9. *Sei K algebraisch abgeschlossen und $A \in K^{n\times n}$. Dann ist A zu einer Matrix der folgenden Form ähnlich:*

$$J = \begin{pmatrix} J_{k_1}(r_1) & & 0 \\ & \ddots & \\ 0 & & J_{k_t}(r_t) \end{pmatrix}$$

Beweis. Satz 17.7 □

Bemerkung 17.9. (i) Man nennt J die ***Jordansche Normalform*** von A. Nach Bemerkung 17.7 ist sie durch A "im Wesentlichen" eindeutig bestimmt.

(ii) $A, B \in K^{n\times n}$ sind genau dann ähnlich, wenn sie die "gleiche" Jordansche Normalform haben. Dies bedeutet, dass sie die gleichen Eigenwerte haben und dass für $i = 1, \ldots, n$ und jeden Eigenwert r von A gilt:

$$\operatorname{rg}((A - r1_n)^i) = \operatorname{rg}((B - r1_n)^i).$$

(iii) Es gibt andere Verfahren, um festzustellen, ob vorgegebene Matrizen ähnlich sind (und die auch über Körpern funktionieren, die nicht algebraisch abgeschlossen sind). Darauf gehen wir jetzt nicht ein.

Beispiel 17.9. (i) Sei K algebraisch abgeschlossen und $A \in K^{n\times n}$. Dann ist A zu A^T ähnlich, denn für $r \in K$ und $i = 1, \ldots, n$ gilt:

$$\operatorname{rg}((A^T - r1_n)^i) = \operatorname{rg}(((A - r \cdot 1_n)^i)^T) = \operatorname{rg}((A - r \cdot 1_n)^i).$$

Die Aussage gilt auch für Körper, die nicht algebraisch abgeschlossen sind, muss dann aber anders bewiesen werden.

(ii) Sei K algebraisch abgeschlossen und $A \in K^{2\times 2}$. Das charakteristische Polynom von A hat also die Form $(X - a)(X - b)$ mit $a, b \in K$. Dann tritt einer der folgenden Fälle auf:

(I) A ist diagonalisierbar.
Dann ist A zu der folgenden Diagonalmatrix ähnlich:

$$D = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}.$$

(II) A ist nicht diagonalisierbar.
Dann ist $a = b$ (vgl. Satz 13.7), und A ist zu der folgenden Matrix ähnlich:

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

(iii) Sei K algebraisch abgeschlossen und $A \in K^{3\times 3}$. Dann ist A zu einer der folgenden Matrizen ähnlich:

(I) $D = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$ mit $a, b, c \in K$; in diesem Fall ist A diagonalisierbar.

(II) $B = \left(\begin{array}{cc|c} a & 0 & 0 \\ 1 & a & 0 \\ \hline 0 & 0 & b \end{array}\right)$ mit $a, b \in K$.

(III) $J = \begin{pmatrix} a & 0 & 0 \\ 1 & a & 0 \\ 0 & 1 & a \end{pmatrix}$ mit $a \in K$.

Bemerkung 17.10. Bekanntlich ist $\mathbb{R}$ nicht algebraisch abgeschlossen. Man kann also die Sätze 17.8 und 17.9 nicht direkt auf reelle Matrizen anwenden. Ohne Beweis sei erwähnt,

dass jede Matrix $A \in \mathbb{R}^{n\times n}$ zu einer Matrix der folgenden Form ähnlich ist:

$$B = \begin{pmatrix} J_{k_1}(r_1) & & & & & & \\ & \ddots & & & & & \\ & & J_{k_s}(r_s) & & & & \\ & & & R_{l_1}(p_1,q_1) & & & \\ & & & & \ddots & & \\ & & & & & R_{l_t}(p_t,q_t) \end{pmatrix};$$

dabei ist jeweils

$$R_l(p,q) = \begin{pmatrix} p & q & & & & & 0 \\ -q & p & \ddots & & & & \\ 0 & 1 & \ddots & \ddots & & & \\ 0 & 0 & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & 0 & 1 & p & q \\ 0 & & & 0 & 0 & -q & p \end{pmatrix} \in \mathbb{R}^{2l\times 2l}$$

Wegen $(X-p)^2 + q^2 = X^2 - 2p + p^2 + q^2$ hat $R_l(p,q)$ das charakteristische Polynom $(X^2 - 2pX + p^2 + q^2)^l$. Die Matrizen $R_{l_i}(p_i,q_i)$ entsprechen also den quadratischen Faktoren in der Faktorisierung des charakteristischen Polynoms von A.

18 Polynome

Definition 18.1. Ein ***Ring*** ist ein Tripel $(R,+,\cdot)$, das aus einer Menge R und Verknüpfungen $+,\cdot$ auf R mit folgenden Eigenschaften besteht:

(i) $(R,+)$ ist abelsche Gruppe.

(ii) $(ab)c = a(bc)$ für $a,b,c \in R$ (Assoziativgesetz)

(iii) $a(b+c) = ab + ac$ für $a,b,c \in R$ (Distributivgesetz)

(iv) $(a+b)c = ac + bc$ für $a,b,c \in R$ (Distributivgesetz)

(v) Es existiert ein Element $1 \in R$ mit $1a = a = a1$ für $a \in R$.

Bemerkung 18.1. (i) Wie üblich zeigt man, dass das Element 1 eindeutig bestimmt ist; man nennt es das ***Einselement*** von R und schreibt $1 = 1_R$.

(ii) Analog enthält R genau ein Element 0 (***Nullelement***) mit $a + 0 = a$ für $a \in R$. Ferner existiert zu jedem $a \in R$ genau ein $-a \in R$ (***negatives*** Element) mit $a + (-a) = 0$. Dabei ist $-(-a) = a$ und $0a = 0 = a0$ und $a(-b) = -(ab) = (-a)b$ für $a,b \in R$. Statt $a + (-b)$ schreibt man auch $a - b$.

(iii) Gilt $ab = ba$ für alle $a, b \in R$, so nennt man R einen ***kommutativen*** Ring.

Beispiel 18.1. (a) Jeder Körper K ist ein kommutativer Ring.

(b) $\mathbb{Z}$ ist ein kommutativer Ring, aber kein Körper.
$\mathbb{N}$ und $\mathbb{N}_0$ sind keine Ringe.

(c) Für $n \in \mathbb{N}$ ist $K^{n\times n}$ ein Ring (***Matrixring***) mit der üblichen Addition und Multiplikation von Matrizen. Für $n \geq 2$ ist $K^{n\times n}$ nicht kommutativ.

(d) $\{0\}$ ist ein Ring.

Definition 18.2. Sei K ein Körper. Ein ***Polynom*** mit Koeffizienten in K ist eine Folge $\varphi = (a_0, a_1, a_2, \dots)$ von Elementen $a_i \in K$ mit $|\{i \in \mathbb{N}_0 : a_i \neq 0\}| < \infty$. Es ist leicht zu sehen, dass diese Polynome einen K-Vektorraum P mit

$$\varphi + \psi := (a_0 + b_0, a_1 + b_1, a_2 + b_2, \dots)$$
$$r\varphi := (ra_0, ra_1, ra_2, \dots)$$

für $\varphi = (a_0, a_1, a_2, \dots), \psi = (b_0, b_1, b_2, \dots) \in P, r \in K$ bilden. Wir definieren eine Multiplikation auf P durch $\varphi\psi := (c_0, c_1, c_2, \dots)$ mit

$$c_0 := a_0b_0, c_1 := a_0b_1 + a_1b_0, \dots, c_i := \sum_{j+k=i} a_jb_k, \dots$$

Satz 18.2. *So wird P zu einem kommutativen Ring mit Nullelement* $(0, 0, 0, \dots)$ *und Einselement* $(1, 0, 0, \dots)$ *Dabei gilt:*

$$r(\varphi\psi) = (r\varphi)\psi = \varphi(r\psi) \qquad (r \in K, \varphi, \psi \in P).$$

Beweis. Seien $\varphi, \psi \in P$ wie oben. Dann ist auch $\varphi\psi \in P$ wegen

$$\{i \in \mathbb{N}_0 : c_i \neq 0\} \subseteq \{j \in \mathbb{N}_0 : a_j \neq 0\} + \{k \in \mathbb{N}_0 : b_k \neq 0\}.$$

Wir rechnen nur das Assoziativgesetz der Multiplikation nach; die übrigen Rechenregeln beweist man analog. Dazu sei $\omega = (c_0, c_1, c_2, \dots) \in P$. Dann ist $\varphi\psi = (d_0, d_1, d_2, \dots)$ mit $d_i = \sum_{j+k=i} a_jb_k$ für alle i, also $(\varphi\psi)\omega = (e_0, e_1, e_2, \dots)$ mit

$$e_i = \sum_{j+k=i} d_jc_k = \sum_{j+k=i} \sum_{l+m=j} a_lb_mc_k = \sum_{l+m+k=i} a_lb_mc_k$$

für alle i. Analog ist $\varphi(\psi\omega) = (f_0, f_1, f_2, \dots)$ mit

$$f_i = \sum_{l+m+k=i} a_lb_mc_k = e_i.$$

□

Bemerkung 18.2. Man nennt P den ***Polynomring*** über K und $(0,0,0,\dots)$ das ***Nullpolynom***, $(1,0,0,\dots)$ das ***Einspolynom***. Ferner nennt man $X := (0,1,0,0,\dots)$ die ***Unbestimmte*** oder ***Variable*** von P.
Für $\varphi = (a_0, a_1, a_2, \dots) \in P$ ist $\varphi X = (0, a_0, a_1, a_2, \dots)$; insbesondere ist $X^2 = (0,0,1,0,0,\dots)$, $X^3 = (0,0,0,1,0,0,\dots)$ usw. Wegen $|\{i \in \mathbb{N}_0 : a_i \neq 0\}| < \infty$ ist also $\varphi = \sum_{i=0}^{\infty} a_i X^i$; dabei ist wie üblich X^0 das Einspolynom. Daher kann man jedes Element in P in der Form $\varphi = \sum_{i=0}^{\infty} a_i X^i$ mit eindeutig bestimmten Koeffizienten $a_i \in K$ schreiben, von denen nur endlich viele von 0 verschieden sind. Dies werden wir in Zukunft stets tun.
Für $\varphi = \sum_{i=0}^{\infty} a_i X^i, \psi = \sum_{i=0}^{\infty} b_i X^i \in P$ und $r \in K$ gilt dann:

$$\varphi = \psi \Leftrightarrow a_i = b_i \text{ für alle } i.$$

$$\varphi + \psi = \sum_{i=0}^{\infty} (a_i + b_i) X^i,$$

$$r\varphi = \sum_{i=0}^{\infty} (r a_i) X^i,$$

$$\varphi\psi = \sum_{i=0}^{\infty} \left(\sum_{j+k=i} a_j b_k \right) X^i.$$

Statt P schreibt man i. Allg. $K[X]$. Ist $0 \neq \varphi = \sum_{i=0}^{\infty} a_i X^i \in K[X]$, so nennt man

$$d := \deg \varphi := \max \{i \in \mathbb{N}_0 : a_i \neq 0\}$$

den ***Grad*** von φ. Dann ist $\varphi = \sum_{i=0}^{d} a_i X^i$. Das Nullpolynom erhält den Grad $-\infty$.

Satz 18.3. *Seien K ein Körper, $\varphi, \psi \in K[X]$ und $0 \neq r \in K$. Dann gilt:*

(i) $\deg(r\varphi) = \deg \varphi$.

(ii) $\deg(\varphi + \psi) \leq \max\{\deg \varphi, \deg \psi\}$.

(iii) $\deg \varphi \neq \deg \psi \Rightarrow \deg(\varphi + \psi) = \max\{\deg \varphi, \deg \psi\}$.

(iv) $\deg(\varphi\psi) = \deg \varphi + \deg \psi$.

Beweis. Als Muster beweisen wir (iv); die übrigen Aussagen zeigt man analog. Wir schreiben $\varphi = \sum_{i=0}^{\infty} a_i X^i, \psi = \sum_{i=0}^{\infty} b_i X^i, \varphi\psi = \sum_{i=0}^{\infty} c_i X^i$ mit $a_i, b_i, c_i \in K$ und setzen $d := \deg \varphi, e := \deg \psi$. Dabei können wir $d \neq -\infty \neq e$ annehmen. Für $i \in \mathbb{N}_0$ ist $c_i = \sum_{j+k=i} a_j b_k$. Wegen $a_j = 0$ für $j > d$ und $b_k = 0$ für $k > e$ ist

$$c_{d+e} = \sum_{j+k=d+e} a_j b_k = \sum_{\substack{j+k=d+e \\ j \leq d, k \leq e}} a_j b_k = a_d b_e \neq 0$$

und $c_i = 0$ für $i > d + e$. Daher ist $\deg(\varphi\psi) = d + e = \deg \varphi + \deg \psi$. □

Bemerkung 18.3. Für $r, s \in K$ gilt:

$$rX^0 + sX^0 = (r+s)X^0, rX^0 - sX^0 = (r-s)X^0, rX^0 \cdot sX^0 = (r \cdot s)X^0.$$

Daher können wir jeweils r mit rX^0 identifizieren und so K als Teilmenge von $K[X]$ auffassen. Man nennt die Elemente in K auch die ***konstanten Polynome*** in $K[X]$.
Sei $0 \neq \varphi = \sum_{i=0}^{\infty} a_i X^i \in K[X]$ mit $d = \deg \varphi$ und $a_d = 1$. Dann nennt man φ ***normiert***. Außerdem betrachtet man das Nullpolynom als normiert.

Satz 18.4. *(Division mit Rest)*
Seien K ein Körper und $\varphi, \psi \in K[X]$ mit $\psi \neq 0$. Dann existieren eindeutig bestimmte $\kappa, \rho \in K[X]$ mit $\varphi = \kappa\psi + \rho$ und $\deg \rho < \deg \psi$.

Definition 18.4. Man nennt κ den ***Quotienten*** und ρ den ***Rest*** bei der Division von φ durch ψ. Im Fall $\rho = 0$ schreibt man auch $\kappa = \frac{\varphi}{\psi}$.

Beweis. (I) Eindeutigkeit: Für $i = 1, 2$ sei $\varphi = \kappa_i\psi + \rho_i$ und $\deg \rho_i < \deg \psi$. Dann ist $\rho_2 - \rho_1 = (\kappa_1 - \kappa_2)\psi$. Im Fall $\kappa_1 \neq \kappa_2$ hätte man den Widerspruch:

$$\deg \psi > \deg(\rho_2 - \rho_1) = \deg(\kappa_1 - \kappa_2)\psi = \deg(\kappa_1 - \kappa_2) + \deg \psi \geq \deg \psi.$$

Also ist $\kappa_1 = \kappa_2$ und $\rho_1 = \varphi - \kappa_1\psi = \varphi - \kappa_2\psi = \rho_2$.

(II) Existenz: Im Fall $\deg \varphi < \deg \psi$ setzt man $\kappa := 0$ und $\rho := \varphi$. Sei also $m := \deg \varphi \geq \deg \psi =: n$. Wir schreiben $\varphi = \sum_{i=0}^{m} a_i X^i$, $\psi = \sum_{j=0}^{n} b_j X^j$ und setzen $\tilde{\varphi} := \varphi - \frac{a_m}{b_n} X^{m-n}\psi$. Dann ist $\deg \tilde{\varphi} < \deg \varphi$. Argumentiert man mit Induktion nach m, so kann man voraussetzen, dass $\tilde{\kappa}, \tilde{\rho} \in K[X]$ mit $\tilde{\varphi} = \tilde{\kappa}\psi + \tilde{\rho}$ und $\deg \tilde{\rho} < n$ existieren. Dann ist aber

$$\varphi = \tilde{\varphi} + \frac{a_m}{b_n} X^{m-n}\psi = \underbrace{\left(\tilde{\kappa} + \frac{a_m}{b_n} X^{m-n}\right)}_{=:\kappa}\psi + \underbrace{\tilde{\rho}}_{=:\rho},$$

wie gewünscht.

□

Beispiel 18.4.

$$\begin{array}{llll|l}
2X^7 + X^6 & +2X^5 & +2X^3 + 2X^2 & +2X + 1 = (X^4 + X^3 + 2X^2 + X + 1) & (2X^3 - X^2 - X + 1) \\
2X^7 + 2X^6 & +4X^5 + 2X^4 & +2X^3 & & +(2X^3 + 2X^2 + 2X) \\
\quad - X^6 & -2X^5 - 2X^4 & \quad + 2X^2 & +2X + 1 & \\
\quad - X^6 & \quad -X^5 - 2X^4 & \quad -X^3 - X^2 & & \\
 & \quad -X^5 & \quad +X^3 + 3X^2 & +2X + 1 & \\
 & \quad -X^5 - X^4 & \quad -2X^3 - X^2 & \quad -X & \\
 & \quad X^4 & \quad +3X^3 + 4X^2 & +3X + 1 & \\
 & \quad X^4 & \quad +X^3 + 2X^2 & \quad +X + 1 & \\
 & & \quad 2X^3 + 2X^2 & +2X &
\end{array}$$

Definition 18.5. Seien K ein Körper und $\varphi, \psi \in K[X]$. Man nennt φ einen ***Teiler*** von ψ und schreibt $\varphi|\psi$, falls $\psi = \varphi\omega$ für ein $\omega \in K[X]$ ist.

Satz 18.5. *Es gelten die folgenden Rechenregeln:*

(i) $\varphi|0, 1|\varphi, \varphi|\varphi$

(ii) $\varphi|\psi \wedge \psi|\omega \Rightarrow \varphi|\omega$

(iii) $\varphi|\psi \Rightarrow c\varphi|d\psi$ *für* $c, d \in K\backslash\{0\}$

(iv) $\varphi|\psi \wedge \psi|\varphi \Rightarrow \exists c \in K\backslash\{0\} : \varphi = c\psi$

(v) $0|\varphi \Leftrightarrow \varphi = 0$

(vi) $\varphi|\psi \wedge \varphi|\omega \Rightarrow \varphi|\alpha\psi + \beta\omega$ *für* $\alpha, \beta \in K[X]$.

Beweis. Als Muster beweisen wir (iv). (Der Rest geht ähnlich.) Sei also $\varphi|\psi$ und $\psi|\varphi$. Dann existieren $\chi, \omega \in K[X]$ mit $\psi = \varphi\omega$ und $\varphi = \psi\chi$. Dann ist $\varphi = \varphi\omega\chi$, d.h. $0 = \varphi(1 - \omega\chi)$. Im Fall $\varphi = 0$ ist auch $\psi = 0$, d.h. $\varphi = 1\psi$. Sei also $\varphi \neq 0$. Nach Satz 18.3 ist dann $1 - \omega\chi = 0$, d.h. $\omega\chi = 1$. Folglich ist $0 = \deg 1 = \deg\omega + \deg\chi$, also $\deg\omega = 0 = \deg\chi$, d.h. $\omega, \chi \in K\backslash\{0\}$. Es ist also $\varphi = c\psi$ mit $c := \chi \in K\backslash\{0\}$. □

Definition 18.6. Seien K ein Körper und $\varphi_1, \ldots, \varphi_n \in K[X]$. Ein Element $\tau \in K[X]$ mit $\tau|\varphi_1, \ldots, \tau|\varphi_n$ nennt man einen ***gemeinsamen Teiler*** von $\varphi_1, \ldots, \varphi_n$. Mit $\mathrm{gT}(\varphi_1, \ldots, \varphi_n)$ bezeichnen wir die Menge aller gemeinsamen Teiler von $\varphi_1, \ldots, \varphi_n$. Ein normiertes Polynom $\delta \in \mathrm{gT}(\varphi_1, \ldots, \varphi_n)$ nennt man ***größten gemeinsamen Teiler*** von $\varphi_1, \ldots, \varphi_n$, falls $\tau|\delta$ für alle $\tau \in \mathrm{gT}(\varphi_1, \ldots, \varphi_n)$ gilt.

Bemerkung 18.6. Sind δ_1, δ_2 größte gemeinsame Teiler von $\varphi_1, \ldots, \varphi_n$, so gilt: $\delta_1|\delta_2|\delta_1$. Nach Satz 18.5 (iv) existieren also ein $c \in K\backslash\{0\}$ mit $\delta_2 = c\delta_1$. Da δ_1 und δ_2 normiert sind, folgt $\delta_1 = \delta_2$. Das bedeutet, dass $\varphi_1, \ldots, \varphi_n$ höchstens einen größten gemeinsamen Teiler δ haben. Man schreibt $\delta = \mathrm{ggT}(\varphi_1, \ldots, \varphi_n)$. Wir untersuchen im Folgenden die Existenz von ggT's.

Satz 18.6. *(Erweiterter Euklidischer Algorithmus, Euklid 365-300)*
Seien K ein Körper und $\alpha, \beta \in K[X]$. *Wir setzen zunächst*

$$(\lambda_0, \mu_0, \nu_0) := (1, 0, \alpha),$$
$$(\lambda_1, \mu_1, \nu_1) := (0, 1, \beta),$$

und $i = 1$. *Im Fall* $\nu_i = 0$ *brechen wir ab. Im Fall* $\nu_i \neq 0$ *liefert die Division mit Rest Polynome* $\kappa_i, \rho_i \in K[X]$ *mit*

$$\nu_{i-1} = \kappa_i\nu_i + \rho_i \text{ und } \deg\rho_i < \deg\nu_i.$$

Wir setzen dann

$$(\lambda_{i+1}, \mu_{i+1}, \nu_{i+1}) := (\lambda_{i-1} - \kappa_i\lambda_i, \mu_{i-1} - \kappa_i\mu_i, \underbrace{\nu_{i-1} - \kappa_i\nu_i}_{=\rho}),$$

erhöhen i um 1 und wiederholen diesen Schritt.
Dieses Verfahren bricht ab, und am Ende existiert ein $c \in K$ mit

$$cv_{i-1} = \mathrm{ggT}(\alpha, \beta) = \lambda_{i-1}\alpha + \mu_{i-1}\beta.$$

Beispiel 18.6. $\alpha = X^4 + X^3 + 2X^2 + X + 1, \beta = X^4 - X^3 + 2X^2 - X + 1$

λ_i	μ_i	ν_i	κ_i
1	0	$X^4 + X^3 + 2X^2 + X + 1$	
0	1	$X^4 - X^3 + 2X^2 - X + 1$	1
1	-1	$2X^3 + 2X$	$\frac{1}{2}X - \frac{1}{2}$
$-\frac{1}{2}X + \frac{1}{2}$	$\frac{1}{2}X + \frac{1}{2}$	$X^2 + 1$	
		0	

Also $X^2 + 1 = \mathrm{ggT}(\alpha, \beta) = (-\frac{1}{2}X + \frac{1}{2})\alpha + (\frac{1}{2}X + \frac{1}{2})\beta$. (Probe!)

Beweis. Wegen $\deg\beta = \deg\nu_1 > \deg\nu_2 > \deg\nu_3 > \dots$ bricht das Verfahren ab. Wir behaupten, dass für alle i gilt:

$$(\star) \qquad \lambda_i\alpha + \mu_i\beta = \nu_i.$$

Für $i = 0, 1$ ist dies sicher richtig. Ferner gilt stets:

$$\begin{aligned}\lambda_{i+1}\alpha + \mu_{i+1}\beta &= (\lambda_{i-1} - \kappa_i\lambda_i)\alpha + (\mu_{i-1} - \kappa_i\mu_i)\beta \\ &= \lambda_{i-1}\alpha + \mu_{i-1}\beta - \kappa_i(\lambda_i\alpha + \mu_i\beta) = \nu_{i-1} - \kappa_i\nu_i = \nu_{i+1}.\end{aligned}$$

Wir behaupten außerdem, dass für alle i gilt:

$$(\star\star) \qquad \mathrm{ggT}(\nu_{i-1}, \nu_i) = \mathrm{ggT}(\alpha, \beta).$$

Für $i = 1$ ist dies sicher richtig. Ferner gilt stets:

$$\mathrm{ggT}(\nu_i, \nu_{i+1}) = \mathrm{ggT}(\nu_i, \nu_{i-1} - \kappa_i\nu_i) = \mathrm{ggT}(\nu_i, \nu_{i-1}) = \mathrm{ggT}(\nu_{i-1}, \nu_i).$$

Damit ist auch $(\star\star)$ bewiesen. Am Ende ist $\nu_i = 0$, d.h.

$$\mathrm{ggT}(\alpha, \beta) = \mathrm{ggT}(\nu_{i-1}, 0) = d\nu_{i-1} \qquad \text{für ein } d \in K.$$

□

Bemerkung 18.7. Aus Satz 18.6 folgt, dass je zwei Polynome in $K[X]$ einen ggT haben. Daraus folgt leicht, dass endlich viele Polynome $\varphi_1, \dots, \varphi_n \in K[X]$ stets einen ggT haben. Genauer gilt:

$$\mathrm{ggT}(\varphi_1, \dots, \varphi_n) = \mathrm{ggT}(\varphi_1, \mathrm{ggT}(\varphi_2, \dots, \varphi_n)).$$

Satz 18.7. *Seien K ein Körper und $\varphi_1, \dots, \varphi_n, \psi \in K[X]$. Genau dann existieren Polynome $\xi_1, \dots, \xi_n \in K[X]$ mit $\xi_1\varphi_1 + \dots + \xi_n\varphi_n = \psi$, wenn $\mathrm{ggT}(\varphi_1, \dots, \varphi_n) | \psi$ gilt.*

Beweis. "⇒": Seien $\xi_1, \ldots, \xi_n \in K[X]$ mit $\psi = \xi_1\varphi_1 + \cdots + \xi_n\varphi_n$. Dann ist jeder gemeinsame Teiler von $\varphi_1, \ldots, \varphi_n$ auch ein Teiler von ψ. Insbesondere gilt: $\operatorname{ggT}(\varphi_1, \ldots, \varphi_n) | \psi$.

"⇐": Sei $\delta := \operatorname{ggT}(\varphi_1, \ldots, \varphi_n) | \psi$, etwa $\psi = \delta\gamma$ mit $\gamma \in K[X]$. Es genügt zu zeigen, dass $\eta_1, \ldots, \eta_n \in K[X]$ mit $\delta = \eta_1\varphi_1 + \cdots + \eta_n\varphi_n$ existieren; denn dann ist

$$\psi = \gamma\delta = \underbrace{(\gamma\eta_1)}_{=:\xi_1}\varphi_1 + \cdots + \underbrace{(\gamma\eta_n)}_{=:\xi_n}\varphi_n.$$

Die Existenz von $\eta_1, \ldots, \eta_n$ zeigen wir induktiv. Im Fall $n = 1$ ist $\delta = \operatorname{ggT}(\varphi_1) = c\varphi_1$ für ein $c \in K$. Im Fall $n = 2$ folgt die Existenz von η_1, η_2 aus Satz 18.6. Im Fall $n > 2$ existieren nach Induktion Polynome $\zeta_2, \ldots, \zeta_n$ mit $\operatorname{ggT}(\varphi_2, \ldots, \varphi_n) = \zeta_2\varphi_2 + \cdots + \zeta_n\varphi_n$. Ferner existieren Polynome $\omega_1, \omega_2 \in K[X]$ mit

$$\begin{aligned}\delta &= \operatorname{ggT}(\varphi_1, \operatorname{ggT}(\varphi_2, \ldots, \varphi_n)) \\ &= \omega_1\varphi_1 + \omega_2 \operatorname{ggT}(\varphi_2, \ldots, \varphi_n) \\ &= \underbrace{\omega_1}_{=\eta_1}\varphi_1 + \underbrace{(\omega_2\zeta_2)}_{=\eta_2}\varphi_2 + \cdots + \underbrace{(\omega_n\zeta_n)}_{=\eta_n}\varphi_n.\end{aligned}$$

□

Satz 18.8. *Seien K ein Körper und $\varphi_1, \ldots, \varphi_n \in K[X]$. Dann gilt:*

$$\operatorname{ggT}(\varphi_1, \ldots, \varphi_n) = 1 \Leftrightarrow \exists \xi_1, \ldots, \xi_n \in K[X] : \xi_1\varphi_1 + \cdots + \xi_n\varphi_n = 1.$$

Definition 18.8. Ggf. nennt man $\varphi_1, \ldots, \varphi_n$ ***teilerfremd***

Beweis. "⇒": Satz 18.7

"⇐": Seien $\xi_1, \ldots, \xi_n \in K[X]$ mit $\xi_1\varphi_1 + \cdots + \xi_n\varphi_n = 1$. Nach Satz 18.7 ist dann $\operatorname{ggT}(\varphi_1, \ldots, \varphi_n) | 1$. Daraus folgt $\operatorname{ggT}(\varphi_1, \ldots, \varphi_n) = 1$.

□

Definition 18.9. Sei K ein Körper. Ein normiertes nichtkonstantes Polynom $\pi \in K[X]$ nennt man ***irreduzibel***, falls π keine Teiler τ mit $0 < \deg\tau < \deg\pi$ hat.

Satz 18.9. *(i) Sei $\pi \in K[X]$ irreduzibel und seien $\alpha, \beta \in K[X]$ mit $\pi | \alpha\beta$. Dann gilt $\pi | \alpha$ oder $\pi | \beta$.*

(ii) Jedes Polynom $\varphi \in K[X] \backslash K$ besitzt mindestens einen irreduziblen Teiler.

Beweis. (i) Im Fall $\pi | \alpha$ sind wir fertig. Sei also $\pi \nmid \alpha$. Dann existieren $\lambda, \mu \in K[X]$ mit $1 = \operatorname{ggT}(\pi, \alpha) = \lambda\pi + \mu\alpha$. Folglich gilt: $\pi | \lambda\pi\beta + \mu\alpha\beta = \beta$.

(ii) Sei $\varphi \in K[X]\backslash K$. Dann ist

$$D := \{\tau \in K[X]\backslash K : \tau \text{ normiert}, \ \tau|\varphi\} \neq \emptyset.$$

Sei $\pi \in D$ so gewählt, dass $\deg \pi$ möglichst klein ist. Dann ist π irreduzibel, denn sonst gäbe es einen Teiler δ von π mit $0 < \deg \delta < \deg \pi$. Wir können annehmen, dass δ normiert ist. Wegen $\delta|\pi|\varphi$ ist dann $\delta \in D$ im Widerspruch zur Wahl von π.

□

Beispiel 18.9. (i) Normierte Polynome vom Grad 1 sind stets irreduzibel.

(ii) Das Polynom $X^2 + 1$ ist irreduzibel in $\mathbb{R}[X]$, aber reduzibel in $\mathbb{C}[X] : X^2 + 1 = (X - i)(X + i)$.

Satz 18.10. *(Eindeutige Primfaktorzerlegung)*
Seien K ein Körper und $\varphi \in K[X]\backslash K$. Dann existieren bis auf die Reihenfolge eindeutig bestimmte irreduzible Polynome $\pi_1, \ldots, \pi_r \in K[X]$ und eine eindeutig bestimmte Konstante $c \in K$ mit $\varphi = c\pi_1 \cdots \pi_r$.

Beweis. (I) Existenz: (Induktion nach $d := \deg \varphi$)
Im Fall $d = 1$ ist $\varphi = c(X - b)$ mit $b, c \in K$, und die Sache ist klar. Sei also $d > 1$. Nach Satz 18.9 besitzt φ einen irreduziblen Teiler π_1. Wir schreiben $\varphi = \pi_1\psi$. Im Fall $\psi \in K$ sind wir fertig. Im Fall $\psi \notin K$ existieren nach Induktion irreduzible Polynome $\pi_2, \ldots, \pi_r \in K[X]$ und ein $c \in K$ mit $\psi = c\pi_2 \cdots \pi_r$. Folglich ist $\varphi = c\pi_1\pi_2 \cdots \pi_r$.

(II) Eindeutigkeit: Sei $c\pi_1 \cdots \pi_r = d\rho_1 \cdots \rho_s$ mit $c, d \in K$ und irreduziblen Polynomen $\pi_1, \ldots, \pi_r, \rho_1, \ldots, \rho_s \in K[X]$. Dann ist $\pi_1|c\pi_1 \cdots \pi_r = d\rho_1 \cdots \rho_s$, also $\pi_1|\rho_i$ für ein $i \in \{1, \ldots, s\}$. Nach Umnummerierung können wir $\pi_1|\rho_1$ annehmen. Da ρ_1 irreduzibel ist, folgt $\pi_1 = \rho_1$. Also ist $0 = \pi_1(c\pi_2 \cdots \pi_r - d\rho_2 \cdots \rho_s)$, d.h. $c\pi_2 \cdots \pi_r = d\rho_2 \cdots \rho_s$. Der Rest ergibt sich induktiv.

□

19 Minimalpolynom

K Körper

Definition 19.1. Für $\varphi = \sum_{i=0}^{n} a_i X^i \in K[X]$ und $b \in K$ setzt man

$$\varphi(b) := \sum_{i=0}^{n} a_i b^i \in K.$$

Man sagt, dass $\phi(b)$ durch ***Einsetzen*** von b in φ entsteht. Ist $\varphi(b) = 0$, so nennt man b eine ***Nullstelle*** von φ in K.

Bemerkung 19.1. Für $\varphi, \psi \in K[X]$ und $r, b \in K$ gilt offenbar:

$$(\varphi + \psi)(b) = \varphi(b) + \psi(b), (\varphi - \psi)(b) = \varphi(b) - \psi(b), (\varphi \cdot \psi)(b) = \varphi(b) \cdot \psi(b), (r\varphi)(b) = r\varphi(b).$$

Satz 19.1. *Für $\varphi \in K[X]$ und $b \in K$ ist $\varphi(b)$ der Rest bei der Division von φ durch $(X - b)$. Insbesondere gilt: $\varphi(b) = 0 \Leftrightarrow X - b | \varphi$.*

Beweis. Division mit Rest liefert $\kappa, \rho \in K[X]$ mit $\varphi = \kappa(X - b) + \rho$; dabei ist $\rho \in K$. Aus der obigen Bemerkung folgt: $\varphi(b) = \kappa(b)(b - b) + \rho = \rho$. Daher gilt die erste Behauptung. Im Fall $\varphi(b) = 0$ ist also $\varphi = \kappa(X - b)$. Umgekehrt gilt im Fall $\varphi = \kappa(X - b) : \varphi(b) = \kappa(b)(b - b) = 0$. □

Satz 19.2. *Jedes Polynom $\varphi \in K[X]$ vom Grad $n \neq -\infty$ besitzt in K höchstens n Nullstellen.*

Beweis. Seien $a_1, \ldots, a_m$ paarweise verschiedene Nullstellen von φ in K. Dann sind $X - a_1, \ldots, X - a_m$ paarweise verschiedene irreduzible Polynome in $K[X]$. Nach Satz 19.1 ist φ durch $X - a_1, \ldots, X - a_m$ teilbar, also nach Satz 18.10 auch durch $(X - a_1) \cdots (X - a_m)$; insbesondere ist $n = \deg \varphi \geq m$. □

Definition 19.3. Für $\varphi = \sum_{i=0}^{d} a_i X^i \in K[X]$ und $A \in K^{n \times n}$ setzt man

$$\varphi(A) := a_d A^d + a_{d-1} A^{d-1} + \cdots + a_1 A + a_0 1_n \in K^{n \times n}.$$

Analog setzt man für jeden K-Vektorraum V und $f \in \operatorname{End}(V)$:

$$\varphi(f) := a_d f^d + a_{d-1} f^{d-1} + \cdots + a_1 f + a_0 \operatorname{id}_V \in \operatorname{End}(V).$$

Satz 19.3. *(Satz von Cayley-Hamilton)*
Für $A \in K^{n \times n}$ gilt: $\chi_A(A) = 0$.

Beweis. Wir schreiben $\chi_A = a_0 + a_1 X + \cdots + a_n X^n$ mit $a_0, \ldots, a_{n-1}, a_n \in K$ und bezeichnen die Adjunkte von $X1_n - A \in K[X]$ mit $\widetilde{X1_n - A}$. Die Koeffizienten von $\widetilde{X1_n - A}$ sind Polynome vom Höchstgrad $n - 1$. Wir schreiben $\widetilde{X1_n - A} = C_0 + C_1 X + \cdots + C_n X^{n-1}$ mit $C_0, \ldots, C_{n-1} \in K^{n \times n}$. Dann gilt:

$$\begin{aligned} a_0 1_n + a_1 1_n X + \cdots + a_n 1_n X^n &= \chi_A \cdot 1_n = |X1_n - A| \cdot 1_n = (X1_n - A)(\widetilde{X1_n - A}) \\ &= (X1_n - A)(C_0 + C_1 X + \cdots + C_{n-1} X^{n-1}) \\ &= -AC_0 + (C_0 - AC_1)X + \cdots + (C_{n-2} - AC_{n-1})X^{n-1} + C_{n-1} X^n. \end{aligned}$$

Koeffizienten-Vergleich liefert:

$$a_0 1_n = -AC_0, a_1 1_n = C_0 - AC_1, \ldots, a_{n-1} 1_n = C_{n-2} - AC_{n-1}, a_n 1_n = C_{n-1}.$$

Daher gilt:

$$\begin{aligned} \chi_A(A) &= a_0 1_n + a_1 A + \cdots + a_n A^n \\ &= -AC_0 + A(C_0 - AC_1) + \cdots + A^{n-1}(C_{n-2} - AC_{n-1}) + A^n C_{n-1} \\ &= 0 \end{aligned}$$

□

Beispiel 19.3. Für $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in K^{2\times 2}$ gilt:

$$\begin{aligned} |X \cdot 1_n - A| &= \begin{vmatrix} X-a & -b \\ -c & X-d \end{vmatrix} = (X-a)(X-d) - bc \\ &= X^2 - (a+d)X + (ad-bc) = X^2 - \operatorname{spur}(A)X + \det(A) \end{aligned}$$

Also gilt:

$$A^2 - \operatorname{spur}(A) \cdot A + \det(A)1_2 = 0. \qquad \text{(Probe!)}$$

Bemerkung 19.3. Nach dem Satz von Cayley-Hamilton sind für $A \in K^{n\times n}$ die Potenzen $A^0 = 1_n, A^1 = A, A^2, \ldots, A^n$ linear abhängig. Sei $m \in \mathbb{N}$ minimal mit der Eigenschaft, dass $A^0, A^1, \ldots, A^m$ linear abhängig sind (also $m \leq n$). Dann existieren $c_o, \ldots, c_{m-1} \in K$ mit

$$A^m + c_{m-1}A^{m-1} + \cdots + c_1 A + c_0 1_n = 0$$

Diese sind eindeutig bestimmt; ist nämlich auch

$$A^m + d_{m-1}A^{m-1} + \cdots + a_1 A + d_0 1_n = 0,$$

so ergibt Subtraktion:

$$(c_{m-1} - d_{m-1})A^{m-1} + \cdots + (c_1 - d_1)A + (c_0 - d_0)1_n = 0.$$

Da $A^0, A^1, \ldots, A^{m-1}$ linear unabhängig sind, folgt

$$c_{m-1} = d_{m-1}, \ldots, c_1 = d_1, c_0 = d_0,$$

wie behauptet. Man nennt $X^m + c_{m-1}X^{m-1} + \cdots + c_1 X + c_0$ das ***Minimalpolynom*** von A.

Beispiel 19.4. Sei

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

Dann sind $1_4, A$ linear unabhängig, aber

$$A^2 = \begin{pmatrix} -1 & 2 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ -2 & 2 & 1 & 2 \\ -2 & 2 & 0 & 3 \end{pmatrix} = 2 \cdot A - 1_4.$$

Also hat A das Minimalpolynom $X^2 - 2X + 1 = (X-1)^2$.

Satz 19.4. *Ähnliche Matrizen besitzen das gleiche Minimal-Polynom.*

Beweis. Sei K ein Körper, und seien $A, B \in K^{n\times n}, S \in \mathrm{GL}(n,K)$ mit $B = S^{-1}AS$. Seien außerdem $a_0, \ldots, a_{m-1} \in K$ mit $A^m + a_{m-1}A_{m-1} + \cdots + a_1A + a_0 1_n = 0$. Dann gilt:

$$\begin{aligned} B^m + a_{m-1}B^{m-1} + \cdots + a_1B + a_0 1_n &= (S^{-1}AS)^m + a_{m-1}(S^{-1}AS)^{m-1} + \cdots + a_1(S^{-1}AS) + a_0 S^{-1}1_n S \\ &= S^{-1}\underbrace{(A^m + a_{m-1}A^{m-1} + \cdots + a_1A + a_0 1_n)}_{=0} S = 0. \end{aligned}$$

□

Bemerkung 19.4. Seien V ein endlich-dimensionaler K-Vektorraum und $f \in \mathrm{End}(V)$ mit Matrix A bzgl. irgendeiner Basis von V. Dann nennt man das Minimalpolynom von A auch das ***Minimalpolynom*** von f. (Nach dem obigen Satz hängt dieses nicht von der Wahl der Basis ab.)

Satz 19.5. *Sei $A \in K^{n\times n}$ mit Minimalpolynom μ_A. Für alle $\varphi \in K[X]$ mit $\varphi(A) = 0$ gilt dann $\mu_A | \varphi$. Insbesondere gilt für das charakteristische Polynom χ_A von A: $\mu_A | \chi_A$.*

Beweis. Division mit Rest liefert $\kappa, \rho \in K[X]$ mit $\varphi = \kappa\mu_A + \rho$ und $\deg\rho < \deg\mu_A$. Dann ist $\rho(A) = (\varphi - \kappa\mu_A)(A) = \underbrace{\varphi(A)}_{=0} - \kappa(A)\underbrace{\mu_A(A)}_{=0} = 0$.
Wegen $\deg\rho < \deg\mu_A$ folgt aus der Definition von μ_A: $\rho = 0$, d.h. $\mu_A | \varphi$. Der Rest ergibt sich aus dem Satz von Cayley-Hamilton. □

Bemerkung 19.5. (i) Man kann zeigen, dass umgekehrt gilt: $\chi_A | \mu_A^n$.

(ii) Entsprechende Aussagen gelten für Minimalpolynome von Endomorphismen.

Beispiel 19.5. In Beispiel 19.4 gilt: $\chi_A = (X-1)^4$.

Satz 19.6. *Seien V ein endlich-dimensionaler K-Vektorraum und $f \in \mathrm{End}(V)$ mit Minimalpolynom μ_f. Für $r \in K$ gilt dann:*

$$r \text{ Eigenwert von } f \Leftrightarrow \mu_f(r) = 0$$

Beweis. "$\Rightarrow$" Sei $0 \neq v \in V$ mit $f(v) = rv$, und sei

$$\mu_f = X^m + c_{m-1}X^{m-1} + \cdots + c_1X + c_0$$

Dann gilt:

$$\begin{aligned} 0 &= (f^m + c_{m-1}f^{m-1} + \cdots + c_1 f + c_0 \mathrm{id}_V)(v) \\ &= f^m(v) + c_{m-1}f^{m-1}(v) + \cdots + c_1 f(v) + c_0 v \\ &= r^m v + c_{m-1}r^{m-1}v + \cdots + c_1 rv + c_0 v = \mu_f(r)v. \end{aligned}$$

Wegen $v \neq 0$ folgt $\mu_f(r) = 0$.

"⇐": Sei χ_f das charakteristische Polynom von f. Nach Satz 19.5 existiert ein $\varphi \in K[X]$ mit $\chi_f = \mu_f \varphi$. Daher ist $\chi_f(r) = \mu_f(r)\varphi(r) = 0\varphi(r) = 0$. Folglich ist r Eigenwert von f.

□

Bemerkung 19.6. (i) Eine entsprechende Aussage gilt für Matrizen.

(ii) Für die Berechnung von Eigenwerten kann man also statt des charakteristische Polynoms auch das Minimalpolynom benutzen. Dies vereinfacht die Rechnung manchmal, da das Minimalpolynom nach 19.5 einen kleineren Grad haben kann als das charakteristische Polynom.

Satz 19.7. *Sei V ein endlich-dimensionaler K-Vektorraum. Ein Endomorphismus f von V mit Minimal-Polynom μ ist genau dann diagonalisierbar, wenn μ sich in der Form $\mu = (X - a_1) \cdots (X - a_k)$ mit paarweise verschiedenen $a_1, \ldots, a_k \in K$ schreiben lässt.*

Beweis. "⇒": Sei f diagonalisierbar. Dann existiert eine Basis von V, bzgl. der die Matrix von f die folgende Form hat:

$$A = \begin{pmatrix} a_1 & & & & & & & & 0 \\ & \ddots & & & & & & & \\ & & a_1 & & & & & & \\ & & & a_2 & & & & & \\ & & & & \ddots & & & & \\ & & & & & a_2 & & & \\ & & & & & & \ddots & & \\ & & & & & & & a_k & \\ & & & & & & & & \ddots \\ 0 & & & & & & & & a_k \end{pmatrix}$$

($a_1, \ldots, a_k \in K$ paarweise verschieden); dabei sind $a_1, \ldots, a_k$ die Eigenwerte von A und f. Nach Satz 19.6 sind $a_1, \ldots, a_k$ Nullstellen von μ, d.h. $\mu = (X - a_1) \cdots (X - a_k)\varphi$ für ein $\varphi \in K[X]$. Andererseits ist offensichtlich $(A - a_1 1_n) \cdots (A - a_k 1_n) = 0$. Aus Satz 19.5 folgt also: $\mu = (X - a_1) \cdots (X - a_k)$.

"⇐": Sei $\mu = (X - a_1) \cdots (X - a_k)$ mit paarweise verschiedenen $a_1, \ldots, a_k \in K$. Für $i = 1, \ldots, k$ hat das Polynom

$$\varphi_i := \prod_{j \neq i} \frac{X - a_j}{a_i - a_j}$$

den Grad $k - 1$ und es gilt:

$$\varphi_i(a_t) = \begin{cases} 0 & \text{für } t \neq i \\ 1 & \text{für } t = i. \end{cases}$$

Daher hat das Polynom $\varphi := \varphi_1 + \cdots + \varphi_k$ höchstens Grad $k-1$, und es gilt $\varphi(a_1) = \cdots = \varphi(a_k) = 1$. Nach Satz 19.2 ist φ das konstante Polynom 1. Daher ist

$$\mathrm{id}_V = \sum_{i=1}^{k} \prod_{j \neq i} \frac{f - a_j \,\mathrm{id}_V}{a_i - a_j}.$$

Für $v \in V$ gilt also:

$$v = \sum_{i=1}^{k} \underbrace{\prod_{j \neq i} \left(\frac{f - a_j \,\mathrm{id}_V}{a_i - a_j} \right)(v)}_{=:v_i}.$$

Dabei ist jeweils

$$(f - a_i \cdot \mathrm{id}_V)(v_i) = \left(\prod_{j \neq i} \frac{1}{a_i - a_j} \right) \underbrace{(f - a_1 \,\mathrm{id}_V) \cdots (f - a_k \,\mathrm{id}_V)}_{=0}(v) = 0,$$

d.h. $v_i \in \mathrm{Ker}\,(f - a_i \cdot \mathrm{id}_V)$. Daher ist V die Summe der Eigenräume zu den Eigenwerten $a_1, \ldots, a_k$ von f. Folglich ist f diagonalisierbar.

□

Bemerkung 19.7. Eine entsprechende Aussage gilt für Matrizen.

Beispiel 19.7. Sei

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

Dann sind 1_4 und A linear unabhängig, aber es ist

$$A^2 = \begin{pmatrix} 1 & -3 & 0 & 3 \\ 0 & 1 & 0 & 0 \\ -3 & -3 & 4 & 3 \\ 0 & -3 & 0 & 4 \end{pmatrix} = 3 \cdot A - 2 \cdot 1_4.$$

Daher ist $X^2 - 3X + 2 = (X-1)(X-2)$ das Minimalpolynom von A. Also ist A diagonalisierbar. (Probe!)

Satz 19.8. *Seien K algebraisch abgeschlossen, $n \in \{2, 3\}$ und $A, B \in K^{n \times n}$. Genau dann sind A und B ähnlich, wenn sie das gleiche charakteristische und das gleiche Minimalpolynom haben.*

Beweis. "$\Rightarrow$": Satz 13.5 und Satz 19.4.

"⇐": (I) $n = 2$: A habe das charakteristische Polynom $\chi = (X-a)(X-b)$ und das Minimalpolynom μ. Im Fall $a \neq b$ ist A zu $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ ähnlich. Sei also $a = b$, d.h. $\chi = (X-a)^2$. Im Fall $\mu = (X-a)^2$ ist A zu $\begin{pmatrix} a & 0 \\ 1 & a \end{pmatrix}$ ähnlich, und im Fall $\mu = X-a$ ist $A = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$.

(II) $n = 3$: A habe das charakteristische Polynom $(X-a)(X-b)(X-c)$ und das Minimalpolynom μ. Im Fall $|\{a,b,c\}| = 3$ ist A zu $\begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$ ähnlich.

Sei jetzt $\chi = (X-a)^2(X-b)$ mit $a \neq b$. Im Fall $\mu = (X-a)(X-b)$ ist A zu $\begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{pmatrix}$ ähnlich, im Fall $\mu = (X-a)^2(X-b)$ zu $\left(\begin{array}{cc|c} a & 0 & 0 \\ 1 & a & 0 \\ \hline 0 & 0 & b \end{array}\right)$.

Schließlich sei $\chi = (X-a)^3$. Im Fall $\mu = X-a$ ist $A = \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix}$, im Fall $\mu = (X-a)^2$ ist A zu $\left(\begin{array}{cc|c} a & 0 & 0 \\ 1 & a & 0 \\ \hline 0 & 0 & a \end{array}\right)$ ähnlich und im Fall $\mu = (X-a)^3$ zu $\begin{pmatrix} a & 0 & 0 \\ 1 & a & 0 \\ 0 & 1 & a \end{pmatrix}$. □

Bemerkung 19.8. Was passiert im Fall $n = 4$?

Bemerkung 19.9. (Die Exponentialfunktion mit Matrizen als Argument)
Im Folgenden identifizieren wir $\mathbb{C}^{n\times n}$ mit $\mathbb{C}^{n^2}$. Man kann zeigen (Analysis), dass für $A \in \mathbb{C}^{n\times n}$ die Reihe

$$e^A := \exp(A) := \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

in $\mathbb{C}^{n\times n}$ konvergiert. Für $A, B \in \mathbb{C}^{n\times n}$ ist i.Allg. $e^{A+B} \neq e^A e^B$. Gilt aber $AB = BA$, so ist auch (ohne Beweis)

$$e^{A+B} = e^A e^B.$$

Ist

$$A = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix}$$

mit quadratischen Matrizen $A_1, \ldots, A_m$, so ist, wie man sich leicht überlegt,

$$e^A = \begin{pmatrix} e^{A_1} & & 0 \\ & \ddots & \\ 0 & & e^{A_m} \end{pmatrix}.$$

Ist insbesondere

$$D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_m \end{pmatrix}$$

eine Diagonalmatrix, so ist

$$e^D = \begin{pmatrix} e^{d_1} & & 0 \\ & \ddots & \\ 0 & & e^{d_m} \end{pmatrix}.$$

Für

$$J_n := \begin{pmatrix} 0 & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & 0 \end{pmatrix} \in \mathbb{C}^{n\times n}$$

ist $J_n^n = 0$, d.h.

$$e^{J_n} = 1_n + J_n + \frac{1}{2}J_n^2 + \frac{1}{6}J_n^3 + \cdots + \frac{1}{(n-1)!}J_n^{n-1}$$
$$= \begin{pmatrix} 1 & & & & & 0 \\ 1 & \ddots & & & & \\ \frac{1}{2} & \ddots & \ddots & & & \\ \frac{1}{6} & \ddots & \ddots & \ddots & & \\ \vdots & \ddots & \ddots & \ddots & \ddots & \\ \frac{1}{(n-1)!} & \cdots & \frac{1}{6} & \frac{1}{2} & 1 & 1 \end{pmatrix}$$

Für $A \in \mathbb{C}^{n\times n}$ und $S \in \mathrm{GL}(n, \mathbb{C})$ ist schließlich

$$e^{S^{-1}AS} = S^{-1}e^A S.$$

Mit diesen Regeln kann man e^A leicht berechnen.

Beispiel 19.9. Sei

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

Die Methoden von Kapitel 17 liefern eine Matrix

$$S = \begin{pmatrix} 1 & 3 & -1 \\ 1 & 2 & -3 \\ 0 & 0 & -1 \end{pmatrix} \in \mathrm{GL}(3, \mathbb{C})$$

mit

$$S^{-1}AS = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & 2 & 0 \\ 0 & 1 & 2 \end{array}\right) =: J.$$

Wir setzen

$$D := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix},\ N := \left(\begin{array}{c|cc} 0 & 0 & 0 \\ \hline 0 & 0 & 0 \\ 0 & 1 & 0 \end{array}\right).$$

Dann ist $J = D + N$ mit $DN = ND$ und

$$e^D = \left(\begin{array}{c|cc} e & 0 & 0 \\ \hline 0 & e^2 & 0 \\ 0 & 0 & e^2 \end{array}\right), e^N = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & 1 & 0 \\ 0 & 1 & 1 \end{array}\right), e^J = e^{D+N} = e^D e^N = \left(\begin{array}{c|cc} e & 0 & 0 \\ \hline 0 & e^2 & 0 \\ 0 & e^2 & e^2 \end{array}\right),$$

$$e^A = e^{SJS^{-1}} = Se^JS^{-1} = \cdots = \begin{pmatrix} -2e + 2e^2 & 3e - 2e^2 & -7e + 5e^2 \\ -2e - e^2 & 3e + e^2 & -7e + e^2 \\ -e^2 & e^2 & -e^2 \end{pmatrix}$$

Beispiel 19.10. Gegeben sei die Differentialgleichung

$$(\star) \qquad x'(t) = ax(t).$$

Dabei sei $a \in \mathbb{C}$ vorgegeben. Gesucht wird eine differenzierbare Funktion

$$x : \mathbb{R} \longrightarrow \mathbb{C}, t \longmapsto x(t),$$

die $(\star)$ erfüllt. Bekanntlich hat die allgemeine Lösung von $(\star)$ die Form

$$x(t) = c \cdot e^{at} \qquad (c \in \mathbb{C}).$$

Wir wollen dies verallgemeinern.

Bemerkung 19.10. (Systeme linearer Differentialgleichungen)
Gegeben sei das folgende System linearer Differentialgleichungen:

$$(\star) \qquad \begin{cases} x_1'(t) = a_{11}x_1(t) + \cdots + a_{1n}x_n(t) \\ \vdots \\ x_n'(t) = a_{n1}x_1(t) + \cdots + a_{nn}x_n(t). \end{cases}$$

Dabei ist $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ vorgegeben. Gesucht werden differenzierbare Funktionen

$$x_i : \mathbb{R} \longrightarrow \mathbb{C}, t \longmapsto x_i(t), \qquad (i = 1, \ldots, n)$$

die $(\star)$ simultan erfüllen. Wir setzen $x := (x_1, \ldots, x_n)$ und schreiben $(\star)$ auch in der Form

$$x'(t)^T = Ax(t)^T.$$

In der Analysis zeigt man, dass die allgemeine Lösung von ($\star$) gegeben wird durch

$$x(t)^T = e^{At}v^T \qquad (v \in \mathbb{C}^n).$$

Die Lösungen bilden einen n-dimensionalen $\mathbb{C}$-Vektorraum $\mathcal{L}$.
Als konkretes Beispiel betrachten wir das folgende System linearer Differentialgleichungen:

$$(\star) \qquad \begin{cases} x_1'(t) = 13x_1(t) - 4x_2(t) \\ x_2'(t) = -4x_1(t) + 7x_2(t). \end{cases}$$

Wir setzen

$$A := \begin{pmatrix} 13 & -4 \\ -4 & 7 \end{pmatrix}.$$

Die Methoden von Kapitel 17 liefern $A = SDS^{-1}$ mit

$$D := \begin{pmatrix} 5 & 0 \\ 0 & 15 \end{pmatrix},\ S := \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}.$$

Also ist

$$e^{At} = e^{SDtS^{-1}} = Se^{Dt}S^{-1} = S\begin{pmatrix} e^{5t} & 0 \\ 0 & e^{15t} \end{pmatrix}S^{-1}$$
$$= \cdots = \frac{1}{5}\begin{pmatrix} e^{5t} + 4e^{15t} & 2e^{5t} - 2e^{15t} \\ 2e^{5t} - 2e^{15t} & 4e^{5t} + e^{15t} \end{pmatrix}$$

Die allgemeine Lösung von ($\star$) hat also die Form

$$\begin{aligned} x_1(t) &= \frac{a}{5}(e^{5t} + 4e^{15t}) + \frac{b}{5}(2e^{5t} - 2e^{15t}) \\ &= \frac{a+2b}{5}e^{5t} + \frac{4a-2b}{5}e^{15t} \\ x_2(t) &= \frac{a}{5}(2e^{5t} - 2e^{15t}) + \frac{b}{5}(4e^{5t} + e^{15t}) \\ &= \frac{2a+4b}{5}e^{5t} + \frac{-2a+b}{5}e^{15t}, \end{aligned}$$

mit $a, b \in \mathbb{C}$. (Probe!)

20 Der Dualraum

K Körper

Definition 20.1. Für jeden K-Vektorraum V nennt man $V^\star := \mathrm{Hom}_K(V, K) := \{\lambda : V \longrightarrow K | \lambda \text{ linear}\}$ den ***Dualraum*** von V. Seine Elemente heißen ***Linearformen*** auf V.

Bemerkung 20.1. (i) Bekanntlich ist $V^\star$ ein K-Vektorraum; für $\lambda, \mu \in V^\star$ und $a \in K$ sind dabei $\lambda + \mu \in V^\star$ und $a\mu \in V^\star$ definiert durch

$$(\lambda + \mu)(v) = \lambda(v) + \mu(v) \text{ und } (a\mu)(v) = a\mu(v)$$

für $v \in V$.

(ii) Sei V endlich-dimensional mit Basis $b_1, \ldots, b_n$. Für $i = 1, \ldots, n$ existiert dann genau ein $\beta_i \in V^\star$ mit $\beta_i(b_j) = \delta_{ij}$ für $j = 1, \ldots, n$; dabei ist

$$\delta_{ij} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases},$$

das Kronecker-Symbol. Die so definierten Linearformen $\beta_1, \ldots, \beta_n$ auf V sind linear unabhängig; sind nämlich $a_1, \ldots, a_n \in K$ mit $0 = \sum_{i=1}^n a_i\beta_i$, so gilt für $j = 1, \ldots, n$:

$$0 = \left(\sum_{i=1}^n a_i\beta_i\right)(b_j) = \sum_{i=1}^n a_i \underbrace{\beta_i(b_j)}_{=\delta_{ij}} = a_j.$$

Die Elemente $\beta_1, \ldots, \beta_n$ bilden sogar eine Basis von $V^\star$; denn für $\lambda \in V^\star$ und $j = 1, \ldots, n$ gilt:

$$\left(\sum_{i=1}^n \lambda(b_i)\beta_i\right)(b_j) = \sum_{i=1}^n \lambda(b_i) \underbrace{\beta_i(b_j)}_{=\delta_{ij}} = \lambda(b_j),$$

d.h. $\lambda = \sum_{i=1}^n \lambda(b_i)\beta_i$.
Man nennt $\beta_1, \ldots, \beta_n$ die zu $b_1, \ldots, b_n$ ***duale Basis***. Insbesondere gilt also im Fall $\dim V < \infty$:

$$\dim V^\star = \dim V$$

(iii) Man nennt $V^{\star\star} := (V^\star)^\star$ den ***Bidualraum*** von V Für $v \in V$ ist die Abbildung $f_v : V^\star \longrightarrow K, \lambda \longmapsto \lambda(v)$ linear; denn für $\lambda, \mu \in V^\star$ und $a, b \in K$ ist

$$f_v(a\lambda + b\mu) = (a\lambda + b\mu)(v) = a\lambda(v) + b\mu(v) = af_v(\lambda) + bf_v(\mu).$$

Daher ist also $f_v \in V^{\star\star}$. Für $v, w \in V$ und $a, b \in K$ ist dabei

$$f_{av+bw} = af_v + bf_w;$$

für $\lambda \in V^\star$ gilt nämlich:

$$f_{av+bw}(\lambda) = \lambda(av + bw) = a\lambda(v) + b\lambda(w) = af_v(\lambda) + bf_w(\lambda) = (af_v + bf_w)(\lambda).$$

Folglich ist die Abbildung

$$f : V \longrightarrow V^{\star\star}, v \longmapsto f_v$$

linear. Im Fall $\dim V < \infty$ ist f injektiv; ist nämlich $v \in V$ mit $v \neq 0$, so kann man v zu einer Basis $b_1 = v, b_2, \ldots, b_n$ von V ergänzen. Nimmt man dann die dazu duale Basis $\beta_1, \ldots, \beta_n$ von $V^\star$, so ist $1 = \beta_1(b_1) = \beta_1(v) = f_v(\beta_1)$, d.h. $f_v \neq 0$.
Wegen $\dim V^{\star\star} = \dim V^\star = \dim V$ ist also die lineare Abbildung $f : V \longrightarrow V^{\star\star}$ sogar bijektiv.

Satz 20.1. *Sei U ein Untervektorraum eines endlich-dimensionalen K-Vektorraums V. Dann ist*

$$U^\perp := \{\lambda \in V^\star : \lambda | U = 0\}$$

ein Untervektorraum von $V^\star$ mit $\dim U + \dim U^\perp = \dim V$.

Beweis. Wir wählen eine Basis $b_1, \ldots, b_m$ von U und ergänzen diese zu einer Basis $b_1, \ldots, b_n$ von V. Die dazu duale Basis von $V^\star$ bezeichnen wir mit $\beta_1, \ldots, \beta_n$. Wie oben gezeigt, ist dann $\lambda = \sum_{i=1}^n \lambda(b_i)\beta_i$ für $\lambda \in V^\star$. Im Fall $\lambda \in U^\perp$ ist $\lambda(b_1) = \cdots = \lambda(b_m) = 0$, also $\lambda = \sum_{i=m+1}^n \lambda(b_i)\beta_i$. Umgekehrt verschwinden $\beta_{m+1}, \ldots, \beta_n$ auf $b_1, \ldots, b_m$, also auch auf U. Daher verschwindet jede Linearkombination von $\beta_{m+1}, \ldots, \beta_n$ auf U. Daher ist $U^\perp$ gerade der von $\beta_{m+1}, \ldots, \beta_n$ aufgespannte Untervektorraum von $V^\star$, und die Behauptung folgt. □

Bemerkung 20.2. Es seien U_1, U_2 Untervektorräume eines endlich-dimensionalen K-Vektorraums V. Mit obigen Bezeichnungen gilt offenbar:

$$(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp.$$

Ferner ist offenbar $U_1^\perp + U_2^\perp \subseteq (U_1 \cap U_2)^\perp$. Wegen

$$\begin{aligned}
\dim (U_1^\perp + U_2^\perp) &= \dim U_1^\perp + \dim U_2^\perp - \dim (U_1^\perp \cap U_2^\perp) \\
&= \dim V - \dim U_1 + \dim V - \dim U_2 - \dim (U_1 + U_2)^\perp \\
&= \dim V - \dim U_1 - \dim U_2 + \dim (U_1 + U_2) \\
&= \dim V - \dim (U_1 \cap U_2) = \dim (U_1 \cap U_2)^\perp.
\end{aligned}$$

folgt also:

$$(U_1 \cap U_2)^\perp = U_1^\perp + U_2^\perp.$$

Bemerkung 20.3. Gegeben seien K-Vektorräume V, W und eine lineare Abbildung $f : V \longrightarrow W$. Für $\mu \in W^\star$ ist dann $\mu \circ f : V \longrightarrow W \longrightarrow K$ linear, d.h. $\mu \circ f \in V^\star$. Man nennt

$$f^\star : W^\star \longrightarrow V^\star, \mu \longmapsto \mu \circ f$$

die zu f ***duale Abbildung***. Sie ist ebenfalls linear; für $\lambda, \mu \in W^\star$ und $a, b \in K$ gilt nämlich:

$$f^\star(a\lambda + b\mu) = (a\lambda + b\mu) \circ f = a(\lambda \circ f) + b(\mu \circ f) = af^\star(\lambda) + bf^\star(\mu).$$

Hat man einen weiteren K-Vektorraum X und eine lineare Abbildung $g : W \longrightarrow X$, so gilt für die lineare Abbildung $g \circ f : V \longrightarrow X$:

$$(g \circ f)^\star = f^\star \circ g^\star \qquad \text{(Achtung!)};$$

denn für $\nu \in X^\star$ ist

$$(g \circ f)^\star(\nu) = \nu \circ (g \circ f) = (\nu \circ g) \circ f = f^\star(\nu \circ g) = f^\star(g^\star(\nu)) = (f^\star \circ g^\star)(\nu).$$

Offenbar ist $(\mathrm{id}_V)^\star = \mathrm{id}_{V^\star}$; denn für $\lambda \in V^\star$ gilt:

$$(\mathrm{id}_V)^\star(\lambda) = \lambda \circ \mathrm{id}_V = \lambda.$$

Ist also f bijektiv, so ist

$$\mathrm{id}_{V^\star} = (\mathrm{id}_V)^\star = (f^{-1} \circ f)^\star = f^\star \circ (f^{-1})^\star$$

und analog $\mathrm{id}_{W^\star} = (f^{-1})^\star \circ f^\star$. Daher ist mit f auch $f^\star$ bijektiv, und es gilt:

$$(f^\star)^{-1} = (f^{-1})^\star.$$

Für lineare Abbildungen $f_1, f_2 : V \longrightarrow W$ und Elemente $a_1, a_2 \in K$ gilt ferner:

$$(a_1 f_1 + a_2 f_2)^\star = a_1 f_1^\star + a_2 f_2^\star;$$

denn für alle $\mu \in W^\star$ gilt:

$$\begin{aligned}(a_1 f_1 + a_2 f_2)^\star(\mu) &= \mu \circ (a_1 f_1 + a_2 f_2) = a_1(\mu \circ f_1) + a_2(\mu \circ f_2) \\ &= a_1 f_1^\star(\mu) + a_2 f_2^\star(\mu) = (a_1 f_1^\star + a_2 f_2^\star)(\mu).\end{aligned}$$

Dies zeigt, dass die Abbildung

$$\mathrm{Hom}_K(V, W) \longrightarrow \mathrm{Hom}_K(W^\star, V^\star), f \longmapsto f^\star$$

linear ist; insbesondere gilt für die Nullabbildung $0_{V,W} : V \longrightarrow W$:

$$0_{V,W}^\star = 0_{W^\star, V^\star}$$

Satz 20.3. *Gegeben seien endlich-dimensionale K-Vektorräume V, W und eine lineare Abbildung $f : V \longrightarrow W$. Für die duale Abbildung $f^\star : W^\star \longrightarrow V^\star$ gilt dann:*

(i) $\mathrm{Ker}(f^\star) = (\mathrm{Bld}\, f)^\perp$.

(ii) $\mathrm{Bld}(f^\star) = (\mathrm{Ker}\, f)^\perp$.

Beweis. (i) Für $\mu \in W^\star$ gilt:
$\mu \in \mathrm{Ker}\, f^\star \Leftrightarrow 0 = f^\star(\mu) = \mu \circ f \Leftrightarrow \mu(f(v)) = 0$ für alle $v \in V \Leftrightarrow \mu|\mathrm{Bld}\, f = 0 \Leftrightarrow \mu \in (\mathrm{Bld}\, f)^\perp$.

(ii) Sei $\lambda \in \mathrm{Bld}(f^\star)$, also $\lambda = f^\star(\mu) = \mu \circ f$ für ein $\mu \in W^\star$. Dann ist $\mathrm{Ker}\, f \subseteq \mathrm{Ker}\, \mu \circ f = \mathrm{Ker}\, \lambda$, d.h. $\lambda|\mathrm{Ker}\, f = 0$ und $\lambda \in (\mathrm{Ker}\, f)^\perp$. Daher ist $\mathrm{Bld}(f^\star) \subseteq (\mathrm{Ker}\, f)^\perp$. Andererseits ist

$$\begin{aligned}\dim \mathrm{Bld}(f^\star) &= \dim W^\star - \dim \mathrm{Ker}(f^\star) \overset{(i)}{=} \dim W - \dim(\mathrm{Bld}\, f)^\perp \\ &\overset{20.1}{=} \dim \mathrm{Bld}\, f = \dim V - \dim \mathrm{Ker}\, f \overset{20.1}{=} \dim(\mathrm{Ker}\, f)^\perp.\end{aligned}$$

Daher gilt: $\mathrm{Bld}(f^\star) = (\mathrm{Ker}\, f)^\perp$.

□

Satz 20.4. *Gegeben seien endlich-dimensionale K-Vektorräume V, W mit Basen $b_1, \dots, b_m$ bzw. $c_1, \dots, c_n$ und eine lineare Abbildung $f : V \longrightarrow W$ mit Matrix $A = (a_{ij})$ bzgl. $b_1, \dots, b_m$ und $c_1, \dots, c_n$. Dann ist A^T die Matrix der dualen Abbildung $f^\star : W^\star \longrightarrow V^\star$ bzgl. der dualen Basen $\gamma_1, \dots, \gamma_n$ von $W^\star$ und $\beta_1, \dots, \beta_m$ von $V^\star$.*

Beweis. Wir schreiben $\underbrace{f^\star(\gamma_j)}_{\gamma_j \circ f} = \sum_{i=1}^n a'_{ij}\beta_i$ mit $a'_{ij} \in K$ für alle i, j. Dann gilt für $k = 1, \dots, n$:

$$\begin{aligned} a'_{kj} &= \sum_{i=1}^n a'_{ij} \underbrace{\beta_i(b_k)}_{=\delta_{ik}} = \left(\sum_{i=1}^n a'_{ij}\beta_i\right)(b_k)(\gamma_j \circ f)(b_k) \\ &= \gamma_j\left(\sum_{i=1}^n a_{ik}c_i\right) = \sum_{i=1}^n a_{ik} \underbrace{\gamma_j(c_i)}_{=\delta_{ij}} = a_{jk}. \end{aligned}$$

□

21 Bilineare Abbildungen

K Körper

Definition 21.1. Seien U, V, W K-Vektorräume. Eine Abbildung $\beta : U \times V \longrightarrow W$ nennt man ***bilinear***, falls für alle $a, a' \in K, u, u' \in U, v, v' \in V$ gilt:

(i) $\beta(au + a'u', v) = a\beta(u, v) + a'\beta(u', v)$;

(ii) $\beta(u, av + a'v') = a\beta(u, v) + a'\beta(u, v')$.

Bemerkung 21.1. (i) Die Nullabbildung $U \times V \longrightarrow W$ ist stets bilinear.

(ii) Für bilineare Abbildungen $\beta, \beta' : U \times V \longrightarrow W$ und Elemente $b, b' \in K$ ist auch

$$b\beta + b'\beta' : U \times V \longrightarrow W, (u, v) \longmapsto b\beta(u, v) + b'\beta'(u, v)$$

bilinear; für alle $a, a' \in K, u, u' \in U, v, v' \in V$ ist nämlich

$$\begin{aligned} (b\beta + b'\beta')(au + a'u', v) &= b\beta(au + a'u', v) + b'\beta'(au + a'u', v) \\ &= b[a\beta(u, v) + a'\beta(u', v)] + b'[a\beta'(u, v) + a'\beta'(u', v)] \\ &= a[b\beta(u, v) + b'\beta'(u, v)] + a'[b\beta(u', v) + b'\beta'(u', v)] \\ &= a(b\beta + b'\beta')(u, v) + a'(b\beta + b'\beta')(u', v) \end{aligned}$$

und analog

$$(b\beta + b'\beta')(u, av + a'v') = \cdots = a(b\beta + b'\beta')(u, v) + a'(b\beta + b'\beta')(u, v').$$

(iii) Nach (i) und (ii) ist die Menge $\mathrm{Bil}(U,V;W)$ aller bilinearen Abbildungen $\beta : U \times V \longrightarrow W$ ein Untervektorraum von $\mathrm{Abb}(U \times V, W)$. Wie groß ist $\mathrm{Bil}(U \times V; W)$?

(iv) Für $\beta \in \mathrm{Bil}(U,V;W)$ und $u \in U$ ist die Abbildung

$$g_\beta(u) : V \longrightarrow W, v \longmapsto \beta(u,v)$$

linear; für $a, a' \in K, v, v' \in V$ gilt nämlich:

$$\begin{aligned}[g_\beta(u)](av + a'v') &= \beta(u, av + a'v') = a\beta(u,v) + a'\beta(u,v') \\ &= a[g_\beta(u)](v) + a'[g_\beta(u)](v').\end{aligned}$$

Insbesondere ist $0 = [g_\beta(u)](0) = \beta(u,0)$.

(v) Für $\beta \in \mathrm{Bil}(U,V;W)$ ist die Abbildung

$$g_\beta : U \longrightarrow \mathrm{Hom}(V,W), u \longmapsto g_\beta(u)$$

linear; für $a, a' \in K, u, u' \in U, v \in V$ gilt nämlich:

$$\begin{aligned}[g_\beta(au + a'u')](v) &= \beta(au + a'u', v) = a\beta(u,v) + a'\beta(u',v) \\ &= a[g_\beta(u)](v) + a'[g_\beta(u')](v) = [ag_\beta(u) + a'g_\beta(u')](v),\end{aligned}$$

d.h.

$$g_\beta(au + a'u') = ag_\beta(u) + a'g_\beta(u').$$

Insbesondere ist $g_\beta(0) = 0$, d.h. $0 = [g_\beta(0)](v) = \beta(0,v)$ für alle $v \in V$.

(vi) Die Abbildung

$$G : \mathrm{Bil}(U,V;W) \longrightarrow \mathrm{Hom}(U, \mathrm{Hom}(V,W)), \beta \longmapsto g_\beta$$

ist linear; für alle $b, b' \in K, \beta, \beta' \in \mathrm{Bil}(U,V;W)$ gilt nämlich:

$$G(b\beta + b'\beta') = g_{b\beta+b'\beta'} \stackrel{!}{=} bg_\beta + b'g_{\beta'} = bG(\beta) + b'G(\beta');$$

denn für $u \in U$ ist

$$g_{b\beta+b'\beta'}(u) \stackrel{!}{=} (bg_\beta + b'g_{\beta'})(u) = bg_\beta(u) + b'g_{\beta'}(u)$$

wegen

$$\begin{aligned}[g_{b\beta+b'\beta'}(u)](v) &= (b\beta + b'\beta')(u,v) = b\beta(u,v) + b'\beta'(u,v) \\ &= b[g_\beta(u)](v) + b'[g_{\beta'}(u)](v) = [bg_\beta(u) + b'g_{\beta'}(u)](v)\end{aligned}$$

für $v \in V$.

(vii) Ferner ist G injektiv; denn ist $\beta \in \mathrm{Bil}(U, V; W)$ mit $0 = G(\beta) = g_\beta$, so ist $0 = g_\beta(u)$ für alle $u \in U$, d.h. $0 = [g_\beta(u)](v) = \beta(u, v)$ für alle $v \in V$. Folglich ist $\beta = 0$.

(viii) Schließlich ist G auch surjektiv; zum Beweis sei $h \in \mathrm{Hom}(U, \mathrm{Hom}(V, W))$. Dann ist die Abbildung

$$\beta : U \times V \longrightarrow W, (u, v) \longmapsto [h(u)](v)$$

bilinear; für $a, a' \in K, u, u' \in U, v, v' \in V$ ist nämlich

$$\begin{aligned}\beta(au + a'u', v) &= [h(au + a'u')](v) = [ah(u) + a'h(u')](v) \\ &= a[h(u)](v) + a'[h(u')](v) = a\beta(u, v) + a'\beta(u', v)\end{aligned}$$

und

$$\begin{aligned}\beta(u, av + a'v') &= [h(u)](av + a'v') = a[h(u)](v) + a'[h(u)](v') \\ &= a\beta(u, v) + a'\beta(u, v').\end{aligned}$$

Ferner ist $G(\beta) = g_\beta \overset{!}{=} h$; denn für $u \in U$ ist $g_\beta(u) = h(u)$ wegen $[g_\beta(u)](v) = \beta(u, v) = [h(u)](v)$ für $v \in V$.

(ix) Wir haben also einen Vektorraum-Isomorphismus

$$G : \mathrm{Bil}(U, V; W) \longrightarrow \mathrm{Hom}(U, \mathrm{Hom}(V, W));$$

insbesondere ist

$$\begin{aligned}\dim \mathrm{Bil}(U, V; W) &= \dim \mathrm{Hom}(U, \mathrm{Hom}(V, W)) \\ &= (\dim U)(\dim \mathrm{Hom}(V, W)) = (\dim U)(\dim V)(\dim W).\end{aligned}$$

Beispiel 21.1. Für $m, n, p \in \mathbb{N}$ ist die Matrixmultiplikation

$$\beta : K^{m\times n} \times K^{n\times p} \longrightarrow K^{m\times p}, (A, B) \longmapsto AB$$

bilinear.

Definition 21.2. Für K-Vektorräume U, V nennt man die Elemente in $\mathrm{Bil}(U, V; K)$ ***Bilinearformen***. Ist $\beta \in \mathrm{Bil}(U, V; K)$ und sind U, V endlich-dimensional mit Basen $a_1, \ldots, a_m$ bzw. $b_1, \ldots, b_n$, so nennt man

$$B := (\beta(a_i, b_j)) \in K^{m\times n}$$

die ***Matrix*** von β bzgl. $a_1, \ldots, a_m$ und $b_1, \ldots, b_n$.

Satz 21.2. *Für endlich-dimensionale K-Vektorräume U, V mit Basen $a_1, \ldots, a_m$ bzw. $b_1, \ldots, b_n$ ist die Abbildung*

$$F : \mathrm{Bil}(U, V; K) \longrightarrow K^{m\times n},$$

die jeder Bilinearform $\beta \in \mathrm{Bil}(U, V; K)$ ihre Matrix bzgl. $a_1, \ldots, a_m$ und $b_1, \ldots, b_n$ zuordnet, ein Vektorraum-Isomorphismus.

Beweis. Seien $r, r' \in K$ und $\beta, \beta' \in \mathrm{Bil}(U, V; K)$. Für $i = 1, \ldots, m$ und $j = 1, \ldots, n$ ist dann

$$(r\beta + r'\beta')(a_i, b_j) = r\beta(a_i, b_j) + r'\beta'(a_i, b_j).$$

Daher ist $F(r\beta + r'\beta') = rF(\beta) + r'F(\beta')$. Folglich ist F linear. Ist $F(\beta) = 0$, so ist $\beta(a_i, b_j) = 0$ für $i = 1, \ldots, m$ und $j = 1, \ldots, n$. Sind $u \in U, v \in V$ beliebig und schreibt man $u = \sum_{i=1}^{m} r_i a_i$, $v = \sum_{j=1}^{n} s_j b_j$ $(r_i, s_j \in K)$, so ist

$$\beta(u, v) = \sum_{i=1}^{m} \sum_{j=1}^{n} r_i s_j \beta(a_i, b_j) = 0.$$

Also ist $\beta = 0$. Damit ist gezeigt: $\mathrm{Ker}\, F = \{0\}$, d.h. F ist injektiv. Wegen $\dim \mathrm{Bil}(U, V; K) = (\dim U)(\dim V) = mn = K^{m \times n}$ ist F auch surjektiv. □

Bemerkung 21.2. Seien $a'_1, \ldots, a'_m$ und $b'_1, \ldots, b'_n$ weitere Basen von U bzw. V. Wir schreiben $a'_i = \sum_{k=1}^{m} r_{ki} a_k$ und $b'_j = \sum_{l=1}^{n} s_{lj} b_l$ $(r_{ki}, s_{lj} \in K)$. Für $\beta \in \mathrm{Bil}(U, V; K)$ und $i = 1, \ldots, m$, $j = 1, \ldots, n$ gilt dann:

$$\beta(a'_i, b'_j) = \sum_{k=1}^{m} \sum_{l=1}^{n} r_{ki} s_{lj} \beta(a_k, b_l) = \sum_{k=1}^{m} \sum_{l=1}^{n} r_{ki} \beta(a_k, b_l) s_{lj}.$$

Sind also B und B' Matrizen von β bzgl. $a_1, \ldots, a_m$ und $b_1, \ldots, b_n$ bzw. $a'_1, \ldots, a'_m$ und $b'_1, \ldots, b'_n$, so ist

$$B' = R^T B S.$$

mit $R := (r_{ij}) \in \mathrm{GL}(m, K)$ und $S := (s_{ij}) \in \mathrm{GL}(n, K)$; insbesondere ist $\mathrm{rg}\, B = \mathrm{rg}\, B'$. Man bezeichnet $\mathrm{rg}\, B$ auch als ***Rang*** von β und schreibt $\mathrm{rg}\, B =: \mathrm{rg}\, \beta$. Wie in LA I gezeigt, kann man Basen von U und V so wählen, dass die Matrix von β bzgl. dieser Basis die folgende Form hat:

$$\begin{pmatrix} 1_t & 0 \\ 0 & 0 \end{pmatrix} \qquad (t = \mathrm{rg}\, \beta).$$

Beispiel 21.2. (i) Für $A = (a_{ij}) \in K^{m \times n}$ ist die Abbildung

$$\beta : K^{m \times 1} \times K^{n \times 1} \longrightarrow K, (x, y) \longmapsto x^T A y$$

bilinear; denn für $a, a' \in K, x, x' \in K^{m \times 1}, y, y' \in K^{n \times 1}$ gilt:

$$\begin{aligned} \beta(ax + a'x', y) &= (ax + a'x')^T A y = (ax^T + a'(x')^T) A y \\ &= ax^T A y + a'(x') A y = a\beta(x, y) + a'\beta(x', y). \end{aligned}$$

und

$$\beta(x, ay + a'y') = \cdots = a\beta(x, y) + a'\beta(x, y').$$

Was ist die Matrix von β bzgl. der Standardbasen von $K^{m\times 1}$ und $K^{n\times 1}$? Für $i = 1, \ldots, m$ und $j = 1, \ldots, n$ ist

$$\beta(e_i, e_j) = e_i^T A e_j = (0, \ldots, 0, 1, 0, \ldots, 0) \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$= (a_{i1}, \ldots, a_{in}) \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = a_{ij}.$$

Daher ist A selbst die Matrix von β bzgl. der Standardbasen.

(ii) Sei V ein beliebiger K-Vektorraum und $V^\star$ der Dualraum von V. Dann ist die Abbildung

$$\beta : V \times V^\star \longrightarrow K, (v, f) \longmapsto f(v)$$

bilinear; für $a, a' \in K, v, v' \in V, f, f' \in V^\star$ gilt nämlich:

$$\beta(av + a'v', f) = f(av + a'v') = af(v) + a'f(v') = a\beta(v, f) + a'\beta(v', f)$$

und

$$\beta(v, af + a'f') = (af + a'f')(v) = af(v) + a'f'(v) = a\beta(v, f) + a'\beta(v, f').$$

Sei jetzt V endlich-dimensional mit Basis $b_1, \ldots, b_m$. Wir bezeichnen die dazu duale Basis von $V^\star$ mit $f_1, \ldots, f_m$. Für $i, j = 1, \ldots, m$ ist dann

$$\beta(b_i, f_j) = f_j(b_i) = \delta_{ij}.$$

Die Matrix von β bzgl. $b_1, \ldots, b_m$ und $f_1, \ldots, f_m$ ist also die Einheitsmatrix.

(iii) Die Abbildung

$$\beta : \mathbb{R}^2 \times \mathbb{R}^3 \longrightarrow \mathbb{R}, (x, y) \longmapsto x_1y_1 + 2x_1y_2 + 3x_2y_3$$

$x = (x_1, x_2) \in \mathbb{R}^2, y = (y_1, y_2, y_3) \in \mathbb{R}^3$ ist bilinear, wie man leicht nachrechnet. Die Matrix von β bzgl. der Basis $(1, 1), (1, -1)$ von $\mathbb{R}^2$ und der Basis $(1, 1, 0), (1, 0, 1), (0, 1, 1)$ von $\mathbb{R}^3$ ist

$$B = \begin{pmatrix} 3 & 4 & 5 \\ 3 & -2 & -1 \end{pmatrix}.$$

Bemerkung 21.3. Seien V ein K-Vektorraum mit Basis $b_1, \ldots, b_n$ und $\beta : V \times V \longrightarrow K$ bilinear. Dann nennt man $B := (\beta(b_i, b_j)) \in K^{n\times n}$ die ***Matrix*** von β bzgl. $b_1, \ldots, b_n$.
Ist $b'_1, \ldots, b'_n$ eine weitere Basis von V und schreibt man $b'_j = \sum_{i=1}^{n} s_{ij} b_i$ $(s_{ij} \in K)$, so ist $S := (s_{ij}) \in \mathrm{GL}(n, K)$, und $B' = S^T B S$ ist nach Bemerkung 21.2 die Matrix von β bzgl. $b'_1, \ldots, b'_n$. Wir werden später versuchen $b'_1, \ldots, b'_n$ so zu wählen, dass die Matrix von β bzgl. $b'_1, \ldots, b'_n$ möglichst "einfach" wird.

Definition 21.3. Man nennt $A, B \in K^{n\times n}$ ***kongruent***, falls $B = U^T A U$ für ein $U \in \mathrm{GL}(n, K)$ ist.

Satz 21.3. *Für $A, B, C \in K^{n\times n}$ gilt:*

(i) (Reflexivität)
A ist zu A kongruent.

(ii) (Symmetrie)
Ist A zu B kongruent, so auch B zu A.

(iii) (Transitivität)
Ist A zu B und B zu C kongruent, so ist auch A zu C kongruent.

Beweis. Routine. □

Definition 21.4. Seien V ein K-Vektorraum und $\beta : V \times V \longrightarrow K$ bilinear mit $\beta(u, v) = \beta(v, u)$ für alle $u, v \in V$. Dann nennt man β ***symmetrisch***.

Bemerkung 21.4. (i) Sei $b_1, \ldots, b_n$ eine Basis von V. Man zeigt leicht, dass β genau dann symmetrisch ist, wenn $\beta(b_i, b_j) = \beta(b_j, b_i)$ für $i, j = 1, \ldots, n$ gilt, d.h. wenn die Matrix $B = (\beta(b_i, b_j))$ von β bzgl. $b_1, \ldots, b_n$ symmetrisch ist.

(ii) Man nennt $b_1, \ldots, b_n$ eine ***Orthogonalbasis*** von (V, β), falls $\beta(b_i, b_j) = 0$ für alle $i \neq j$ ist. Das bedeutet, dass die Matrix von β bzgl. $b_1, \ldots, b_n$ eine Diagonalmatrix ist.

(iii) Man nennt $b_1, \ldots, b_n$ eine ***Orthonormalbasis*** von (V, β), falls $\beta(b_i, b_j) = \delta_{ij}$ für $i, j = 1, \ldots, n$ ist. Das bedeutet, dass die Matrix von β bzgl. $b_1, \ldots, b_n$ die Einheitsmatrix ist.

Bemerkung 21.5. In LA I hatten wir Körper mit $2, 3, 4$ Elementen konstruiert:

(I)

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

(II)

$+$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$\cdot$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

(III)

$+$	0	1	a	b
0	0	1	a	b
1	1	0	b	a
a	a	b	0	1
b	b	a	1	0

$\cdot$	0	1	a	b
0	0	0	0	0
1	0	1	a	b
a	0	a	b	1
b	0	b	1	a

Dabei gilt:

(I) $1 + 1 = 0$.

(II) $1 + 1 + 1 = 0$.

(III) $1 + 1 = 0$.

Definition 21.5. Die kleinste natürliche Zahl n mit

$$\underbrace{1 + 1 + \cdots + 1}_{n \text{ Summanden}} = 0 \text{ in } K$$

nennt man ***Charakteristik*** von K. Man schreibt $n = \operatorname{char}(K)$. Existiert kein solches $n \in \mathbb{N}$, so sagt man, K hat die Charakteristik 0, und schreibt $\operatorname{char} K = 0$.

Beispiel 21.5. In (I) und (III) ist $\operatorname{char} K = 2$, in (II) ist $\operatorname{char} K = 3$. Außerdem ist $\operatorname{char} \mathbb{Q} = \operatorname{char} \mathbb{R} = \operatorname{char} \mathbb{C} = 0$.

Satz 21.5. *Die Charakteristik von K ist stets 0 oder eine Primzahl.*

Beweis. Sei $n := \operatorname{char} K = ab$ mit $ab \in \mathbb{N}$. Dann ist

$$0 = \underbrace{1 + 1 + \cdots + 1}_{n \text{ Summanden}} = \underbrace{(1 + \cdots + 1)}_{a \text{ Summanden}} \underbrace{(1 + \cdots + 1)}_{b \text{ Summanden}},$$

also

$$0 = \underbrace{1 + \cdots + 1}_{a \text{ Summanden}} \quad \text{oder} \quad 0 = \underbrace{1 + \cdots + 1}_{b \text{ Summanden}}$$

Nach Wahl von n folgt $n = a$ oder $n = b$. □

Satz 21.6. *Sei* $\operatorname{char} K \neq 2$, *$V$ ein endlich-dimensionaler K-Vektorraum und $\beta : V \times V \longrightarrow K$ bilinear und symmetrisch. Dann existiert eine Orthogonalbasis von (V, β).*

Beweis. (Induktion nach $n := \dim V$)
Im Fall $n = 1$ ist nichts zu tun. Sei also $n > 1$. Im Fall $\beta = 0$ ist auch nichts zu tun. Sei also $\beta \neq 0$. Dann existieren $u, v \in V$ mit $a := \beta(u, v) \neq 0$. Wir setzen $w := a^{-1}v$. Dann ist $\beta(u, w) = 1$. Im Fall $\beta(x, x) = 0$ für alle $x \in V$ hätte man den Widerspruch

$$0 = \beta(u + w, u + w) = \underbrace{\beta(u, u)}_{0} + \underbrace{\beta(u, w)}_{1} + \underbrace{\beta(w, u)}_{1} + \underbrace{\beta(w, w)}_{0} = 1 + 1.$$

Also existiert ein $b_1 \in V$ mit $\beta(b_1, b_1) \neq 0$; insbesondere ist $b_1 \neq 0$. Wir ergänzen b_1 zu einer Basis $b_1, c_2, \ldots, c_n$ von V. Für $i = 2, \ldots, n$ sei

$$d_i := c_i - \frac{\beta(b_1, c_i)}{\beta(b_1, b_1)} b_1.$$

Dann ist

$$\beta(b_1, d_i) = \beta(b_1, c_i) - \frac{\beta(b_1, c_i)}{\beta(b_1, b_1)} \beta(b_1, b_1) = 0.$$

Außerdem bilden $b_1, d_2, \ldots, d_n$ eine Basis von V. Wir setzen $U := \operatorname{Span}(d_2, \ldots, d_n)$. Dann ist $\beta(b_1, x) = 0$ für alle $x \in U$. Die Einschränkung

$$\gamma : U \times U \longrightarrow K, (x, y) \longmapsto \beta(x, y)$$

ist auch bilinear und symmetrisch. Nach Induktion existiert eine Orthogonalbasis $b_2, \ldots, b_n$ von (U, γ). Dann bilden $b_1, b_2, \ldots, b_n$ eine Orthogonalbasis von (V, β). □

Satz 21.7. *Sei* $\operatorname{char} K \neq 2$ *und* $B \in K^{n \times n}$ *symmetrisch. Dann ist* B *zu einer Diagonalmatrix kongruent.*

Beweis. Sei $e_1, \ldots, e_n$ die Standardbasis von $V := K^n$. Nach Satz 21.2 existiert genau eine Bilinearform $\beta : V \times V \longrightarrow K$ mit Matrix B bzgl. $e_1, \ldots, e_n$. Mit B ist auch β symmetrisch. Nach Satz 21.6 existiert zu (V, β) eine Orthogonalbasis $b_1, \ldots, b_n$ Wir schreiben $b_j = \sum_{i=1}^{n} s_{ij} e_i$ $(s_{ij} \in K)$. Dann ist $S := (s_{ij}) \in \mathrm{GL}(n, K)$, und $S^T B S$ ist die Matrix von β bzgl. $b_1, \ldots, b_n$, d.h. $S^T B S$ ist eine Diagonalmatrix. □

Beispiel 21.7. Sei

$$B := \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 2 & 3 & 4 \end{pmatrix} \in \mathbb{R}^{3 \times 3}.$$

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

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 & 5 \\ 1 & 0 & 3 \\ 5 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 5 \\ 2 & 0 & 6 \\ 10 & 6 & 8 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 10 \\ 2 & 0 & 12 \\ 10 & 12 & 16 \end{pmatrix}$$

$$\begin{pmatrix}1&0&0\\-1&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}2&2&10\\2&0&12\\10&12&16\end{pmatrix}\begin{pmatrix}1&-1&0\\0&1&0\\0&0&1\end{pmatrix}=\begin{pmatrix}2&2&10\\0&-2&2\\10&12&16\end{pmatrix}\begin{pmatrix}1&-1&0\\0&1&0\\0&0&1\end{pmatrix}=\begin{pmatrix}2&0&10\\0&-2&2\\10&2&16\end{pmatrix}$$
$$\begin{pmatrix}1&0&0\\0&1&0\\-5&0&1\end{pmatrix}\begin{pmatrix}2&0&10\\0&-2&2\\10&2&16\end{pmatrix}\begin{pmatrix}1&0&-5\\0&1&0\\0&0&1\end{pmatrix}=\begin{pmatrix}2&0&10\\0&-2&2\\0&2&-34\end{pmatrix}\begin{pmatrix}1&0&-5\\0&1&0\\0&0&1\end{pmatrix}=\begin{pmatrix}2&0&0\\0&-2&2\\0&2&-34\end{pmatrix}$$
$$\begin{pmatrix}1&0&0\\0&1&0\\0&1&1\end{pmatrix}\begin{pmatrix}2&0&0\\0&-2&2\\0&2&-34\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&1\\0&0&1\end{pmatrix}=\begin{pmatrix}2&0&0\\0&-2&2\\0&0&-32\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&1\\0&0&1\end{pmatrix}=\begin{pmatrix}2&0&0\\0&-2&0\\0&0&-32\end{pmatrix}$$

Insgesamt gilt also:

$$\begin{pmatrix}1&1&0\\-1&1&0\\-6&-4&2\end{pmatrix}\begin{pmatrix}0&1&2\\1&0&3\\2&3&4\end{pmatrix}\begin{pmatrix}1&-1&-6\\1&1&-4\\0&0&2\end{pmatrix}=\begin{pmatrix}2&0&0\\0&-2&0\\0&0&-32\end{pmatrix}$$

Bemerkung 21.7. Die Einträge in der Diagonalmatrix sind in der Regel nicht eindeutig bestimmt, auch nicht bis auf Reihenfolge.

Definition 21.8. Eine ***quadratische Form*** auf einem K-Vektorraum V ist eine Abbildung $q : V \longrightarrow K$ mit folgenden Eigenschaften:

(i) $q(ax) = a^2 q(x)$ für alle $a \in K, x \in V$.

(ii) Die Abbildung $\beta_q : V \times V \longrightarrow K, (x, y) \longmapsto q(x + y) - q(x) - q(y)$ ist bilinear.

Bemerkung 21.8. (a) Die obige Bilinearform β_q auf V ist offensichtlich symmetrisch.

(b) Umgekehrt liefert jede (symmetrische) Bilinearform γ auf V eine quadratische Form

$$q_\gamma : V \longrightarrow K, x \longmapsto \gamma(x, x);$$

denn für $a, a' \in K, x, x', y, y' \in V$ gilt:

$$\begin{aligned}
&q_\gamma(ax) = \gamma(ax, ax) = a^2\gamma(x, x) = a^2 q_\gamma(x),\\
&q_\gamma(ax + a'x' + y) - q_\gamma(ax + a'x') - q_\gamma(y)\\
&= \gamma(ax + a'x' + y, ax + a'x' + y) - \gamma(ax + a'x', ax + a'x') - \gamma(y, y)\\
&= \gamma(ax + a'x', y) + \gamma(y, ax + a'x')\\
&= a\gamma(x, y) + a'\gamma(x', y) + a\gamma(y, x) + a'\gamma(y, x')\\
&= a[\gamma(x, y) + \gamma(y, x)] + a'[\gamma(x', y) + \gamma(y, x')]\\
&= a[\gamma(x + y, x + y) - \gamma(x, x) - \gamma(y, y)] + a'[\gamma(x' + y, x' + y) - \gamma(x', x') - \gamma(y, y)]\\
&= a[q_\gamma(x + y) - q_\gamma(x) - q_\gamma(y)] + a'[q_\gamma(x' + y) - q_\gamma(x') - q_\gamma(y)].
\end{aligned}$$

(c) Sei q eine quadratische Form auf V, β_q die entsprechende Bilinearform auf V und $q' := q_{\beta_q}$ die in (b) definierte quadratische Form. Für $x \in V$ ist dann

$$q'(x) = \beta_q(x, x) = q(x + x) - q(x) - q(x) = 2q(x).$$

(d) Umgekehrt sei γ eine symmetrische Bilinearform auf V, q_γ die in (b) definierte quadratische Form und $\gamma' := \beta_{q_\gamma}$ die entsprechende Bilinearform aus der Definition. Für $x, y \in V$ gilt dann:

$$\begin{aligned}\gamma'(x,y) &= q_\gamma(x+y) - q_\gamma(x) - q_\gamma(y) = \gamma(x+y, x+y) - \gamma(x,x) - \gamma(y,y) \\ &= \gamma(x,y) + \gamma(y,x) = 2\gamma(x,y).\end{aligned}$$

(e) Aus (c) und (d) folgt, dass im Fall $\operatorname{char} K \neq 2$ eine Bijektion zwischen der Menge der symmetrischen Bilinearformen auf V und der Menge der quadratischen Formen auf V existiert. Aussagen und Begriffe für symmetrische Bilinearformen übertragen sich auf diese Weise auf quadratische Formen.

Bemerkung 21.9. Sei $\operatorname{char} K \neq 2$. Wir betrachten eine Abbildung q der Form

$$q : K^n \longrightarrow K, x = (x_1, \ldots, x_n) \longmapsto \sum_{i,j=1}^{n} a_{ij} x_i x_j;$$

dabei seien die Elemente $a_{ij} \in K (i, j = 1, \ldots, n)$ vorgegeben. Indem wir notfalls a_{ij} durch $\frac{a_{ij}+a_{ji}}{2}$ ersetzen, können wir $a_{ij} = a_{ji}$ annehmen. Dann ist $A := (a_{ij}) \in K^{n\times n}$ symmetrisch. Für $x \in K^n$ ist also $q(x) = xAx^T$, d.h. q ist eine quadratische Form; denn q kommt von der Bilinearform $\beta : K^n \times K^n \longrightarrow K, (x, y) \longmapsto xAy^T$ (vgl. Beispiel 21.2). Nach Satz 21.7 existieren eine Diagonalmatrix $D = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} \in K^{n\times n}$ und eine Matrix $S = (s_{ij}) \in \mathrm{GL}(n, K)$ mit $A = S^T DS$. Dann ist

$$q(x) = \underbrace{xS^T}_{y} DSx^T = \sum_{i=1}^{n} d_i y_i^2 \text{ mit } y_i = \sum_{j=1}^{n} s_{ij} x_j$$

für $i = 1, \ldots, n$. Auf diese Weise kann man $q(x)$ als Linearkombination von Quadraten schreiben.

Beispiel 21.9. Wir betrachten die Abbildung $q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ mit

$$q(x) = 2x_1^2 + 2x_1x_2 + 2x_1x_3 - x_2^2 + 2x_2x_3 + 2x_3^2$$

für $x = (x_1, x_2, x_3) \in \mathbb{R}^3$. Es gilt:

$$\begin{aligned} q(x) &= 2(x_1^2 + x_1x_2 + x_1x_3) - x_2^2 + 2x_2x_3 + 2x_3^2 \\ &= 2\underbrace{(x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3)}_{y_1}{}^2 - \frac{1}{2}x_2^2 - \frac{1}{2}x_3^2 - x_2x_3 - x_2^2 + 2x_2x_3 + 2x_3^2 \\ &= 2y_1^2 - \frac{3}{2}x_2^2 + x_2x_3 + \frac{3}{2}x_3^2 = 2y_1^2 - \frac{3}{2}(x_2^2 - \frac{2}{3}x_2x_3) + \frac{3}{2}x_3^2 \\ &= 2y_1^2 - \frac{3}{2}\underbrace{(x_2 - \frac{1}{3}x_3)}_{y_2}{}^2 + \frac{1}{6}x_3^2 + \frac{3}{2}x_3^2 = 2y_1^2 - \frac{3}{2}y_2^2 + \frac{5}{3}\underbrace{x_3^2}_{y_3^2}. \end{aligned}$$

(Probe!)
Dies nennt man ***quadratische Ergänzung***.

Bemerkung 21.10. Sei wieder $\operatorname{char} K \neq 2$. Gegeben sei eine Abbildung der Form

$$f : K^n \longrightarrow K, x = (x_1, \ldots, x_n) \longmapsto \sum_{i,j=1}^{n} a_{ij}x_ix_j + \sum_{k=1}^{n} b_kx_k + c;$$

dabei sind die Elemente $a_{ij}, b_k, c \in K$ vorgegeben. Wie oben können wir $a_{ij} = a_{ji}$ für $i, j = 1, \ldots, n$ annehmen. Wir setzen

$$\begin{aligned} a_{k,n+1} := a_{n+1,k} &:= \frac{b_k}{2} \qquad (k = 1, \ldots, n) \\ a_{n+1,n+1} &:= c. \end{aligned}$$

Dann ist $q : K^{n+1} \longrightarrow K, x = (x_1, \ldots, x_{n+1}) \longmapsto \sum_{i,j=1}^{n+1} a_{ij}x_ix_j$ eine quadratische Form mit $f(x_1, \ldots, x_n) = q(x_1, \ldots, x_n, 1)$ für $x_1, \ldots, x_n \in K$. Wir können also q mit den obigen Methoden behandeln und dann auf f rückschließen.

Beispiel 21.10. Wir betrachten die Abbildung

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}, (x, y) \longmapsto x^2 + 2y^2 + 2xy + 4y + 3.$$

Wir definieren

$$q : \mathbb{R}^3 \longrightarrow \mathbb{R}, (x, y, z) \longmapsto x^2 + 2y^2 + 2xy + 4yz + 3z^2.$$

Dann gilt:

$$\begin{aligned} q(x, y, z) &= (x^2 + 2xy + 2y^2) + 4yz + 3z^2 = (x + y)^2 + y^2 + 4yz + 3z^2 \\ &= (x + y)^2 + (y^2 + 4yz) + 3z^2 = (x + y)^2 + (y + 2z)^2 - z^2. \end{aligned}$$

Also ist

$$f(x, y) = (x + y)^2 + (y + 2)^2 - 1. \qquad \text{(Probe!)}$$

22 Reelle symmetrische Bilinearformen

Definition 22.1. Sei V ein $\mathbb{R}$-Vektorraum und $\beta : V \times V \longrightarrow \mathbb{R}$ bilinear und symmetrisch. Man nennt β

(i) ***positiv definit*** (Skalarprodukt), falls $\beta(v, v) > 0$ für alle $v \in V \backslash \{0\}$ ist.

(ii) ***positiv semidefinit***, falls $\beta(v, v) \geq 0$ für alle $v \in V$ ist.

(iii) ***negativ definit***, falls $\beta(v, v) < 0$ für alle $v \in V \backslash \{0\}$ ist.

(iv) ***negativ semidefinit***, falls $\beta(v, v) \leq 0$ für alle $v \in V$ ist.

(v) ***indefinit***, falls $u, v \in V$ mit $\beta(u, u) > 0 > \beta(v, v)$ existieren.

Satz 22.1. *(Trägheitssatz von Sylvester)*
Sei V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum und $\beta : V \times V \longrightarrow \mathbb{R}$ bilinear und symmetrisch. Dann existiert eine Basis von V, bzgl. der die Matrix von β die folgende Form hat:

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

mit k Einsen, l Minus-Einsen und m Nullen auf der Diagonalen. Dabei ist k das Maximum der Dimensionen aller Untervektorräume U von V mit der Eigenschaft, dass die Einschränkung $U \times U \longrightarrow \mathbb{R}$ von β positiv definit ist. Analog ist l das Maximum der Dimensionen aller Untervektorräume U von V mit der Eigenschaft, dass die Einschränkung $U \times U \longrightarrow \mathbb{R}$ von β negativ definit ist. Insbesondere sind k, l, m unabhängig von der Wahl der Basis.

Beweis. Nach Satz 21.6 existiert eine Orthogonalbasis $c_1, \ldots, c_n$ von (v, β). Durch Umnummerierung können wir erreichen:

$$\begin{aligned} &\beta(c_i, c_i) > 0 \text{ für } i = 1, \ldots, k \\ &\beta(c_i, c_i) < 0 \text{ für } i = k+1, \ldots, k+l \\ &\beta(c_i, c_i) = 0 \text{ für } i = k+l+1, \ldots, n. \end{aligned}$$

Für $i = 1, \ldots, k+l$ sei $a_i := \sqrt{|\beta(c_i, c_i)|}$. Dann bilden

$$b_1 := a_1^{-1} c_1, \ldots, b_{k+l} := a_{k+l}^{-1} c_{k+l}, b_{k+l+1} := c_{k+l+1}, \ldots, b_n := c_n$$

eine Basis von V, bzgl. der die Matrix von β die gewünschte Form hat. Wir setzen $W := \text{Span}\,(b_1, \ldots, b_k)$. Ist $0 \neq w \in W$ und schreibt man $w = \sum_{i=1}^{k} r_i b_i \qquad (r_i \in \mathbb{R})$, so gilt:

$$\beta(w, w) = \sum_{i,j=1}^{k} r_i r_j \beta(b_i, b_j) = \sum_{i=1}^{k} r_i^2 > 0.$$

Daher ist $\dim W = k$, und die Einschränkung $W \times W \longrightarrow \mathbb{R}$ von β ist positiv definit.
Wir nehmen an, dass es einen Untervektorraum U von V mit $\dim U > k$ gibt, so dass die Einschränkung $U \times U \longrightarrow \mathbb{R}$ von β positiv definit ist. Aus Dimensionsgründen ist $U \cap \text{Span}\,(b_{k+1}, \ldots, b_n) \neq 0$. Sei $0 \neq x \in U \cap \text{Span}\,(b_{k+1}, \ldots, b_n)$. Wir schreiben $x = \sum_{i=k+1}^{n} s_i b_i$ $(s_i \in \mathbb{R})$. Dann ist

$$0 < \beta(x, x) = -\sum_{i=k+1}^{k+l} s_i^2 \leq 0.$$

Widerspruch. Analog für l. □

Bemerkung 22.1. Man nennt $k - l$ die ***Signatur*** von β. Andererseits ist $k + l = \text{rg}\,(\beta)$. Wegen $(k + l) + (k - l) = 2k$ und $(k + l) - (k - l) = 2l$ sind k und l durch Rang und Signatur eindeutig bestimmt.

Beispiel 22.1. Der euklidische Raum $\mathbb{R}^3$ hat Rang 3 und Signatur 3:

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

Der ***Minkowski-Raum*** $\mathbb{R}^4$ der speziellen Relativitätstheorie hat Rang 4 und Signatur 2:

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

Satz 22.2. *Seien V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum, $\beta : V \times V \longrightarrow \mathbb{R}$ bilinear und symmetrisch und $B = (b_{ij})$ die Matrix von β bzgl. einer Basis $b_1, \ldots, b_n$ von V. Genau dann ist β positiv definit, wenn für $m = 1, \ldots, n$ gilt:*

$$d_m := \begin{vmatrix} b_{11} & \ldots & b_{1m} \\ \vdots & & \vdots \\ b_{m1} & \ldots & b_{mm} \end{vmatrix} > 0.$$

Beweis. Für $m = 1, \ldots, n$ sei $V_m := \operatorname{Span}(b_1, \ldots, b_m)$ und $\beta_m : V_m \times V_m \longrightarrow \mathbb{R}$ die Einschränkung von β. Dann ist

$$B_m := \begin{pmatrix} b_{11} & \ldots & b_{1m} \\ \vdots & & \vdots \\ b_{m1} & \ldots & b_{mm} \end{pmatrix}$$

die Matrix von β_m bzgl. der Basis $b_1, \ldots, b_m$ von V_m.

"$\Rightarrow$": Sei β positiv definit. Für $m = 1, \ldots, n$ ist dann auch β_m positiv definit. Nach Sylvester besitzt (V_m, β_m) eine Orthonormalbasis. Daher existiert ein $S \in \mathrm{GL}(m, \mathbb{R})$ mit $S^T B_m S = 1_m$. Folglich ist

$$1 = \det 1_m = \det S^T B_m S = (\det S)^2 (\det B_m),$$

d.h. $\det B_m > 0$.

"$\Leftarrow$": Sei $d_m > 0$ für $m = 1, \ldots, n$. Im Fall $n = 1$ ist $V = \mathbb{R}b_1$ und

$$\beta(rb_1, rb_1) = r^2 \beta(b_1, b_1) = r^2 b_{11} = r^2 d_1 > 0$$

für $r \in \mathbb{R}\backslash\{0\}$. Also ist β positiv definit.
Sei also $n > 1$ und die Behauptung für $n - 1$ bereits bewiesen. Dann ist β_{n-1} bereits positiv definit. Nach Sylvester existiert ein Basis von V, bzgl. der die Matrix von β folgende Form hat:

$$C = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ & 1 & \\ 0 & & c \end{pmatrix}, \; c \in \{0, 1, -1\}.$$

Andererseits existiert ein $S \in \mathrm{GL}(n, \mathbb{R})$ mit $C = S^T B S$. Folglich ist

$$c = \det C = \det(S^T B S) = (\det S)^2 (\det B) > 0.$$

Also ist $c = 1$. Nach Sylvester ist β positiv definit.

□

Bemerkung 22.2. Es folgt, dass β genau dann negativ definit ist, wenn gilt:

$$d_1 < 0, d_2 > 0, d_3 < 0, d_4 > 0, \ldots$$

Beispiel 22.2. Sei $V := \mathbb{R}^n$, und sei $\alpha_n : V \times V \longrightarrow \mathbb{R}$ die symmetrische Bilinearform mit Matrix

$$A_n = \begin{pmatrix} 1 & -\frac{1}{2} & & 0 \\ -\frac{1}{2} & 1 & \ddots & \\ & \ddots & \ddots & -\frac{1}{2} \\ 0 & & -\frac{1}{2} & 1 \end{pmatrix}$$

bzgl. der Standardbasis. Dann gilt:

$$|A_1| = 1, |A_2| = \begin{vmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{vmatrix} = \frac{3}{4}.$$

Im Fall $n > 2$ liefert die Entwicklung nach der ersten Zeile und ersten Spalte:

$$|A_n| = |A_{n-1}| + \frac{1}{2}\left(-\frac{1}{2}\right)|A_{n-2}| = |A_{n-1}| - \frac{1}{4}|A_{n-2}|.$$

Ein einfacher Induktionsbeweis ergibt:

$$|A_n| = \frac{n+1}{2^n} > 0.$$

Also ist α_n positiv definit.
Diese und ähnliche Bilinearformen spielen eine wichtige Rolle bei endlichen "Spiegelungsgruppen".

Bemerkung 22.3. Gegeben sei eine symmetrische Matrix $B \in \mathbb{R}^{n\times n}$. Nach der Hauptachsentransformation (vgl. Satz 16.5) existiert eine Matrix $S \in O(n)$ mit der Eigenschaft, dass $D := S^T BS$ eine Diagonalmatrix ist, etwa

$$D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix};$$

dabei sind $d_1, \ldots, d_n$ die Eigenwerte von D und B.
Man kann annehmen, dass $d_1, \ldots, d_k$ positiv, $d_{k+1}, \ldots, d_{k+l}$ negativ und $d_{k+l+1}, \ldots, d_n$ gleich Null sind. Daher sind B und D zu der folgenden Matrix kongruent:

$$\begin{pmatrix} 1_k & & 0 \\ & -1_l & \\ 0 & & 0_m \end{pmatrix}$$

Die zu B gehörende symmetrische Bilinearform β ist also genau dann positiv definit (bzw. negativ definit), wenn die Eigenwerte $d_1, \ldots, d_n$ von B positiv (bzw. negativ) sind. Man sagt dann auch, dass die Matrix B ***positiv*** (bzw. ***negativ***) ***definit*** ist.

Bemerkung 22.4. (Anwendung in der Analysis)
Die Funktion $f : U \longrightarrow \mathbb{R}$ sei in der Umgebung U des Punktes $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ definiert und zweimal stetig differenzierbar mit

$$\frac{\partial f}{\partial x_1}(a) = \cdots = \frac{\partial f}{\partial x_n}(a) = 0.$$

Außerdem sei die symmetrische Matrix

$$J_f(a) := \left(\frac{\partial f}{\partial x_i \partial x_j}(a)\right) \in \mathbb{R}^{n\times n}$$

positiv definit (bzw. negativ definit). Dann besitzt f in a ein strenges lokales Minimum (bzw. Maximum). Das bedeutet $f(b) > f(a)$ (bzw. $f(b) < f(a)$) für alle $b \neq a$ in einer Umgebung von a. Ist $J_f(a)$ indefinit, so besitzt f in a kein lokales Extremum.

Bemerkung 22.5. Es seien $V := \mathbb{R}^4$ der Minkowski-Raum und β die entsprechende Bilinearform auf V. Eine lineare Abbildung $f : V \longrightarrow V$ mit

$$\beta(f(x), f(y)) = \beta(x, y) \qquad \text{für alle } x, y \in V$$

nennt man ***Lorentztransformation***. Die Menge aller Lorentztransformationen $f : V \longrightarrow V$ ist eine Gruppe bzgl. der Komposition von Abbildungen, die ***Lorentzgruppe*** von V. Wie sieht die Matrix $A = (a_{ij})$ einer Lorentztransformation $f : V \longrightarrow V$ bzgl. der Standardbasis e_1, e_2, e_3, e_4 von V aus? Wir brauchen

$$\beta(f(e_j), f(e_l)) = \beta(e_j, e_l) \qquad \text{für } j, l = 1, 2, 3, 4.$$

Dabei gilt:

$$\begin{aligned}\beta(f(e_j), f(e_l)) &= \beta\left(\sum_{i=1}^{4} a_{ij}e_i, \sum_{k=1}^{4} a_{kl}e_k\right)\\ &= \sum_{i,k=1}^{4} a_{ij}a_{kl}\beta(e_i, e_k) = \sum_{i=1}^{4} a_{ij}a_{il}\beta(e_i, e_i)\\ &= a_{1j}a_{1l} + a_{2j}a_{2l} + a_{3j}a_{3l} - a_{4j}a_{4l}\end{aligned}$$

und

$$\beta(e_j, e_l) = \begin{cases} 0 & j \neq l \\ 1 & j = l \in \{1, 2, 3\} \\ -1 & j = l = 4 \end{cases}$$

23 Lineare Algebra und Codes

Einsatzgebiete von Codes:
sichere Datenübertragung und Datenspeicherung (Satellit-Erde, Arbeitsspeicher-Festplatte, CD, ...)
Aufgabe: Erkennung und automatische Korrektur von Fehlern.

$$\underset{\text{Nachricht}}{\text{Absender}} \xrightarrow[\uparrow\ \text{Störungen (Rauschen)}]{\text{Kanal}} \underset{\text{gestörte Nachricht}}{\text{Empfänger}}$$

KOMME HEUTE NACHT → KOMME HEUTE NICHT

KOMME HEXTE NACHT

KOMME HIUTE NACHT

Wiederholt man jede Nachricht mehrfach, so können einzelne Fehler erkannt und korrigiert werden; aber der Aufwand ist hoch.
Ziel der Codierungstheorie: Gleicher Effekt mit weniger Aufwand.

$$\text{Nachricht} \longrightarrow \boxed{\text{Codierung}} \xrightarrow[\uparrow\ \text{Rauschen}]{\text{Kanal}} \boxed{\text{Decodierung}} \longrightarrow \text{Nachricht}$$

Modell:

- Nachrichten bestehen aus Nullen und Einsen
- Unterteile Nachricht in "Worte" fester Länge k

$$0110|\underbrace{1000}_{k=4}|0100|1001$$

- Die Worte fasst man als Elemente im $\mathbb{F}_2$-Vektorraum $\mathbb{F}_2^k$ auf. ($\mathbb{F}_2 = \{0, 1\}$ Körper)
- Jedes Wort wird bei der Codierung mit Zusatzinformationen versehen und erst dann übertragen.

Einfache Methode:
Hänge an jedes Wort ein ***Prüfbit*** an, das die Summe der Einträge enthält.

$$01100|10001|01001|10010$$

Dann können einzelne Fehler erkannt, aber nicht korrigiert werden:

$$\begin{aligned} 10101 &\longrightarrow 00101 \\ &\longrightarrow 11101 \\ &\ldots \end{aligned}$$

Verfeinerte Methode:
Hänge an jedes Wort mehrere Prüfbits an:

$$\underbrace{\underbrace{\square}_{k\ \text{Informationsbits}}\ \underbrace{\square}_{n-k\ \text{Prüfbits}}}_{\text{Länge } n}$$

Beispiel:
Hamming-Code (1948)
eingesetzt in IBM- Großrechnern und Telefonschaltzentralen

$k = 4, n = 7$
Nachricht
$(x_1, x_2, x_3, x_4) \in \mathbb{F}_2^4 \overset{\text{Codierung}}{\longrightarrow} \underbrace{(x_1, x_2, x_3, x_4, x_5, x_6, x_7)}_{\text{“Codewort”}} \in \mathbb{F}_2^7.$
Regel: $x_5 = x_1 + x_2 + x_3, x_6 = x_1 + x_2 + x_4, x_7 = x_1 + x_3 + x_4$
Beispiel: $(1, 0, 1, 0) \longmapsto (1, 0, 1, 0, 0, 1, 0)$
Beachte: Die Codierung ist eine lineare Abbildung $\mathbb{F}_2^4 \longrightarrow \mathbb{F}_2^7$.
Decodierung:

$$H = \left(\begin{array}{cccc|ccc} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \qquad \text{“Hamming-Matrix”}$$

- Berechne
$$H \begin{pmatrix} x_1 \\ \vdots \\ x_7 \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}.$$
- Falls $\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, so ist x ein Codewort.
- Im Fall $\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ ist $\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}$ eine der Spalten von H, sagen wir die i-te. Wir ändern dann den i-ten Eintrag von $(x_1, \ldots, x_7)$ und erhalten ein Codewort.

$x' = (1, 0, 1, 0, 1, 1, 0), Hx' = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ = 5. Spalte von $H \overset{\text{Decodierung}}{\longrightarrow} (1, 0, 1, 0, 0, 1, 0) = x$

$y' = (0, 0, 1, 1, 0, 0, 0), Hy' = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ = 2. Spalte von $H \longrightarrow (0, 1, 1, 1, 0, 0, 0) = y$

Probe: $Hy = 0$
$x = (x_1, \ldots, x_7)$ beliebiges Codewert $\overset{\text{Kanal}}{\longrightarrow} x' = (x_1', \ldots, x_7')$.
Bei der Übertragung trete genau ein Fehler auf, und zwar an Position i.
Dann ist $x' = x + e_i$ ($e_i = i$-ter Vektor der Standardbasis)
Folglich ist $Hx' = \underbrace{Hx}_{=0} + He_i = He_i = i$-te Spalte von H.
Man kann also die Position erkennen, bei der der Fehler aufgetreten ist.

Definition 23.1. Ein (***binärer***) ***Code*** der ***Länge*** n ist ein Untervektorraum C von $\mathbb{F}_2^n$. Die Elemente in C nennt man ***Codewörter***.

Beispiel 23.1. Hamming-Code $C = \{x \in \mathbb{F}_2^7 : Hx = 0\}$: Länge $n = 7$, Dimension $k = 4$.
Anzahl der Elemente in $\mathbb{F}_2^7 : 2^7 = 128$
Anzahl der Elemente in $C : 2^4 = 16$
Sei $x \in C$ fest.
Anzahl der Elemente, die sich von x an genau einer Position unterscheiden: 7
Anzahl dieser Elemente insgesamt: $16 \cdot 7 = 112$ (Kein Element wird mehrfach gezählt.)
Beachte: $16 + 112 = 128$.
Daher ist jedes Element in $\mathbb{F}_2^7$ entweder ein Codewort, oder es unterscheidet sich von einem Codewort an genau einer Position.

Etwas Geometrie:
Für $v = (v_1, \ldots, v_n), w = (w_1, \ldots, w_n) \in \mathbb{F}_2^n$ setze allgemein:

$$d(v, w) := |\{i : v_i \neq w_i\}|$$
$$= \text{Anzahl der Positionen, an denen sich } v \text{ und } w \text{ unterscheiden}$$

d ***Hamming-Distanz***
$v = (1, 0, 0, 1, 0, 1, 1), w = (1, 1, 0, 0, 1, 1, 0), d(v, w) = 4.$

Eigenschaften: $d(v, w) \geq 0$
$d(v, w) = 0 \Leftrightarrow v = w$
$d(v, w) = d(w, v)$
$d(u, w) \leq d(u, v) + d(v, w)$ (Dreiecksungleichung)

d Metrik auf $\mathbb{F}_2^n$, $(\mathbb{F}_2^n, d)$ metrischer Raum.
Kugel um $v \in \mathbb{F}_2^n$ mit Radius $r > 0$:

$$K(v, r) = \{w \in \mathbb{F}_2^n : d(v, w) \leq r\}.$$

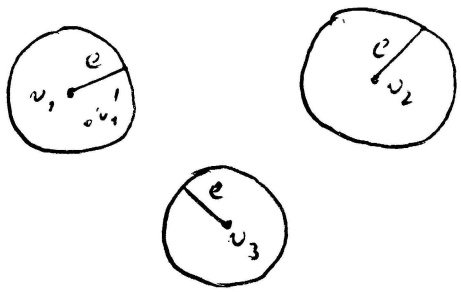
Anzahl der Elemente in $K(v, r)$ für $r \in \mathbb{N}, r \leq n$:

$$|K(v, r)| = 1 + n + \binom{n}{2} + \cdots + \binom{n}{r}.$$

Minimalabstand eines Codes $C \subseteq \mathbb{F}_2^n$:

$$\delta = \delta(C) = \min\{d(v, w) : v, w \in C, v \neq w\}.$$

Wichtig: Hat C Minimalabstand $2e + 1$, so kann C e Fehler korrigieren:



Da sich die Kugeln nicht schneiden, gilt: $2^k \cdot |K(v,e)| \leq 2^n$.

Definition 23.2. Ein binärer Code C (Länge n, Dimension k, Minimalabstand $\delta = 2e+1$) heißt ***perfekt***, falls $2^k|K(v,e)| = 2^n$ gilt.

Bemerkung 23.2. Das bedeutet, dass die Kugeln vom Radius e um Codewörter ganz $\mathbb{F}_2^n$ überdecken. (Der Raum ist also "perfekt" ausgenutzt.)

Beispiel 23.2. Der Hamming-Code ist perfekt mit Minimalabstand 3:

$$(0,0,0,0,0,0,0), (1,1,0,0,0,0,1) \in C$$

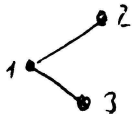
Er kann also genau 1 Fehler korrigieren.

Frage: Wie sehen perfekte Codes mit Minimalabstand 5 aus? (Solche Codes können also 2 Fehler korrigieren.)

Sei $C \subseteq \mathbb{F}_2^n$ perfekter Code, Dimension k, Minimalabstand 5.
Beobachtung: $2^k \cdot |K(v,2)| = 2^n$, also $|K(v,2)| = 2^{n-k}$ Potenz von 2
Andererseits: $|K(v,2)| = 1 + n + \binom{n}{2} = \frac{n^2+n+2}{2}$.
Daher ist auch $n^2 + n + 2 = 2^{n-k+1}$ Potenz von 2. (z.B. $n = 1, 2, 5, \dots$)
Ordne die Elemente in $\mathbb{F}_2^n : w_1 = 0 \dots 0, w_2 = 0 \dots 01, \dots, w_{2^n} = 1 \dots 1$.
Definiere Matrix $A_n = (a_{ij}) \in \mathbb{R}^{2^n} \times \mathbb{R}^{2^n}$ durch

$$a_{ij} := \begin{cases} 1 & \text{falls } d(w_i, w_j) = 1 \\ 0 & \text{sonst.} \end{cases}$$

Randbemerkung: A_n ist die ***Inzidenzmatrix*** eines ***Graphen*** Γ:
Ecken: Elemente in $\mathbb{F}_2^n$
Kanten: w_i— $w_j \Leftrightarrow d(w_i, w_j) = 1$.



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

$$A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

allgemein: $A_{n+1} = \left(\begin{array}{c|c} A_n & 1_n \\ \hline 1_n & A_n \end{array}\right)$
Beachte: A_n ist symmetrisch, also diagonalisierbar. (Hauptachsentransformation)

Es existiert also eine Basis aus Eigenvektoren $v_1, \ldots, v_{2^n}$ zu Eigenwerten $r_1, \ldots, r_{2^n}$. Sei $v := v_i$ und $r = r_i$. Dann:

$$\begin{pmatrix} A_n & 1_n \\ 1_n & A_n \end{pmatrix} \begin{pmatrix} v \\ v \end{pmatrix} = \begin{pmatrix} rv + v \\ v + rv \end{pmatrix} = (r+1) \begin{pmatrix} v \\ v \end{pmatrix}$$
$$\begin{pmatrix} A_n & 1_n \\ 1_n & A_n \end{pmatrix} \begin{pmatrix} v \\ -v \end{pmatrix} = \begin{pmatrix} rv - v \\ v - rv \end{pmatrix} = (r-1) \begin{pmatrix} v \\ -v \end{pmatrix}$$

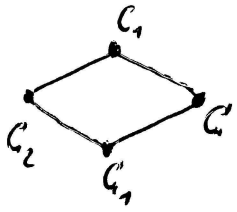
Jeder Eigenvektor von A_n liefert also 2 Eigenvektoren von A_{n+1}. Man sieht leicht, dass diese linear unabhängig sind, also eine Basis bilden.

Eigenwerte: $n = 1 : A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ Eigenwerte: 1,-1

$n = 2 : 2, 0, -2$
$n = 3 : 3, 1, -1, -3$
allgemein: $\{n, n-2, n-4, \ldots, -n\}$
Jetzt nummerieren wir die Elemente in $\mathbb{F}_2^n$ anders:
Erst die Elemente in C
Dann die Elemente mit Abstand 1 zu den Elementen in C: C_1
Schließlich die Elemente mit Abstand 2 zu den Elementen in C: C_2
Wir definieren analog die Matrix A':

$$A' = \left(\begin{array}{c|c|c} 0 & \star & 0 \\ \hline \star & 0 & \star \\ \hline 0 & \star & \star \end{array}\right)$$

Wie hängen A und A' zusammen?
Durch Zeilen- und Spaltenvertauschungen: $A' = PAP^{-1}$, P Permutationsmatrix
Insbesondere sind A und A' ähnlich, haben also die gleichen Eigenwerte.
Beobachtung: In A und A' ist jede Zeilensumme gleich n. Betrachte die Zeilensummen in den Teilmatrizen:



	C	C_1	C_2
C	0	n	0
C_1	1	0	$n-1$
C_2	0	2	$n-2$

Betrachte Zeilensummenmatrix

$$Z = \begin{pmatrix} 0 & n & 0 \\ 1 & 0 & n-1 \\ 0 & 2 & n-2 \end{pmatrix} \in \mathbb{R}^{3\times 3}$$

Beachte: Ist $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ Eigenvektor von Z zum Eigenwert r, so ist $\begin{pmatrix} a \\ \vdots \\ a \\ b \\ \vdots \\ b \\ c \\ \vdots \\ c \end{pmatrix}$ Eigenvektor von A'

zum Eigenwert r. Daher liegt jeder Eigenwert von Z in $\{n, n-2, n-4, \ldots, -n\}$.
Charakteristisches Polynom von Z:

$$\begin{vmatrix} x & -n & 0 \\ -1 & x & 1-n \\ 0 & -2 & x-n+2 \end{vmatrix} = \cdots = x^3 + (2-n)x^2 + (2-3n)x + n^2 - 2n$$

$$= (x-n)[(x-n)^2 + (2n+2)(x-n) + n^2 + n + 2]$$
$$= y[y^2 + (2n+2)y + n^2 + n + 2] \text{ nach Substitution } y = x - n$$
$$= y(y-y_1)(y-y_2)$$

$y_1 y_2 = n^2 + n + 2$ Potenz von 2
Also: $y_1 = \pm 2^a$, $y_2 = \pm 2^b$, o.B.d.A. $a \leq b$.
Beachte: y_1, y_2 negativ, außerdem: $y_1 + y_2 = -(2n+2)$
Im Fall $3 \leq a \leq b$ ist $2n+2$ durch 8, also $n+1$ durch 4 teilbar. Dann ist aber $n^2 + n + 2 = n(n+1) + 2$ nicht durch 4 teilbar. Widerspruch.
Also ist $a \leq 2$.

$a = 0$: $y_1 = -1$: $0 = (-1)^2 - (2n+2) + n^2 + n + 2 = n^2 - n + 1$ (geht nicht)

$a = 1$: $y_1 = -2$: $0 = (-2)^2 - (4n+4) + n^2 + n + 2 = n^2 - 3n + 2 \Rightarrow n \in \{1, 2\}$.

$a = 2$: $y_1 = -4$: $0 = (-4)^2 - (8n+8) + n^2 + n + 2 = n^2 - 7n + 10 \Rightarrow n \in \{2, 5\}$.

$n = 2$: $00, 01, 10, 11$ $\quad C = \{00\}$

$n = 5$: $00000, \ldots, 11111$ $\quad C = \{00000, 11111\}$.

Fazit: Es gibt keine "interessanten" perfekten Codes mit Minimalabstand 5 (2-fehler-korrigierend).

Frage: Gibt es "interessante" perfekte Codes mit Minimalabstand 7?
Ja: im Wesentlichen genau einen:
den ***Golay-Code*** der Länge 23 und der Dimension 12
Beachte: $2^{23} = 2^{12}\left[1 + 23 + \binom{23}{2} + \binom{23}{3}\right]$

Satz 23.3. *"Interessante" perfekte Codes mit größerem Minimalabstand existieren nicht.*

Beweis. Vorlesung Codierungstheorie □

Bemerkung 23.4. (i) Die ***Symmetriegruppe*** des Golay-Codes ist die ***sporadische Mathieugruppe*** M_{23}.

(ii) Auf einer CD verwendet man zwei (nichtperfekte) Codes:
Längen $28, 32$ Dimensionen $24, 28$ Minimalabstand 5
Es wird nicht mit dem Körper $\mathbb{F}_2$, sondern mit dem Körper $\mathbb{F}_{256}$ gearbeitet.

24 Unitäre Vektorräume

Definition 24.1. Gegeben seien $\mathbb{C}$-Vektorräume U, V. Eine Abbildung $\sigma : U \times V \longrightarrow \mathbb{C}$ nennt man ***Sesquilinearform***, falls für alle $z, z' \in \mathbb{C}, u, u' \in U, v, v' \in V$ gilt:

(i) $\sigma(zu + z'u', v) = z\sigma(u, v) + z'\sigma(u', v)$;

(ii) $\sigma(u, zv + z'v') = \overline{z}\sigma(u, v) + \overline{z'}\sigma(u, v')$.

Hier bezeichnet $\overline{z} = a - ib$ wie üblich die zu $z = a + ib$ $(a, b \in \mathbb{R})$ konjugiert komplexe Zahl.

Bemerkung 24.1. (i) Ggf. ist $\sigma(0, v) = \sigma(0 \cdot 0, v) = 0\sigma(0, v) = 0$ für alle $v \in V$ und analog $\sigma(u, 0) = 0$ für alle $u \in U$.

(ii) Für Basen $a_1, \ldots, a_m$ von U und $b_1, \ldots, b_n$ von V nennt man

$$S := (\sigma(a_i, b_j)) \in \mathbb{C}^{m \times n}$$

die ***Matrix*** von σ bzgl. $a_1, \ldots, a_m$ und $b_1, \ldots, b_n$.

(iii) Seien $a'_1, \ldots, a'_m$ und $b'_1, \ldots, b'_n$ weitere Basen von U bzw. V. Wir schreiben $a'_i = \sum_{k=1}^{m} r_{ki} a_k$ und $b'_j = \sum_{l=1}^{n} t_{lj} b_l$ mit Elementen $r_{ki}, t_{lj} \in \mathbb{C}$. Dann ist (vgl. Bemerkung 21.2)

$$\boxed{S' := R^T S \overline{T}}$$

die Matrix von σ bzgl. $a'_1, \ldots, a'_m$ und $b'_1, \ldots, b'_n$; dabei ist $R := (r_{ij}) \in \mathrm{GL}(m, \mathbb{C})$ und $T = (t_{ij}) \in \mathrm{GL}(n, \mathbb{C})$. Wegen $\mathrm{rg}(S') = \mathrm{rg}(S)$ kann man den ***Rang*** von σ durch $\mathrm{rg}(\sigma) := \mathrm{rg}(S)$ definieren. Es ist klar, dass man Basen von U und V so wählen kann, dass die Matrix von σ bzgl. dieser Basen folgende Form hat:

$$\begin{pmatrix} 1_r & 0 \\ 0 & 0 \end{pmatrix} \qquad (r = \mathrm{rg}(\sigma)).$$

(iv) Für Sesquilinearformen $\sigma, \sigma' : U \times V \longrightarrow \mathbb{C}$ und $w, w' \in \mathbb{C}$ ist auch die Abbildung

$$w\sigma + w'\sigma' : U \times V \longrightarrow \mathbb{C}, (u, v) \longmapsto w\sigma(u, v) + w'\sigma'(u, v),$$

sesquilinear; denn für $u, u' \in U, v, v' \in V, z, z' \in \mathbb{C}$ gilt:

$$\begin{aligned}(w\sigma + w'\sigma')(u, zv + z'v') &= w\sigma(u, zv + z'v') + w'\sigma'(u, zv + z'v') \\ &= w[\overline{z}\sigma(u, v) + \overline{z'}\sigma(u, v)] + w'[\overline{z}\sigma'(u, v) + \overline{z'}\sigma'(u', v)] \\ &= \overline{z}[w\sigma(u, v) + w'\sigma'(u, v)] + \overline{z'}[w\sigma(u, v') + w'\sigma'(u, v')] \\ &= \overline{z}[w\sigma + w'\sigma'](u, v) + \overline{z'}[w\sigma + w'\sigma'](u, v')\end{aligned}$$

und

$$(w\sigma + w'\sigma')(zu + z'u', v) = \cdots = z[w\sigma + w'\sigma'](u, v) + z'[w\sigma + w'\sigma'](u', v).$$

Daher bilden die Sesquilinearformen $\sigma : U \times V \longrightarrow \mathbb{C}$ einen Untervektorraum $\mathrm{Ses}(U, V; \mathbb{C})$ von $\mathrm{Abb}(U \times V, \mathbb{C})$.

Satz 24.1. *Für $\mathbb{C}$-Vektorräume U, V mit Basen $a_1, \ldots, a_m$ bzw. $b_1, \ldots, b_n$ ist die Abbildung*

$$F : \mathrm{Ses}(U, V; \mathbb{C}) \longrightarrow \mathbb{C}^{m \times n}, \sigma \longmapsto (\sigma(a_i, b_j))$$

ein Vektorraum-Isomorphismus; insbesondere ist $\dim(\mathrm{Ses}(U, V; \mathbb{C})) = mn = (\dim U)(\dim V)$.

Beweis. Wie im Beweis von Satz 21.2 zeigt man, dass F linear und injektiv ist. Zum Beweis der Surjektivität sei $S = (s_{ij}) \in \mathbb{C}^{m \times n}$. Man rechnet leicht nach, dass die Abbildung $\sigma : U \times V \longrightarrow \mathbb{C}$ mit

$$\sigma\left(\sum_{i=1}^{m} w_i a_i, \sum_{j=1}^{n} z_j b_j\right) := \sum_{i=1}^{m} \sum_{j=1}^{n} w_i \overline{z}_j s_{ij} \qquad (w_i, z_j \in \mathbb{C})$$

sesquilinear mit Matrix S bzgl. $a_1, \ldots, a_m$ und $b_1, \ldots, b_n$ ist. Daher ist F auch surjektiv. □

Beispiel 24.1. (i) Für jede Matrix $A \in \mathbb{C}^{m \times n}$ ist die Abbildung

$$\sigma : \mathbb{C}^m \times \mathbb{C}^n \longrightarrow \mathbb{C}, (x, y) \longmapsto xA\overline{y}^T,$$

sesquilinear (vgl. Beispiel 21.2). Wie dort ist A selbst die Matrix von σ bzgl. der Standardbasen von $\mathbb{C}^m$ und $\mathbb{C}^n$.

(ii) Die Abbildung

$$\sigma : \mathbb{C}^3 \times \mathbb{C}^2 \longrightarrow \mathbb{C}, (x, y) \longmapsto x_2\overline{y_1} + ix_3\overline{y_2},$$

$x = (x_1, x_2, x_3) \in \mathbb{C}^3, y = (y_1, y_2) \in \mathbb{C}^2$ ist sesquilinear; die Matrix von σ bzgl. der Basen $(1, 0, 0), (0, i, 0), (i, 0, 1)$ von $\mathbb{C}^3$ und $(1, i), (i, 1)$ von $\mathbb{C}^2$ ist

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

Bemerkung 24.2. (i) Sei V ein $\mathbb{C}$-Vektorraum mit Basis $b_1, \ldots, b_n$ und $\sigma : V \times V \longrightarrow \mathbb{C}$ sesquilinear; dann nennt man $S := (\sigma(b_i, b_j)) \in \mathbb{C}^{n\times n}$ die ***Matrix*** von σ bzgl. $b_1, \ldots, b_n$.

(ii) Ist $b'_1, \ldots, b'_n$ eine weitere Basis von V und schreibt man $b'_j = \sum_{i=1}^n r_{ij} b_i \qquad (r_{ij} \in \mathbb{C})$, so ist $R := (r_{ij}) \in \mathrm{GL}(n, \mathbb{C})$, und $S' = R^T S \overline{R}$ ist die Matrix von σ bzgl. $b'_1, \ldots, b'_n$.

(iii) Man nennt $A, B \in \mathbb{C}^{n\times n}$ ***unitär-kongruent***, falls $B = R^T A \overline{R}$ für ein $R \in \mathrm{GL}(n, \mathbb{C})$ ist. Diese Relation auf $\mathbb{C}^{n\times n}$ ist wie üblich reflexiv, symmetrisch und transitiv.

Definition 24.3. Seien V ein $\mathbb{C}$-Vektorraum und $\sigma : V \times V \longrightarrow \mathbb{C}$ sesquilinear mit $\sigma(u, v) = \overline{\sigma(v, u)}$ für alle $u, v \in V$. Dann nennt man σ ***hermitesch*** (C. Hermite, 1822-1901).

Bemerkung 24.3. (i) Eine Matrix $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ mit $\overline{A}^T = A$ nennt man ***hermitesch***; dabei ist $\overline{A} := (\overline{a_{ij}})$.

(ii) Sei $b_1, \ldots, b_n$ eine Basis von V. Man zeigt leicht, dass eine Sesquilinearform σ auf V genau dann hermitesch ist, wenn die Matrix von σ bzgl. $b_1, \ldots, b_n$ hermitesch ist (vgl. Bemerkung 21.4).

(iii) Wie üblich definiert man ***Orthogonalbasen*** und ***Orthonormalbasen*** von (V, σ).

(iv) Für jede hermitesche Sesquilinearform $\sigma : V \times V \longrightarrow \mathbb{C}$ und alle $v \in V$ gilt: $\sigma(v, v) = \overline{\sigma(v, v)}$, d.h. $\sigma(v, v) \in \mathbb{R}$. Man nennt σ ***positiv definit***, falls $\sigma(v, v) > 0$ für alle $v \in V \setminus \{0\}$ ist. Analog hat man die Begriffe ***negativ definit***, ***positiv (negativ) semidefinit***, ***indefinit***.

Satz 24.3. *(Trägheitssatz von Sylvester)*
Seien V ein endlich-dimensionaler $\mathbb{C}$-Vektorraum und $\sigma : V \times V \longrightarrow \mathbb{C}$ sesquilinear und hermitesch. Dann existiert eine Basis von V, bzgl. der die Matrix von σ folgende Form hat:

$$\begin{pmatrix} 1_k & & 0 \\ & -1_l & \\ 0 & & 0_m \end{pmatrix}.$$

Dabei ist k das Maximum der Dimensionen aller Untervektorräume U von V mit der Eigenschaft, dass die Einschränkung $U \times U \longrightarrow \mathbb{C}$ von σ positiv definit ist. Analog ist l das Maximum der Dimensionen aller Untervektorräume U von V mit der Eigenschaft, dass die Einschränkung $U \times U \longrightarrow \mathbb{C}$ von σ negativ definit ist. Insbesondere sind k, l, m unabhängig von der Wahl der Basis.

Beweis. Wir zeigen zunächst durch Induktion nach $n := \dim V$ die Existenz einer Orthogonalbasis. Im Fall $n = 1$ ist nichts zu tun. Sei also $n > 1$. Im Fall $\sigma = 0$ ist auch nichts zu tun. Seien also $x, y \in V$ mit $a := \sigma(x, y) \neq 0$. Setzt man $z := \overline{a}^{-1} y$, so ist $\sigma(x, z) = 1$.
Im Fall $\sigma(v, v) = 0$ für alle $v \in V$ hätte man den Widerspruch

$$0 = \sigma(x + z, x + z) = \underbrace{\sigma(x, x)}_{=0} + \underbrace{\sigma(x, z)}_{=1} + \underbrace{\sigma(x, z)}_{=1} + \underbrace{\sigma(z, z)}_{=0} = 2.$$

Also existiert ein $b_1 \in V$ mit $\sigma(b_1, b_1) \neq 0$; insbesondere ist $b_1 \neq 0$. Wir ergänzen b_1 zu einer Basis $b_1, c_2, \ldots, c_n$ von V. Für $i = 2, \ldots, n$ sei

$$d_i := c_i - \frac{\sigma(c_i, b_1)}{\sigma(b_1, b_1)} b_1.$$

Wie im Beweis von Satz 21.6 ist dann $\sigma(d_i, b_1) = 0$. Außerdem bilden $b_1, d_2, \ldots, d_n$ eine Basis von V. Dann ist $\sigma(x, b_1) = 0$ für alle $x \in U := \text{Span}\,(d_2, \ldots, d_n)$. Die Einschränkung

$$\tau : U \times U \longrightarrow \mathbb{C}, (u, v) \longmapsto \sigma(u, v)$$

ist auch sesquilinear und hermitesch. Nach Induktion existiert eine Orthogonalbasis $b_2, \ldots, b_n$ von (U, τ). Dann ist $b_1, b_2, \ldots, b_n$ eine Orthogonalbasis von (V, σ). Nach Umnummerierung können wir annehmen:

$$\begin{aligned}
&\sigma(b_i, b_i) > 0 \text{ für } i = 1, \ldots, k \\
&\sigma(b_i, b_i) < 0 \text{ für } i = k+1, \ldots, k+l \\
&\sigma(b_i, b_i) = 0 \text{ für } i = k+l+1, \ldots, n.
\end{aligned}$$

Dann bilden

$$a_1 := \frac{b_1}{\sqrt{|\sigma(b_1, b_1)|}}, \ldots, a_{k+l} := \frac{b_{k+l}}{\sqrt{|\sigma(b_{k+l}, b_{k+l})|}}, a_{k+l+1} := b_{k+l+1}, \ldots, a_n = b_n$$

eine Basis von V, bzgl. der die Matrix von σ die gewünschte Form hat. Wir setzen $W := \text{Span}\,(a_1, \ldots, a_k)$. Ist $0 \neq w \in W$ und schreibt man $w = \sum_{j=1}^{k} r_j a_j \qquad (r_j \in \mathbb{C})$, so gilt:

$$\sigma(w, w) = \sum_{i,j=1}^{k} r_i \overline{r_j} \sigma(a_i a_j) = \sum_{j=1}^{k} r_j \overline{r_j} > 0.$$

Daher ist $\dim W = k$, und die Einschränkung $W \times W \longrightarrow \mathbb{C}$ von σ ist positiv definit.
Wie im Beweis von Satz 22.1 zeigt man, dass für jeden Untervektorraum U von V mit der Eigenschaft, dass die Einschränkung $U \times U \longrightarrow \mathbb{C}$ von σ positiv definit ist, gilt: $\dim U \leq k$. Analog beweist man die Eindeutigkeit von l. □

Beispiel 24.3. Gegeben sei die hermitesche Matrix

$$S := \begin{pmatrix} 1 & i \\ -i & -1 \end{pmatrix} \in \mathbb{C}^{2\times 2}.$$

Gesucht wird eine Matrix $R \in \text{GL}\,(2, \mathbb{C})$ mit $R^T S \overline{R} = \begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}$.

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

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

Also:

$$R = \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 & \frac{i}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} \end{pmatrix}. \qquad \text{(Probe!)}$$

Bemerkung 24.4. Seien V ein endlich-dimensionaler $\mathbb{C}$-Vektorraum und $\sigma : V \times V \longrightarrow \mathbb{C}$ sesquilinear und hermitesch. Wie in Bemerkung 22.1 definiert man die ***Signatur*** von σ.

Definition 24.4. Ein ***unitärer Vektorraum*** ist ein Paar (V, σ), das aus einem $\mathbb{C}$-Vektorraum V und einer positiv definiten hermiteschen Sesquilinearform σ auf V besteht. Statt $\sigma(u, v)$ schreibt man oft auch $(u|v)$. Man spricht dann von einem ***hermiteschen Skalarprodukt*** auf V.

Beispiel 24.4. Für $n \in \mathbb{N}$ ist $\mathbb{C}^n$ mit dem ***kanonischen hermiteschen Skalarprodukt*** $((w_1, \ldots, w_n), (z_1, \ldots, z_n)) \longmapsto w_1\overline{z_1} + \cdots + w_n\overline{z_n}$ ein unitärer Vektorraum.

Bemerkung 24.5. (i) Nach dem Trägheitssatz von Sylvester enthält jeder unitäre Vektorraum V eine Orthonormalbasis (ONB) $b_1, \ldots, b_n$; d.h.

$$(b_i|b_j) = \delta_{ij} \qquad \text{für } i, j = 1, \ldots, n.$$

(ii) Wie bei euklidischen Vektorräumen beweist man für jeden unitären Vektorraum V die ***Cauchy-Schwarz-Ungleichung*** (CSU):

$$(x|y)\overline{(x|y)} \leq (x|x)(y|y) \qquad (x, y \in V).$$

Genau dann gilt Gleichheit, wenn x und y linear abhängig über $\mathbb{C}$ sind.

(iii) Daher kann man für jeden unitären Vektorraum V durch $\|x\| := \sqrt{(x|x)}$ für $x \in V$ eine ***Norm*** $\|\cdot\| : V \longrightarrow \mathbb{R}$ mit den üblichen Eigenschaften definieren. Eine entsprechende Metrik $d : V \times V \longrightarrow \mathbb{R}$ definiert man dann durch $d(x, y) := \|x - y\|$ für $x, y \in V$. Dabei gelten ebenfalls die üblichen Eigenschaften.

(iv) Jeden $\mathbb{C}$-Vektorraum V kann man nach Einschränkung der Skalarmultiplikation auch als $\mathbb{R}$-Vektorraum betrachten. Bilden $b_1, \ldots, b_n$ eine $\mathbb{C}$-Basis von V, so bilden $b_1, ib_1, \ldots, b_n, ib_n$ eine $\mathbb{R}$-Basis von V; insbesondere ist $\dim_{\mathbb{R}} V = 2 \dim_{\mathbb{C}} V$.

(v) Sei jetzt V ein unitärer Vektorraum. Wir setzen $\rho(z, w) := \mathrm{Re}\,(z|w)$ und $\tau(z, w) := \mathrm{Im}\,(z|w)$ für alle $z, w \in V$. Man rechnet leicht nach, dass ρ ein Skalarprodukt auf dem $\mathbb{R}$-Vektorraum V ist. Auf diese Weise kann man V als euklidischen Vektorraum ansehen. Für alle $z, w \in V$ ist

$$\tau(z, w) = \mathrm{Im}\,(z|w) = \mathrm{Re}\,(-i(z|w)) = \mathrm{Re}\,(z|iw) = \rho(z, iw).$$

Daher ist $(\cdot|\cdot)$ durch ρ eindeutig bestimmt.

Definition 24.6. Seien V ein unitärer Vektorraum und $X \subseteq V$. Dann nennt man

$$X^\perp := \{v \in V : (v|x) = 0 \text{ für alle } x \in X\}$$

den ***Orthogonalraum*** von X in V.

Bemerkung 24.6. (i) Stets ist $X^\perp$ ein Untervektorraum von V.

(ii) Ist $\dim V < \infty$ und ist X ein Untervektorraum von V, so gilt:

$$V = X \oplus X^\perp \text{ und } X^{\perp\perp} = X;$$

insbesondere ist $\dim V = \dim X + \dim X^\perp$. Man nennt $X^\perp$ auch das ***orthogonale Komplement*** von X in V.

(iii) Ist $\dim V < \infty$, so gilt für beliebige Untervektorräume U_1, U_2 von V:

$$(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp \text{ und } (U_1 \cap U_2)^\perp = U_1^\perp + U_2^\perp.$$

Definition 24.7. Eine ***Isometrie*** zwischen unitären Vektorräumen U, V ist eine lineare Abbildung $f : U \longrightarrow V$ mit $(f(u)|f(u')) = (u|u')$ für alle $u, u' \in U$.

Bemerkung 24.7. (i) Ggf. ist f injektiv; denn aus $f(u) = 0$ folgt $(u|u) = (f(u)|f(u)) = (0|0) = 0$, d.h. $u = 0$.

(ii) Für jeden unitären Vektorraum W und jede weitere Isometrie $g : V \longrightarrow W$ ist auch $g \circ f : U \longrightarrow W$ eine Isometrie.

(iii) Stets ist id_V eine Isometrie.

(iv) Für jede bijektive Isometrie $f : U \longrightarrow V$ ist auch die Umkehrabbildung $f^{-1} : V \longrightarrow U$ eine Isometrie.

(v) Für jeden unitären Vektorraum V bilden die bijektiven Isometrien $f : V \longrightarrow V$ eine Gruppe $U(V)$ bzgl. der Komposition von Abbildungen. Diese bezeichnet man als ***unitäre Gruppe*** von V. Die Elemente in $U(V)$ nennt man auch ***unitäre Transformationen***.

(vi) Seien V, W endlich-dimensionale unitäre Vektorräume, und sei $f : V \longrightarrow W$ linear mit Matrix $A = (a_{ij}) \in \mathbb{C}^{n\times m}$ bzgl. Orthonormalbasen $a_1, \ldots, a_m$ von V und $b_1, \ldots, b_n$ von W. Genau dann ist f eine Isometrie, wenn gilt:

$$(f(a_i)|f(a_j)) = (a_i|a_j) \qquad \text{für } i, j = 1, \ldots, m.$$

Dabei ist

$$(f(a_i)|f(a_j)) = \left(\sum_{k=1}^{n} a_{ki} b_k \Big| \sum_{l=1}^{n} a_{lj} b_l\right) = \sum_{k,l=1}^{n} a_{ki}\overline{a_{lj}}(b_k|b_l)$$
$$= \sum_{k=1}^{n} a_{ki}\overline{a_{kj}}.$$

Daher ist f genau dann eine Isometrie, wenn gilt: $\boxed{A^T\overline{A} = 1_m.}$

(vii) Man nennt eine Matrix $A \in \mathbb{C}^{n\times n}$ ***unitär***, wenn gilt: $A^T\overline{A} = 1_n$ (d.h. $A^{-1} = \overline{A}^T$). Die unitären Matrizen $A \in \mathbb{C}^{n\times n}$ bilden eine Gruppe $U(n) = U(n,\mathbb{C})$ bzgl. der Multiplikation von Matrizen. Diese bezeichnet man als ***unitäre Gruppe*** des ***Grades*** n (über $\mathbb{C}$). Für $A \in U(n)$ gilt:

$$1 = (\det A)(\overline{\det A}), \text{ d.h.} |\det A| = 1.$$

(viii) Sei V ein unitärer Vektorraum mit Orthonormalbasis $b_1, \ldots, b_n$, und sei $f \in \mathrm{End}\,(V)$ mit Matrix A bzgl. $b_1, \ldots, b_n$. Dann gilt:

$$f \in U(V) \Leftrightarrow A \in U(n).$$

Ggf. ist $|\det f| = |\det A| = 1$.

Definition 24.8. Man nennt unitäre Vektorräume V, W ***isometrisch isomorph***, wenn eine bijektive Isometrie $f : V \longrightarrow W$ existiert.

Bemerkung 24.8. (i) Wie üblich ist diese Relation reflexiv, symmetrisch und transitiv.

(ii) Wie bei euklidischen Vektorräumen zeigt man, dass jeder unitäre Vektorraum der Dimension $n < \infty$ zum $\mathbb{C}^n$ mit dem kanonischen hermiteschen Skalarprodukt isometrisch isomorph ist.

(iii) Daher sind je zwei unitäre Vektorräume der gleichen endlichen Dimension stets isometrisch isomorph.

25 Adjungierte Abbildungen

Satz 25.1. *Seien V, W endlich-dimensionale unitäre Vektorräume und $f \in \mathrm{Hom}\,(V, W)$. Dann existiert zu jedem $w \in W$ genau ein $f^\star(w) \in V$ mit*

$$(f^\star(w)|v) = (w|f(v)) \qquad \textit{für alle } v \in V.$$

Die so definierte Abbildung $f^\star : W \longrightarrow V$ ist auch linear.

Beweis. Seien $x_1, x_2 \in V$ mit $(x_1|v) = (w|f(v)) = (x_2|v)$ für alle $v \in V$. Dann ist

$$(x_1 - x_2|x_1 - x_2) = (x_1|x_1 - x_2) - (x_2|x_1 - x_2) = 0,$$

also $x_1 - x_2 = 0$ und $x_1 = x_2$. Dies beweist die Eindeutigkeit von $f^\star(w)$.
Zum Beweis der Existenz sei $a_1, \ldots, a_m$ eine Orthonormalbasis von V. Wir setzen $f^\star(w) := \sum_{k=1}^m (w|f(a_k))a_k$. Für $l = 1, \ldots, m$ gilt dann:

$$(f^\star(w)|a_l) = \sum_{k=1}^m (w|f(a_k))(a_k|a_l) = (w|f(a_l)).$$

Daraus folgt leicht: $(f^\star(w)|v) = (w|f(v))$ für alle $v \in V$. Die Linearität von $f^\star$ ist nach Definition klar. □

Definition 25.1. Man nennt $f^\star$ die zu f ***adjungierte*** (lineare) Abbildung.

Bemerkung 25.1. (i) Für $f_1, f_2 \in \mathrm{Hom}\,(V,W)$ und $z_1, z_2 \in \mathbb{C}$ gilt:

$$(z_1 f_1 + z_2 f_2)^\star = \overline{z_1} f_1^\star + \overline{z_2} f_2^\star;$$

denn für alle $v \in V, w \in W$ gilt:

$$\begin{aligned}((z_1 f_1 + z_2 f_2)^\star(w)|v) &= (w|(z_1 f_1 + z_2 f_2)(v)) = (w|z_1 f_1(v) + z_2 f_2(v)) \\ &= \overline{z_1}(w|f_1(v)) + \overline{z_2}(w|f_2(v)) = \overline{z_1}(f_1^\star(w)|v) + \overline{z_2}(f_2^\star(w)|v) \\ &= (\overline{z_1} f_1^\star(w) + \overline{z_2} f_2^\star(w)|v) = ((\overline{z_1} f_1^\star + \overline{z_2} f_2^\star)(w)|v).\end{aligned}$$

(ii) Sei $f \in \mathrm{Hom}\,(V,W)$ mit Matrix $A = (a_{ij}) \in \mathbb{C}^{n\times m}$ bzgl. Orthonormalbasen $a_1, \ldots, a_m$ von V und $b_1, \ldots, b_n$ von W. Dann ist $\overline{A}^T$ die Matrix von $f^\star$ bzgl. $b_1, \ldots, b_n$ und $a_1, \ldots, a_m$; schreibt man nämlich $f^\star(b_j) = \sum_{i=1}^m x_{ij} a_i$ mit $x_{ij} \in \mathbb{C}$, so gilt für alle j, k:

$$\begin{aligned}x_{kj} &= \sum_{i=1}^m x_{ij}(a_i|a_k) = (f^\star(b_j)|a_k) = (b_j|f(a_k)) = \left(b_j \Big| \sum_{l=1}^n a_{lk} b_l\right) \\ &= \sum_{l=1}^n \overline{a_{lk}}(b_j|b_l) = \overline{a_{jk}}.\end{aligned}$$

Definition 25.2. Seien V ein endlich-dimensionaler unitärer Vektorraum und $f \in \mathrm{End}\,(V)$ mit $f^\star \circ f = f \circ f^\star$. Dann nennt man f ***normal***.

Satz 25.2. *Seien V ein endlich-dimensionaler unitärer Vektorraum und $f \in \mathrm{End}\,(V)$. Genau dann ist f normal, wenn V eine Orthonormalbasis besitzt, die aus Eigenvektoren von f besteht. Ggf. ist also die Matrix von f bzgl. dieser Basis eine Diagonalmatrix.*

Beweis. "$\Rightarrow$": Sei f normal. Nach dem Fundamentalsatz der Algebra besitzt f einen Eigenwert $c \in \mathbb{C}$. Sei b ein entsprechender Eigenvektor. Wir können $\|b\| = 1$ annehmen. Dann gilt:

$$\begin{aligned}(f^\star(b) - \overline{c}b|f^\star(b) - \overline{c}b) &= (f^\star(b)|f^\star(b)) - \overline{c}(b|f^\star(b)) - c(f^\star(b)|b) + c\overline{c}(b|b) \\ &= (b|f(f^\star(b))) - (b|f^\star(cb)) - c(b|f(b)) + c(b|cb) \\ &= (b|f^\star(f(b)) - f^\star(f(b))) = 0,\end{aligned}$$

d.h. $f^\star(b)) = \overline{c}b$. Insbesondere ist $f(\mathbb{C}b) = \mathbb{C}b$ und $f^\star(\mathbb{C}b) \subseteq \mathbb{C}b$. Dann ist $V = \mathbb{C}b \oplus W$ mit $W := (\mathbb{C}b)^\perp$. Wie üblich zeigt man: $f(W) \subseteq W$ und $f^\star(W) \subseteq W$. Nun ist W ein unitärer Vektorraum, und für die Einschränkung $f|W : W \longrightarrow W$ gilt: $(f|W)^\star = f^\star|W$. Insbesondere ist $f|W \in \mathrm{End}\,(W)$ normal. Induktiv kann man annehmen, dass eine Orthonormalbasis $b_2, \ldots, b_n$ von W existiert, die aus Eigenvektoren von $f|W$ besteht. Dann ist $b_1 := b, b_2, \ldots, b_n$ eine Orthonormalbasis von V, die aus Eigenvektoren von f besteht.

"$\Leftarrow$": Umgekehrt sei $b_1, \ldots, b_n$ eine Orthonormalbasis von V, die aus Eigenvektoren von f besteht. Die Matrix A von f bzgl. $b_1, \ldots, b_n$ ist dann eine Diagonalmatrix. Folglich ist auch $\overline{A}^T$ eine Diagonalmatrix; insbesondere ist $A\overline{A}^T = \overline{A}^T A$. Nach Bemerkung 25.1 ist $\overline{A}^T$ die Matrix von $f^\star$ bzgl. $b_1, \ldots, b_n$. Außerdem sind $A\overline{A}^T$ und $\overline{A}^T A$ die Matrizen von $f \circ f^\star$ bzw. $f^\star \circ f$ bzgl. $b_1, \ldots, b_n$. Daher ist $f \circ f^\star = f^\star \circ f$, d.h. f ist normal.

□

Beispiel 25.2. Die lineare Abbildung $f : \mathbb{C}^3 \longrightarrow \mathbb{C}^3, (x, y, z) \longmapsto (y + z, -x + 2z, -x - 2y)$ hat die folgende Matrix bzgl. der Standardbasis:

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

Wegen $\overline{A}^T = -A$ ist $A \cdot \overline{A}^T = -A^2 = \overline{A}^T A$. Fasst man $\mathbb{C}^3$ in üblicher Weise als unitären Vektorraum auf, so ist also $f \circ f^\star = f^\star \circ f$, d.h. f ist normal.
Man rechnet leicht nach, dass A die Eigenwerte 0, $\sqrt{6}i$, $-\sqrt{6}i$ mit zugehörigen Eigenvektoren

$$(2, -1, 1),\ (-\frac{2}{5} - \frac{\sqrt{6}}{5}i, \frac{1}{5} - \frac{2\sqrt{6}}{5}i, 1),\ (-\frac{2}{5} + \frac{\sqrt{6}}{5}i, \frac{1}{5} + \frac{2\sqrt{6}}{5}i, 1)$$

hat. Normiert man diese Vektoren noch, so erhält man eine Orthonormalbasis von $\mathbb{C}^3$, bzgl. der f die folgende Matrix hat:

$$D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sqrt{6}i & 0 \\ 0 & 0 & -\sqrt{6}i \end{pmatrix}.$$

Definition 25.3. Eine Matrix $A \in \mathbb{C}^{n \times n}$ mit $A\overline{A}^T = \overline{A}^T A$ nennt man ***normal***.

Satz 25.3. *Eine Matrix $A \in \mathbb{C}^{n \times n}$ ist genau dann normal, wenn eine Matrix $S \in U(n, \mathbb{C})$ existiert mit der Eigenschaft, dass $S^{-1}AS$ eine Diagonalmatrix ist.*

Beweis. Routine □

Definition 25.4. Es seien V ein unitärer Vektorraum und $f \in \operatorname{End}(V)$ mit $(f(x)|y) = (x|f(y))$ für alle $x, y \in V$. Dann nennt man f ***selbstadjungiert***.

Bemerkung 25.4. Im Fall $\dim V < \infty$ ist f genau dann selbstadjungiert, wenn $f^\star = f$ gilt. Ggf. ist $f^\star \circ f = f^2 = f \circ f^\star$, d.h. f ist auch normal.

Satz 25.4. *Seien V ein endlich-dimensionaler unitärer Vektorraum und $f \in \operatorname{End}(V)$. Genau dann ist f selbstadjungiert, wenn V eine Orthonormalbasis bestitz, die aus Eigenvektoren von f zu reellen Eigenwerten besteht.*

Beweis. "⇒": Sei f selbstadjungiert, also auch normal. Nach Satz 25.2 existiert eine Orthonormalbasis $b_1, \ldots, b_n$ von V, die aus Eigenvektoren von f besteht. Die Matrix $A = (a_{ij})$ von f bzgl. $b_1, \ldots, b_n$ ist also eine Diagonalmatrix. Wegen $f^\star = f$ ist $\overline{A}^T = A$; insbesondere gilt: $a_{11}, a_{22}, \ldots, a_{nn} \in \mathbb{R}$. Diese Zahlen sind aber gerade die Eigenwerte von f.

"⇐": Es sei $b_1, \ldots, b_n$ eine Orthonormalbasis von V, die aus Eigenvektoren von f zu reellen Eigenwerten besteht. Die Matrix A von f bzgl. $b_1, \ldots, b_n$ ist also eine Diagonalmatrix mit reellen Koeffizienten. Daher ist $\overline{A}^T = A$, d.h. $f^\star = f$.

□

Satz 25.5. *Eine Matrix $A \in \mathbb{C}^{n\times n}$ ist genau dann hermitesch, wenn eine Matrix $S \in U(n, \mathbb{C})$ existiert mit der Eigenschaft, dass $S^{-1}AS$ eine Diagonalmatrix mit reellen Koeffizienten ist.*

Beweis. Routine; denn hermitesche Matrizen ($\overline{A}^T = A$) entsprechen selbstadjungierten linearen Abbildungen. ($f^\star = f$). □

Satz 25.6. *Es seien V ein endlich-dimensionaler unitärer Vektorraum und $f \in \mathrm{End}\,(V)$. Genau dann ist $f \in U(V)$, wenn eine Orthonormalbasis von V existiert, die aus Eigenvektoren von f zu Eigenwerten vom Betrag 1 besteht.*

Bemerkung 25.6. Es gilt $f \in U(V) \Leftrightarrow (x|y) = \underbrace{(f(x)|f(y))}_{=(f^\star(f(x))|y)}$ für alle $x, y \in V \Leftrightarrow f^\star(f(x)) = x$ für alle $x \in V \Leftrightarrow f$ bijektiv und $f^\star = f^{-1}$.
Ggf. ist $f \circ f^\star = \mathrm{id}_V = f^\star \circ f$, d.h. f ist auch normal.

Beweis. "⇒": Sei $f \in U(V)$, also auch f normal. Nach Satz 25.2 existiert eine Orthonormalbasis $b_1, \ldots, b_n$ von V, die aus Eigenvektoren von f besteht. Die Matrix A von f bzgl. $b_1, \ldots, b_n$ ist also eine Diagonalmatrix. Wegen $f^\star \circ f = \mathrm{id}_V$ ist $\overline{A}^T A = 1_n$. Daher haben die Koeffizienten auf der Hauptdiagonalen von A den Betrag 1. Diese sind aber gerade die Eigenwerte von f.

"⇐": Sei $b_1, \ldots, b_n$ eine Orthonormalbasis von V, die aus Eigenvektoren von f zu Eigenwerten vom Betrag 1 besteht. Die Matrix A von f bzgl. $b_1, \ldots, b_n$ ist dann eine Diagonalmatrix mit $\overline{A}^T A = 1_n$. Daher ist $f^\star \circ f = \mathrm{id}_V$, d.h. $f \in U(V)$.

□

Satz 25.7. *Eine Matrix $A \in \mathbb{C}^{n\times n}$ ist genau dann unitär, wenn eine Matrix $S \in U(n, \mathbb{C})$ existiert mit der Eigenschaft, dass $S^{-1}AS$ eine Diagonalmatrix mit Koeffizienten vom Betrag 1 auf der Hauptdiagonalen ist.*

Beweis. Routine. □

Bemerkung 25.8. Gegeben sei ein unitärer Vektorraum mit Basis $a_1, \ldots, a_n$. Wie in euklidischen Vektorräumen kann man aus $a_1, \ldots, a_n$ eine Orthonormalbasis $b_1, \ldots, b_n$ von V mit

$$\text{Span}\,(b_1, \ldots, b_k) = \text{Span}\,(a_1, \ldots, a_k) \text{ für } k = 1, \ldots, n$$

konstruieren (***Gram-Schmidt-Verfahren***).

Satz 25.8. *(Schur 1875-1941)*
Es seien V ein endlich-dimensionaler unitärer Vektorraum und $f \in \text{End}\,(V)$. Dann existiert eine Orthonormalbasis von V, bzgl. der die Matrix von f eine untere Dreiecksmatrix ist:

$$\begin{pmatrix} \star & & 0 \\ \vdots & \ddots & \\ \star & \ldots & \star \end{pmatrix}.$$

Beweis. Wir wählen eine Basis $a_1, \ldots, a_n$ von V, bzgl. der die Matrix J von f Jordansche Normalform hat; insbesondere ist J eine untere Dreiecksmatrix, d.h.

$$f(a_k) \in \text{Span}\,(a_k, \ldots, a_n) \text{ für } k = 1, \ldots, n.$$

Das Gram-Schmidt-Verfahren macht aus $a_n, a_{n-1}, \ldots, a_1$ eine Orthonormalbasis $b_n, b_{n-1}, \ldots, b_1$ von V mit $\text{Span}\,(b_n, \ldots, b_k) = \text{Span}\,(a_n, \ldots, a_k)$ für $k = 1, \ldots, n$. Daher ist

$$f(b_k) \in f(\text{Span}\,(a_n, \ldots, a_k)) \subseteq \text{Span}\,(a_n, \ldots, a_k) = \text{Span}\,(b_n, \ldots, b_k)$$

für $k = 1, \ldots, n$. Die Matrix von f bzgl. $b_1, \ldots, b_n$ ist also eine untere Dreiecksmatrix. □

Beispiel 25.8. $V = \mathbb{C}^3$ mit kanonischem hermiteschen Skalarprodukt
$f : \mathbb{C}^3 \longrightarrow \mathbb{C}^3, (x, y, z) \longmapsto (2z, x + y - z, -x + 3z)$ linear
Matrix von f bzgl. Standardbasis: $A = \begin{pmatrix} 0 & 0 & 2 \\ 1 & 1 & -1 \\ -1 & 0 & 3 \end{pmatrix}$
charakteristisches Polynom: $(X - 2)(X - 1)^2$
Eigenwerte: $2, 1$
Eigenräume $E_2(f) = \mathbb{C}(1, 0, 1), E_1(f) = \mathbb{C}(0, 1, 0)$
$\text{Ker}\,(f - \text{id}_V)^2 = \mathbb{C}(0, 1, 0) + \mathbb{C}(2, 0, 1)$
Basis von V: $a_1 = (1, 0, 1), a_2 = (2, 0, 1), a_3 = (0, 1, 0)$
Matrix von f bzgl. a_1, a_2, a_3: $J = \left(\begin{array}{c|cc} 2 & 0 & 0 \\ \hline 0 & 1 & 0 \\ 0 & 1 & 1 \end{array}\right)$ Jordansche Normalform
Gram-Schmidt-Verfahren: $b_3 = (0, 1, 0), b_2 = \frac{1}{\sqrt{5}}(2, 0, 1), b_1 = \frac{1}{\sqrt{5}}(-1, 0, 2)$ Orthonormalbasis von V
Matrix von f bzgl. b_1, b_2, b_3: $B = \begin{pmatrix} 2 & 0 & 0 \\ 3 & 1 & 0 \\ -\frac{3}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 1 \end{pmatrix}$
(Probe!)

Satz 25.9. *Zu jeder Matrix $A \in \mathbb{C}^{n\times n}$ existiert eine Matrix $S \in U(n,\mathbb{C})$ mit der Eigenschaft, dass $S^{-1}AS$ eine untere Dreiecksmatrix ist.*

Beweis. Wie üblich. □

26 Untergruppen, Nebenklassen, Normalteiler

Bemerkung 26.1. Wir wissen, dass eine ***Gruppe*** eine Menge mit einer assoziativen Verknüpfung ist, die (genau) ein ***neutrales Element*** $e = 1$ [$1g = g = g1$ für alle $g \in G$] enthält und zu jedem $g \in G$ ein ***inverses Element*** g^{-1} [$gg^{-1} = 1 = g^{-1}g$]. Die ***Ordnung*** $|G|$ ist die Anzahl der Elemente in G. Ist $ab = ba$ für alle $a, b \in G$, so heißt G ***abelsch*** oder ***kommutativ***.

Definition 26.1. Eine nichtleere Teilmenge H einer Gruppe G nennt man ***Untergruppe*** von G, falls gilt:

(i) $a, b \in H \Rightarrow ab \in H$.

(ii) $a \in H \Rightarrow a^{-1} \in H$.

Man schreibt dann $H \leq G$, im Fall $H \neq G$ auch $H < G$.

Bemerkung 26.2. (i) Sei $H \leq G$. Dann existiert ein Element $a \in H$. Nach Definition ist $a^{-1} \in H$ und $1 = aa^{-1} \in H$. Daher ist H selbst eine Gruppe mit der entsprechend eingeschränkten Verknüpfung, und die neutralen Elemente von G und H stimmen überein.

(ii) Eine nichtleere Teilmenge H einer Gruppe G ist genau dann eine Untergruppe von G, wenn gilt:

$$(\star) \qquad a, b \in H \Rightarrow ab^{-1} \in H.$$

Dies ist eine leichte Übung.

(iii) Für Untergruppen H, K von G ist offensichtlich $H \cap K \leq G$. Allgemeiner ist $\bigcap_{i\in I} H_i \leq G$ für jede Familie $(H_i)_{i\in I}$ von Untergruppen H_i von G.

Beispiel 26.2. (i) $(\mathbb{Z}, +) \leq (\mathbb{Q}, +) \leq (\mathbb{R}, +) \leq (\mathbb{C}, +)$.

(ii) $(\mathbb{Q} \setminus \{0\}, \cdot) \leq (\mathbb{R} \setminus \{0\}, \cdot) \leq (\mathbb{C} \setminus \{0\}, \cdot)$.

(iii) $n \in \mathbb{N} \Rightarrow O(n, \mathbb{R}) \leq \mathrm{GL}(n, \mathbb{R}), U(n, \mathbb{C}) \leq \mathrm{GL}(n, \mathbb{C})$.

(iv) Für jeden Körper K und jeden K-Vektorraum V ist $\mathrm{GL}(V) \leq \mathrm{Sym}(V)$. [Für jede Menge X ist $\mathrm{Sym}(X)$ die ***symmetrische Gruppe*** auf X, d.h. die Gruppe aller bijektiven Abbildungen (***Permutationen***) $g : X \longrightarrow X$]

(v) Für jede Gruppe G ist $\{1\} \leq G$ und $G \leq G$. Man nennt $1 := \{1\}$ die ***triviale*** Untergruppe von G. Untergruppen $H < G$ nennt man ***echte*** Untergruppen von G.

(vi) Für jede Gruppe G und jedes $a \in G$ ist

$$\langle a \rangle := \{a^n : n \in \mathbb{Z}\} \leq G;$$

man nennt $\langle a \rangle$ die von a ***erzeugte zyklische*** Untergruppe von G. Im Fall $G = (\mathbb{Z}, +)$ und $a = 5$ ist also

$$\langle a \rangle = \{0, \pm 5, \pm 10, \dots\} = 5\mathbb{Z}.$$

Im Fall $G = \mathrm{Sym}\,(3) := \mathrm{Sym}\,(\{1,2,3\})$ und $a = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$ ist $a^2 = 1, a^3 = a, \dots$ und $a = a^{-1}$, d.h. $\langle a \rangle = \{1, a\}$.

Definition 26.3. Seien G eine Gruppe und $H \leq G$. Für $a \in G$ nennt man $aH := \{ah : h \in H\}$ die ***Linksnebenklasse*** und $Ha := \{ha : h \in H\}$ die ***Rechtsnebenklasse*** von a nach H. Mit $G/H := \{aH : a \in G\}$ und $H\backslash G := \{Ha : a \in G\}$ bezeichnen wir die Menge aller Linksnebenklassen bzw. Rechtsnebenklassen von G nach H.

Bemerkung 26.3. (i) Für $a \in G$ ist

$$Ha^{-1} = \{ha^{-1} : h \in H\} = \{k^{-1}a^{-1} : k \in H\} = \{(ak)^{-1} : k \in H\}.$$

Für jede Linksnebenklasse X nach H ist also $X^{-1} := \{x^{-1} : x \in X\}$ eine Rechtsnebenklasse nach H. Die Anzahl aller Linksnebenklassen nach H in G stimmt also mit der Anzahl aller Rechtsnebenklassen nach H in G überein. Diese bezeichnet man als ***Index*** von H in G. Man schreibt dafür $|G : H|$.

(ii) Für $a, b \in G$ gilt:

$$aH \cap bH \neq \emptyset \Leftrightarrow aH = bH \Leftrightarrow a^{-1}b \in H;$$

Zum Beweis sei zunächst $c \in aH \cap bH$. Wir schreiben $c = ah_0$ mit $h_0 \in H$. Für $h \in H$ ist dann $ch = ah_0h \in aH$ und $ah = ch_0^{-1}h \in cH$. Also ist $cH = aH$ und analog $cH = bH$, d.h. $aH = bH$.
Sei jetzt $aH = bH$. Dann ist $a = a \cdot 1 \in aH = bH$, also $a = bh$ für ein $h \in H$. Daher ist $a^{-1}b = h^{-1} \in H$.
Sei jetzt $a^{-1}b \in H$. Dann ist $b = a \cdot a^{-1}b \in aH \cap bH$.

(iii) Für $a \in G$ ist die Abbildung

$$H \longrightarrow aH, h \longmapsto ah,$$

bijektiv; insbesondere ist $|aH| = |H|$.

Satz 26.3. *(Lagrange 1736-1813)*
Für jede Untergruppe H einer Gruppe G gilt:

$$|G| = |G : H| \cdot |H|;$$

insbesondere sind $|H|$ und $|G : H|$ im Fall $|G| < \infty$ Teiler von $|G|$.

Beweis. Nach den obigen Bemerkungen liegt jedes Element aus G in genau einer Linksnebenklasse nach H, und jede Linksnebenklasse nach H enthält genau $|H|$ Elemente. □

Beispiel 26.3. (i) $G = \mathrm{Sym}\,(3), H = \langle a \rangle$ mit $a = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$.
Dann ist $|G| = 3! = 6$ und $|H| = 2$, also $|G : H| = 3$. (Probe!)

(ii) $G = \mathbb{Z}, H = 5\mathbb{Z}$. Die Linksnebenklassen sind

$$\begin{aligned} 0 + 5\mathbb{Z} &= \{\dots, -10, -5, 0, 5, 10, \dots\} \\ 1 + 5\mathbb{Z} &= \{\dots, -9, -4, 1, 6, 11, \dots\} \\ 2 + 5\mathbb{Z} &= \{\dots, -8, -3, 2, 7, 12, \dots\} \\ 3 + 5\mathbb{Z} &= \{\dots, -7, -2, 3, 8, 13, \dots\} \\ 4 + 5\mathbb{Z} &= \{\dots, -6, -1, 4, 9, 14, \dots\} \end{aligned}$$

Daher ist $|\mathbb{Z} : 5\mathbb{Z}| = 5$. Analog ist $|\mathbb{Z} : n\mathbb{Z}| = n$ für $n \in \mathbb{N}$.

Satz 26.4. *Für eine Untergruppe H einer Gruppe G sind äquivalent:*

(1) $aH = Ha$ für alle $a \in G$.

(2) $aHa^{-1} = H$ für alle $a \in G$.

(3) $aHa^{-1} \subseteq H$ für alle $a \in G$.

Beweis.

(1) $\Leftrightarrow$ (2) $\Rightarrow$ (3): klar.

(3) $\Rightarrow$ (2): Sei (3) erfüllt und $a \in G$. Dann gilt:

$$H = a \underbrace{a^{-1} H (a^{-1})^{-1}}_{\subseteq H} a^{-1} \subseteq aHa^{-1}.$$

□

Definition 26.4. Ggf. nennt man H ***normal*** in G oder ***Normalteiler*** von G. Man schreibt $H \trianglelefteq G$, im Fall $H \neq G$ auch $H \triangleleft G$.

Beispiel 26.4. (i) Stets ist $1 \trianglelefteq G$ und $G \trianglelefteq G$.

(ii) Aus $H \trianglelefteq G$ und $K \trianglelefteq G$ folgt $H \cap K \trianglelefteq G$; denn für $a \in G$ gilt:

$$a(H \cap K) = aH \cap aK = Ha \cap Ka = (H \cap K)a.$$

Allgemeiner ist $\bigcap_{i \in I} H_i \trianglelefteq G$ für jede Familie $(H_i)_{i \in I}$ von Normalteilern H_i von G.

(iii) In einer abelschen Gruppe ist jede Untergruppe normal.

(iv) Jede Untergruppe H von G mit $|G : H| = 2$ ist normal in G; denn die einzigen Linksnebenklassen (Rechtsnebenklassen) nach H in G sind H und $G \setminus H$.

(v) $G = \mathrm{Sym}\,(3), H = \langle b \rangle$ mit $b = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$.
Dann ist $H = \{1, b, b^2\}$ (nachrechnen!), also $|H| = 3$ und $|G : H| = 2$ nach Lagrange. Folglich ist $H \trianglelefteq G$.

(vi) $G = \mathrm{Sym}\,(3), H = \langle a \rangle$ mit $a = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$. Dann ist $H \not\trianglelefteq G$ wegen

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} H = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right\} \neq \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \right\} = H \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

Satz 26.5. *Für jeden Normalteiler N einer Gruppe G wird die Menge G/N aller (Links-) Nebenklassen von G nach N zu einer Gruppe, wenn man definiert:*

$$(aN)(bN) := abN \qquad \text{für } a, b \in G.$$

Definition 26.5. Man nennt G/N die ***Faktorgruppe*** von G nach N.

Beweis. Wir zeigen zunächst, dass die Verknüpfung in G/N "wohldefiniert" ist, d.h. nicht von der Schreibweise der Nebenklassen abhängt. Dazu seien $a, a', b, b' \in G$ mit $aN = a'N$ und $bN = b'N$. Dann ist $a^{-1}a' \in N$, also auch $b^{-1}a^{-1}a'b \in N$. Folglich ist $abN = a'bN$ und analog $a'bN = a'b'N$. Offensichtlich ist die Verknüpfung in G/N assoziativ, neutrales Element ist $1N = N$, und $(aN)^{-1} = a^{-1}N$ für $a \in G$. □

Beispiel 26.5. $G = \mathbb{Z}, H = 5\mathbb{Z}$. Dann ist

$$G/H = \mathbb{Z}/5\mathbb{Z} = \{0 + 5\mathbb{Z}, 1 + 5\mathbb{Z}, 2 + 5\mathbb{Z}, 3 + 5\mathbb{Z}, 4 + 5\mathbb{Z}\}$$

eine Gruppe der Ordnung 5 bzgl. +. Z.B. ist

$$(3 + 5\mathbb{Z}) + (4 + 5\mathbb{Z}) = 7 + 5\mathbb{Z} = 2 + 5\mathbb{Z}.$$

Definition 26.6. Seien G, H Gruppen. Eine Abbildung $f : G \longrightarrow H$ mit $f(ab) = f(a)f(b)$ für alle $a, b \in G$ nennt man ***Homomorphismus***.

Bemerkung 26.6. (i) Ggf. ist

$$f(1_G) = f(1_G)1_H = f(1_G)f(1_G)f(1_G)^{-1} = f(1_G \cdot 1_G)f(1_G)^{-1} = 1_H.$$

Ferner gilt für alle $g \in G$:

$$\begin{aligned} f(g^{-1}) = f(g^{-1})1_H = f(g^{-1})f(g)f(g)^{-1} = f(g^{-1}g)f(g)^{-1} \\ = f(1_G)f(g)^{-1} = 1_H f(g)^{-1} = f(g)^{-1}. \end{aligned}$$

(ii) Sind G, H, K Gruppen und $f : G \longrightarrow H, g : H \longrightarrow K$ Homomorphismen, so ist auch $g \circ f : G \longrightarrow K$ ein Homomorphismus; denn für $a, b \in G$ gilt:

$$(g \circ f)(ab) = g(f(ab)) = g(f(a)f(b)) = g(f(a))g(f(b)) = (g \circ f)(a) \cdot (g \circ f)(b).$$

Beispiel 26.6. (i) Für jede Gruppe G ist die konstante Abbildung $G \longrightarrow G, a \longmapsto 1$ ein Homomorphismus.

(ii) Die Abbildung $f : (\mathbb{R}, +) \longrightarrow (\mathbb{R} \setminus \{0\}, \cdot), x \longmapsto 2^x$ ist wegen $f(x + y) = 2^{x+y} = 2^x 2^y = f(x)f(y)$ für $x, y \in \mathbb{R}$ ein Homomorphismus.

(iii) Für $n \in \mathbb{N}$ und jeden Körper K ist die Abbildung

$$\mathrm{GL}\,(n, K) \longrightarrow K \setminus \{0\}, A \longmapsto \det A$$

ein Homomorphismus.

Satz 26.6. *Für jeden Homomorphismus von Gruppen $f : G \longrightarrow H$ gilt:*

(i) Ist $U \leq G$, so ist $f(U) \leq H$; insbesondere ist $\mathrm{Bld}\,(f) = f(G) \leq H$.

(ii) Ist $U \lhd G$, so ist $f(U) \lhd f(G)$, aber nicht unbedingt $f(U) \lhd H$.

(iii) Ist $V \leq H$, so ist $f^{-1}(V) \leq G$.

(iv) Ist $V \lhd H$, so ist $f^{-1}(V) \lhd G$.

Beweis. (i) Wegen $U \neq \emptyset$ ist $f(U) \neq \emptyset$. Seien $x, y \in f(U)$. Wir schreiben $x = f(a), y = f(b)$ mit $a, b \in U$. Dann gilt:

$$xy^{-1} = f(a)f(b)^{-1} = f(a)f(b^{-1}) = f(\underbrace{ab^{-1}}_{\in U}) \in f(U).$$

(ii) Seien $y \in f(G)$ und $y = f(x)$ mit $x \in G$. Dann gilt für $u \in U$:

$$yf(u)y^{-1} = f(x)f(u)f(x)^{-1} = f(\underbrace{xux^{-1}}_{\in U}) \in f(U).$$

Also ist $f(U) \lhd f(G)$.

Seien $G = \left\langle \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \right\rangle, H = \mathrm{Sym}\,(3)$ und $f : G \longrightarrow H$ die Inklusionsabbildung. Dann ist $G \lhd G$, aber $f(G) = G \not\lhd H$.

(iii) Wegen $f(1_G) = 1_H \in V$ ist $1_G \in f^{-1}(V)$, also $f^{-1}(V) \neq \emptyset$. Seien $a, b \in f^{-1}(V)$, d.h. $f(a), f(b) \in V$. Dann ist $f(ab^{-1}) = f(a)f(b)^{-1} \in V$, d.h. $ab^{-1} \in f^{-1}(V)$.

(iv) Seien $a \in G$ und $b \in f^{-1}(V)$, d.h. $f(b) \in V$. Dann ist $f(aba^{-1}) = f(a)f(b)f(a)^{-1} \in V$, d.h. $aba^{-1} \in f^{-1}(V)$.

□

Bemerkung 26.7. Für jeden Homomorphismus von Gruppen $f : G \longrightarrow H$ ist

$$\mathrm{Ker}\,(f) := f^{-1}(\{1_H\}) = \{a \in G : f(a) = 1_H\} \lhd G;$$

man nennt $\mathrm{Ker}\,(f)$ den ***Kern*** von f.

Beispiel 26.7. (i) Für $n \in \mathbb{N}$ und jeden Körper K ist

$$\mathrm{SL}\,(n, K) = \{A \in \mathrm{GL}\,(n, K) : \det A = 1\} \lhd \mathrm{GL}\,(n, K).$$

Man nennt $\mathrm{SL}\,(n, K)$ die ***spezielle lineare Gruppe*** des Grades n über K.

(ii) Analog ist

$$SO(n, \mathbb{R}) := SO(n) := \{A \in O(n) : \det A = 1\} \lhd O(n);$$

man nennt $SO(n)$ die ***spezielle orthogonale Gruppe*** des Grades n (über $\mathbb{R}$).

(iii) Analog ist

$$SU(n, \mathbb{C}) := SU(n) := \{A \in U(n) : \det A = 1\} \lhd U(n);$$

man nennt $SU(n, \mathbb{C})$ die ***spezielle unitäre Gruppe*** des Grades n (über $\mathbb{C}$).

Satz 26.7. *Ein Homomorphismus von Gruppen $f : G \longrightarrow H$ ist genau dann injektiv, wenn $\mathrm{Ker}\,(f) = \{1_G\}$ ist.*

Beweis. "⇒": Sei $f : G \longrightarrow H$ injektiv. Wegen $f(1_G) = 1_H$ ist $1_G \in \mathrm{Ker}\,(f)$, d.h. $\{1_G\} \subseteq \mathrm{Ker}\,(f)$.
Sei umgekehrt $x \in \mathrm{Ker}\,(f)$. Dann ist $f(x) = 1_H = f(1_G)$, also $x = 1_G$ wegen der Injektivität von f. Daher ist $\mathrm{Ker}\,(f) = \{1_G\}$.

"⇐": Sei $\mathrm{Ker}\,(f) = \{1_G\}$. Für Elemente $x, y \in G$ mit $f(x) = f(y)$ gilt dann: $1 = f(x)f(y)^{-1} = f(xy^{-1})$. Daher ist $xy^{-1} \in \mathrm{Ker}\,(f) = \{1_G\}$, d.h. $xy^{-1} = 1_G$ und $x = y$.

□

Definition 26.8. Einen injektiven Homomorphismus nennt man ***Monomorphismus***, einen surjektiven Homomorphismus ***Epimorphismus*** und einen bijektiven Homomorphismus ***Isomorphismus***.

Beispiel 26.8. (i) Für jede Untergruppe H einer Gruppe G ist die Inklusionsabbildung $H \longrightarrow G$ ein Monomorphismus.

(ii) Für jeden Normalteiler N von G ist die Abbildung

$$f : G \longrightarrow G/N, g \longmapsto gN,$$

ein Epimorphismus; man nennt f den ***kanonischen Epimorphismus*** von G auf G/N. Für $a \in G$ gilt:

$$a \in \mathrm{Ker}\,(f) \Leftrightarrow aN = 1N \Leftrightarrow a \in N.$$

Daher ist $\mathrm{Ker}\,(f) = N$.

(iii) Stets ist $\mathrm{id}_G : G \longrightarrow G$ ein Isomorphismus.

(iv) Ist $f : G \longrightarrow H$ ein Isomorphismus, so auch $f^{-1} : H \longrightarrow G$; denn für $a, b \in H$ gilt:

$$f^{-1}(ab) = f^{-1}(f(f^{-1}(a)) \cdot f(f^{-1}(b))) = f^{-1}(f(f^{-1}(a) \cdot f^{-1}(b))) = f^{-1}(a)f^{-1}(b).$$

Bemerkung 26.8. Man nennt Gruppen G, H ***isomorph*** und schreibt $G \cong H$, wenn ein Isomorphismus $f : G \longrightarrow H$ existiert. Wie üblich ist diese Relation reflexiv, symmetrisch und transitiv.

Beispiel 26.9. (i) Seien K ein Körper und V ein K-Vektorraum mit Basis $b_1, \ldots, b_n$. Dann ist die Abbildung

$$F : \mathrm{GL}\,(V) \longrightarrow \mathrm{GL}\,(n, K),$$

die jedem Vektorraum-Isomorphismus $f : V \longrightarrow V$ seine Matrix bzgl. $b_1, \ldots, b_n$ zuordnet, ein Gruppenisomorphismus; insbesondere gilt:

$$\mathrm{GL}\,(V) \cong \mathrm{GL}\,(n, K).$$

(ii) Analog gilt für jeden euklidischen Vektorraum V der Dimension $n < \infty$:

$$O(V) \cong O(n).$$

(iii) Analog gilt für jeden unitären Vektorraum V der Dimension $n < \infty$:

$$U(V) \cong U(n).$$

Satz 26.9. *(Homomorphiesatz)*
Für jeden Homomorphismus von Gruppen $f : G \longrightarrow H$ ist die Abbildung

$$F : G/\mathrm{Ker}\,(f) \longrightarrow \mathrm{Bld}\,(f),\ a\,\mathrm{Ker}\,(f) \longmapsto f(a),$$

ein Isomorphismus von Gruppen; insbesondere gilt:

$$G/\mathrm{Ker}\,(f) \cong \mathrm{Bld}\,(f).$$

Beweis. Wir zeigen zunächst, dass F wohldefiniert ist; dazu seien $a, a' \in G$ mit $a\,\mathrm{Ker}\,(f) = a'\,\mathrm{Ker}\,(f)$, d.h. $a^{-1}a' \in \mathrm{Ker}\,(f)$. Dann ist $1 = f(a^{-1}a') = f(a)^{-1}f(a')$, d.h. $f(a') = f(a)$.
Für $a, b \in G$ ist

$$F(a\,\mathrm{Ker}\,(f) \cdot b\,\mathrm{Ker}\,(f)) = F(ab\,\mathrm{Ker}\,(f) = f(ab) = f(a)f(b) = F(a\,\mathrm{Ker}\,(f))F(b\,\mathrm{Ker}\,(f)).$$

Daher ist F ein Homomorphismus.
Seien $a, b \in G$ mit $F(a\,\mathrm{Ker}\,(f)) = F(b\,\mathrm{Ker}\,(f))$, d.h. $f(a) = f(b)$. Dann ist $f(a^{-1}b) = f(a)^{-1}f(b) = 1$, d.h. $a^{-1}b \in \mathrm{Ker}\,(f)$ und $a\,\mathrm{Ker}\,(f) = b\,\mathrm{Ker}\,(f)$. Daher ist F injektiv.
Offensichtlich ist F auch surjektiv. □

Beispiel 26.10. (i) Für $n \in \mathbb{N}$ und jeden Körper K folgt aus dem Homomorphie-Satz:

$$\mathrm{GL}\,(n, K)/\,\mathrm{SL}\,(n, K) \cong K \setminus \{0\}.$$

(ii) Analog ist $O(n)/SO(n) \cong \{1, -1\}$; insbesondere ist $|O(n) : SO(n)| = 2$.

(iii) Analog ist $U(n)/SU(n) \cong \{z \in \mathbb{C} : |z| = 1\}$.

Definition 26.11. Ein ***Endomorphismus*** (bzw. ***Automorphismus***) einer Gruppe G ist ein Homomorphismus (bzw. Isomorphismus) $f : G \longrightarrow G$.

Bemerkung 26.11. Man zeigt leicht, dass

$$\mathrm{Aut}\,(G) := \{f : G \longrightarrow G | f \text{ Automorphismus}\}$$

eine Untergruppe von $\mathrm{Sym}\,(G)$ ist. Man nennt $\mathrm{Aut}\,(G)$ die ***Automorphismengruppe*** von G.

Beispiel 26.11. Für jedes Element x einer Gruppe G ist die Abbildung

$$f_x : G \longrightarrow G, a \longmapsto xax^{-1},$$

ein Automorphismus von G; denn für $a, b \in G$ gilt:

$$f_x(a)f_x(b) = xax^{-1}xbx^{-1} = xabx^{-1} = f_x(ab).$$

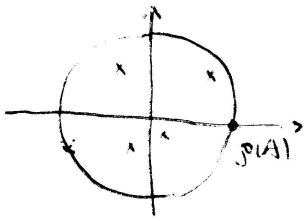
Aus $f_x(a) = f_x(b)$ folgt $xax^{-1} = xbx^{-1}$, also $a = b$. Ferner ist $f_x(x^{-1}cx) = xx^{-1}cxx^{-1} = c$ für $c \in G$.
Man nennt f_x den von x induzierten ***inneren*** Automorphismus von G.

27 Nichtnegative Matrizen

Definition 27.1. Für $A \in \mathbb{C}^{n\times n}$ nennt man

$$\rho(A) := \max\{|r| : r \text{ Eigenwert von } A\}$$

den ***Spektralradius*** von A.



Satz 27.1. *Für $A \in \mathbb{C}^{n\times n}$ mit $\rho(A) < 1$ gilt:*

(i) Die Folge $(A^k)_{k\in\mathbb{N}}$ konvergiert in $\mathbb{C}^{n\times n}$ gegen die Nullmatrix.

(ii) Die Reihe $\sum_{k=0}^{\infty} A_k$ konvergiert in $\mathbb{C}^{n\times n}$ gegen $(1_n - A)^{-1}$; insbesondere ist $1_n - A \in \mathrm{GL}(n, \mathbb{C})$.

Bemerkung 27.1. Bekanntlich sind alle Normen auf $\mathbb{C}^{n\times n}$ äquivalent; daher kommt es bei der Konvergenz nicht auf die Wahl der Norm an.

Beweis. (i) Sei $S \in \mathrm{GL}(n, \mathbb{C})$ derart, dass $J := S^{-1}AS$ Jordan-Normalform hat. Dann ist

$$\lim_{k\to\infty} A^k = \lim_{k\to\infty} (SJS^{-1})^k = \lim_{k\to\infty} SJ^kS^{-1} = S(\lim_{k\to\infty} J^k)S^{-1},$$

da die Multiplikation mit S und S^{-1} stetig ist. Wegen $\rho(A) = \rho(J)$ genügt also zu zeigen: $\lim_{k\to\infty} J^k = 0$. Dabei können wir annehmen, dass J selbst ein Jordan-Block ist:

$$J = \begin{pmatrix} r & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & r \end{pmatrix} \text{ mit } r \in \mathbb{C}, |r| < 1.$$

Für $k \geq n$ gilt dann:

$$J^k = \begin{pmatrix} r^k & & & & 0 \\ kr^{k-1} & \ddots & & & \\ \binom{k}{2} r^{k-2} & & \ddots & & \\ \vdots & & & \ddots & \\ \binom{k}{n-1} r^{k-n+1} & \cdots & \cdots & \cdots & r^k \end{pmatrix} \xrightarrow{k\to\infty} 0.$$

(ii) Beim Beweis der Konvergenz von $\sum_{k=0}^{\infty} A^k$ können wir wie in (i) annehmen, dass A Jordan-Normalform hat und sogar ein Jordan-Block ist. Dann müssen wir zeigen, dass für $|r| < 1$ die Reihen

$$\sum_{k=0}^{\infty} r^k, \sum_{k=1}^{\infty} kr^{k-1}, \sum_{k=2}^{\infty} \binom{k}{2} r^{k-2}, \ldots, \sum_{k=n}^{\infty} \binom{k}{n} r^{k-n}$$

konvergieren. Dies ist aber klar (Analysis!).
Also konvergiert die Reihe. Für ihren Grenzwert gilt wie üblich:

$$(1_n - A)\sum_{k=0}^{\infty} A^k = \sum_{k=0}^{\infty} A^k - \sum_{k=0}^{\infty} A^{k+1} = 1_n.$$

Folglich ist $\sum_{k=0}^{\infty} A^k = (1_n - A)^{-1}$.

□

Definition 27.2. Für $A = (a_{ij}), B = (b_{ij}) \in \mathbb{R}^{m\times n}$ schreiben wir

(i) $A \leq B$ (oder $B \geq A$), falls $a_{ij} \leq b_{ij}$ für alle i, j ist.

(ii) $A < B$ (oder $B > A$), falls $a_{ij} < b_{ij}$ für alle i, j ist.

Im Fall $A \geq 0$ nennt man A ***nichtnegativ***, im Fall $A > 0$ ***positiv***.

Beispiel 27.2. (i) Für beliebige $C = (c_{ij}) \in \mathbb{C}^{m\times n}$ ist $C^+ := (|c_{ij}|) \geq 0$.

(ii) Matrizen, die Übergangswahrscheinlichkeiten enthalten, sind stets nichtnegativ.

(iii) Matrizen, deren Koeffizienten gewisse Anzahlen sind, sind stets nichtnegativ.

Definition 27.3. Eine quadratische Matrix $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ heißt ***zerlegbar***, wenn eine disjunkte Zerlegung $\{1, \ldots, n\} = I \dot\cup J$ existiert mit $I \neq \emptyset \neq J$ und $a_{ij} = 0$ für alle $i \in I, j \in J$. Andernfalls heißt A ***unzerlegbar***.

Bemerkung 27.3. A ist also genau dann zerlegbar, wenn eine Permutationsmatrix $P \in \mathbb{C}^{n\times n}$ existiert derart, dass $P^{-1}AP$ die folgende Form hat:

$$P^{-1}AP = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}.$$

Dies ist genau dann der Fall, wenn eine Permutationsmatrix $Q \in \mathbb{C}^{n\times n}$ existiert derart, dass $Q^{-1}AQ$ die folgende Form hat:

$$Q^{-1}AQ = \begin{pmatrix} R & S \\ 0 & T \end{pmatrix}.$$

[Zum Beweis betrachte man $Q := P\begin{pmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{pmatrix}$.]
Daher ist A genau dann zerlegbar, wenn A^T zerlegbar ist.

Beispiel 27.3. $A \in \mathbb{R}^{n\times n}$ positiv $\Rightarrow A$ unzerlegbar.

Satz 27.3. *Sei $A \in \mathbb{R}^{n\times n}$ unzerlegbar mit $A \geq 0$. Dann ist $(1_n + A)^{n-1} > 0$.*

Beweis. Es genügt zu zeigen, dass $(1_n + A)^{n-1}y > 0$ für alle $y \in \mathbb{R}^{n\times 1}$ mit $0 \neq y \geq 0$ gilt; denn für $i = 1, \ldots, n$ ist $(1_n + A)^{n-1}e_i$ die i-te Spalte von $(1_n + A)^{n-1}$. [Dabei ist $e_1, \ldots, e_n$ die Standardbasis von $\mathbb{R}^{n\times 1}$.] Sei also $y \in \mathbb{R}^{n\times 1}$ mit $0 \neq y \geq 0$. Dann genügt es zu zeigen, dass $z := (1_n + A)y = y + Ay$ mehr von 0 verschiedene Koeffizienten enthält als y. Wir nehmen also an, dass y und $z = y + Ay$ gleich viele [und damit die gleichen] von 0 verschiedenen Koeffizienten haben. Nach Permutation der Koeffizienten können wir annehmen:

$$y = \begin{pmatrix} u \\ 0 \end{pmatrix}, z = \begin{pmatrix} v \\ 0 \end{pmatrix} \text{ mit } u > 0, v > 0.$$

Entsprechend schreiben wir:

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}.$$

Dann gilt:

$$\begin{pmatrix} v \\ 0 \end{pmatrix} = \begin{pmatrix} u \\ 0 \end{pmatrix} + \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} u \\ 0 \end{pmatrix} = \begin{pmatrix} u + A_{11}u \\ A_{21}u \end{pmatrix},$$

d.h. $A_{21}u = 0$. Wegen $u > 0$ folgt $A_{21} = 0$. Also ist A zerlegbar. Widerspruch. □

Satz 27.4. *(Perron-Frobenius)*
Sei $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ unzerlegbar mit $A \geq 0$. Dann ist $\rho(A)$ Eigenwert von A, der entsprechende Eigenraum ist 1-dimensional und wird von einem positiven Vektor aufgespannt.

Beweis. Es sei $M := \{x \in \mathbb{R}^{n\times 1} : 0 \neq x \geq 0\}$. Für $x = (x_1, \ldots, x_n)^T \in M$ setzen wir

$$r_x := \min\Big\{\frac{1}{x_i} \underbrace{\sum_{k=1}^{n} a_{ik}x_k}_{i\text{-ter Koeffizient von } Ax} : i = 1, \ldots, n, x_i \neq 0\Big\}.$$

Dann ist $r_x \geq 0$ und $r_x x_i \leq \sum_{k=1}^{n} a_{ik}x_k$ für $i = 1, \ldots, n$, d.h. $r_x x \leq Ax$. Genauer gilt:

$$r_x = \max\{r \in \mathbb{R} : rx \leq Ax\}.$$

Wir wollen zeigen, dass die Funktion

$$f : M \longrightarrow \mathbb{R}, x \longmapsto r_x,$$

ihr Maximum annimmt. Für $s \in \mathbb{R}$ mit $s > 0$ ist offenbar $r_{sx} = r_x$. Daher brauchen wir bei der Suche nach einem Maximum von f nur die Elemente in

$$N := \{x \in \mathbb{R}^{n\times 1} : x \geq 0 \text{ und } (x|x) = 1\}$$

zu betrachten; dabei ist $(\cdot|\cdot)$ das Standardskalarprodukt auf dem $\mathbb{R}^{n\times 1}$. Die Menge N ist kompakt, aber f ist nicht unbedingt stetig auf N. Daher kann man das übliche Argument

aus der Analysis nicht direkt anwenden. Als Bild von N unter einer stetigen Abbildung ist

$$P := \{(1+A)^{n-1}x : x \in N\}$$

auch kompakt, besteht aber nach Satz 27.3 aus lauter positiven Elementen. Für $x \in N$ ist $y := (1+A)^{n-1}x \in P$, und aus $r_x x \leq Ax$ folgt durch Multiplikation mit $(1_n + A)^{n-1} > 0$:

$$r_x y \leq Ay, \text{ d.h. } r_x \leq r_y.$$

Auf der Suche nach einem Maximum von f können wir uns also auf die Menge P beschränken. Da P aus lauter positiven Elementen besteht, ist f auf P stetig, nimmt also dort ihr Maximum r an. Die Elemente $z \in M$ mit $r_z = r$ nennen wir ***extremal***. Wir behaupten:

(a) $r > 0$, und r ist Eigenwert von A.

(b) Jedes extremale $z \in M$ ist positiv und Eigenvektor von A zum Eigenwert r.

Zum Beweis sei $u := (1, \ldots, 1)^T$. Dann ist $r_u = \min\{\sum_{k=1}^n a_{ik} : i = 1, \ldots, n\}$. Da A keine Nullzeile enthalten kann (sonst wäre A zerlegbar) ist $r_u > 0$, also auch $r \geq r_u > 0$.
Sei jetzt $z \in M$ extremal. Dann ist $x := (1_n + A)^{n-1}z > 0$.
Annahme: $Az \neq rz$. Dann ist $Az - rz \geq 0$, also $(1_n + A)^{n-1}(Az - rz) > 0$, d.h. $Ax - rx > 0$.
Daher existiert ein $\epsilon > 0$ mit $(r+\epsilon)x < Ax$, und wir haben den Widerspruch $r_x \geq r + \epsilon > r$.
Also ist $Az = rz$, d.h. z ist Eigenvektor von A zum Eigenwert r. Daher gilt:

$$0 < x = (1_n + A)^{n-1}z = (1+r)^{n-1}z.$$

Folglich ist auch $z > 0$, und (a) und (b) sind bewiesen.
Als nächstes zeigen wir: $r = \rho(A)$.
Zum Beweis sei $s \in \mathbb{C}$ ein beliebiger Eigenwert von A und $y \in \mathbb{C}^{n\times 1}$ ein entsprechender Eigenvektor, d.h. $Ay = sy$. In dieser Gleichung gehen wir zu den Beträgen über und erhalten $|s|y^+ \leq Ay^+$. Folglich ist $|s| \leq r_{y^+} \leq r$.
Wir müssen noch zeigen, dass der Eigenraum von A zum Eigenwert r eindimensional ist. Sei also $0 \neq y \in \mathbb{C}^{n\times 1}$ mit $Ay = ry$. Die obigen Überlegungen zeigen, daß $y^+ \in M$ extremal ist. Nach (b) ist $y^+ > 0$, d.h. $y = (y_1, \ldots, y_n)^T$ mit $y_i \neq 0$ für $i = 1, \ldots, n$.
Sei auch $0 \neq y' = (y'_1, \ldots, y'_n) \in \mathbb{C}^{n\times 1}$ mit $Ay' = ry'$. Dann ist auch $y'' := y' - \frac{y'_1}{y_1}y \in \mathbb{C}^{n\times 1}$ mit $Ay'' = ry''$. Wegen $y'' = (0, \star, \ldots, \star)^T$ zeigen die obigen Überlegungen: $y'' = 0$, d.h. $y' \in \mathbb{C}y$. Damit ist der Satz bewiesen. □

Satz 27.5. *Sei $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ unzerlegbar mit $A \geq 0$. Dann gilt:*

(i) Ist $Aw = sw$ mit $s \in \mathbb{R}, w \in \mathbb{R}^{n\times 1}$ und $0 \neq w \geq 0$, so ist $s = \rho(A)$.

(ii) $\rho(A)$ ist einfache Nullstelle des charakteristischen Polynoms χ_A von A.

Bemerkung 27.5. Die Jordan-Normalform von A hat also die Form

$$J = \begin{pmatrix} \rho(A) & 0 \\ 0 & B \end{pmatrix},$$

wobei $\rho(A)$ kein Eigenwert von B ist.

Beweis. (i) Sicher ist $A^T \in \mathbb{R}^{n\times n}$ unzerlegbar mit $A^T \geq 0$ und $\rho(A^T) = \rho(A)$. Nach Satz 27.4 existiert ein $z = (z_1, \ldots, z_n)^T \in \mathbb{R}^{n\times 1}$ mit $z > 0$ und $A^T z = \rho(A)z$. Schreibt man $w = (w_1, \ldots, w_n)^T$, so gilt für das Standardskalarprodukt auf $\mathbb{R}^{n\times 1}$:

$$\begin{aligned} s(z|w) = (z|sw) = (z|Aw) &= \sum_{j=1}^{n} z_j \left(\sum_{i=1}^{n} a_{ji} w_i \right) \\ &= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ji} z_j \right) w_i = (A^T z|w) = (\rho(A)z|w) = \rho(A)(z|w). \end{aligned}$$

Wegen $z > 0$ und $w \geq 0$ ist dabei $(z|w) > 0$, d.h. $s = \rho(A)$.

(ii) Die Abbildung $g : V := \mathbb{C}^{n\times 1} \longrightarrow V, x \longmapsto Ax$ ist linear. Nach Satz 27.4 existiert ein $y = (y_1, \ldots, y_n)^T \in \mathbb{R}^{n\times 1}$ mit $y > 0$ und $\operatorname{Ker}(g - \rho(A)\operatorname{id}_V) = \mathbb{C}y$. Wir müssen zeigen:

$$\operatorname{Ker}(g - \rho(A)\operatorname{id}_V)^2 = \operatorname{Ker}(g - \rho(A)\operatorname{id}_V).$$

Sei also $w = (w_1, \ldots, w_n)^T \in \operatorname{Ker}(g - \rho(A)\operatorname{id}_V)^2$. Dann ist

$$(g - \rho(A)\operatorname{id}_V)(w) \in \operatorname{Ker}(g - \rho(A)\operatorname{id}_V) = \mathbb{C}y,$$

d.h. $(g - \rho(A)\operatorname{id}_V)(w) = ay$ für ein $a \in \mathbb{C}$. Wie in (i) existiert ein $z = (z_1, \ldots, z_n)^T \in \mathbb{R}^{n\times 1}$ mit $z > 0$ und $A^T z = \rho(A)z$. Daher gilt für das kanonische hermitesche Skalarprodukt auf $\mathbb{C}^{n\times 1}$:

$$\begin{aligned} a(y|z) = (ay|z) &= ((g - \rho(A)\operatorname{id}_V)(w)|z) = ((A - \rho(A)1_n)w|z) = (Aw|z) - \rho(A)(w|z) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} w_j z_i - \rho(A)(w|z) = \sum_{j=1}^{n} w_j \left(\sum_{i=1}^{n} a_{ij} z_i \right) - \rho(A)(w|z) \\ &= (w|A^T z) - \rho(A)(w|z) = (w|A^T z - \rho(A)z) = (w|0) = 0. \end{aligned}$$

Wegen $y > 0$ und $z > 0$ ist $(y|z) > 0$, d.h. $a = 0$ und damit $w \in \operatorname{Ker}(g - \rho(a)\operatorname{id}_V)$.

□

Satz 27.6. *Seien $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ unzerlegbar und $C := (c_{ij}) \in \mathbb{C}^{n\times n}$ mit $(0 \leq)C^+ \leq A$. Dann ist $\rho(C) \leq \rho(A)$. Ist γ ein Eigenwert von C mit $|\gamma| = \rho(A)$, so ist $C = \frac{\gamma}{\rho(A)} DAD^{-1}$ mit einer Diagonalmatrix $D \in \mathbb{C}^{n\times n}$, deren Koeffizienten auf der Hauptdiagonalen alle den Betrag 1 haben.*

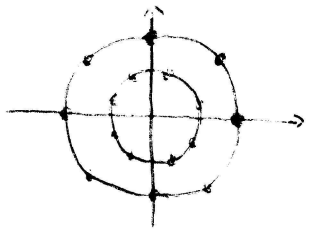
Beweis. Sei $\gamma \in \mathbb{C}$ Eigenwert von C, und sei $0 \neq y \in \mathbb{C}^{n\times 1}$ mit $Cy = \gamma y$. Dann ist $|\gamma|y^+ \leq C^+y^+ \leq Ay^+$. Mit den Bezeichnungen aus dem Beweis von Satz 27.4 ist also $|\gamma| \leq r_{y^+} \leq \rho(A)$. Dies zeigt: $\rho(C) \leq \rho(A)$.
Im Folgenden sei $|\gamma| = \rho(A)$. Dann ist y^+ ein extremaler Eigenvektor von A; wie im Beweis von Satz 27.4 gezeigt, ist also $y^+ > 0$. Ferner ist $Ay^+ = C^+y^+ = \rho(A)y^+$. Wegen $y^+ > 0$ und $C^+ \leq A$ folgt $C^+ = A$. Wir schreiben $y = (y_1, \ldots, y_n)^T$ mit $y_j = |y_j|e^{i\varphi_j}$ und $\varphi_j \in \mathbb{R}$ für $j = 1, \ldots, n$. Wir setzen

$$D := \begin{pmatrix} e^{i\varphi_1} & & 0 \\ & \ddots & \\ 0 & & e^{i\varphi_n} \end{pmatrix}.$$

Dann ist $y = Dy^+$. Wegen $Cy = \gamma y$ folgt $CDy^+ = \gamma Dy^+$, d.h. $D^{-1}CDy^+ = \gamma y^+$ und $\underbrace{\frac{\rho(A)}{\gamma}D^{-1}CD}_{=:F}\, y^+ = \rho(A)y^+ = Ay^+ = C^+y^+$. Nach Konstruktion ist $F^+ = C^+ = A$, d.h. $Fy^+ = F^+y^+$. Wegen $y^+ > 0$ folgt $F = F^+ = A$, d.h. $A = \frac{\rho(A)}{\gamma}D^{-1}CD$. □

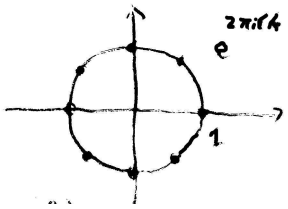
Satz 27.7. *Sei $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ unzerlegbar mit $A \geq 0$. Ferner habe A genau k Eigenwerte $r_1, \ldots, r_k \in \mathbb{C}$ vom Betrag $\rho(A)$. Bei geeigneter Nummerierung ist dann $r_j = e^{2\pi ij/k}\rho(A)$ für $j = 1, \ldots, k$. Ferner sind $r_1, \ldots, r_k$ einfache Nullstellen des charakteristischen Polynoms χ_A von A, und für jeden Eigenwert $r \in \mathbb{C}$ von A ist auch $e^{2\pi i/k}r$ ein Eigenwert von A.*

Bemerkung 27.7. Das bedeutet, dass die Menge der Eigenwerte von A invariant unter Drehungen um den Winkel $\frac{2\pi}{k}$ ist.



Beweis. Für $j = 1, \ldots, k$ sei $r_j := e^{i\varphi_j}\rho(A)$ mit $\varphi_j \in \mathbb{R}$. Wir wenden Satz 27.6 mit $C := A$ und $\gamma := r_j$ an. Es existiert also eine Diagonalmatrix $D_j \in \mathbb{C}^{n\times n}$, deren Koeffizienten auf der Hauptdiagonalen alle Betrag 1 haben, mit $A = e^{i\varphi_j}D_jAD_j^{-1}$. Die Eigenwerte von A stimmen also mit denen von $D_jAD_j^{-1} = e^{-i\varphi_j}A$ überein. Für $l = 1, \ldots, k$ ist also $e^{-i\varphi_j}e^{i\varphi_l}\rho(A) = e^{i(\varphi_l-\varphi_j)}\rho(A)$ Eigenwert von A, d.h. $e^{i(\varphi_l-\varphi_j)} \in \{e^{i\varphi_1}, \ldots, e^{i\varphi_k}\} =: U$. Dies zeigt, dass U eine Untergruppe der Ordnung k von $\mathbb{C} \setminus \{0\}$ ist.
Für $v \in U$ ist $vU = 1U = U$. Folglich ist $\prod_{u\in U} u = \prod_{u\in U} vu = v^k \prod_{u\in U} u$, d.h. $v^k = 1$. Die Elemente in U sind also Lösungen der Gleichung $X^k = 1$ in $\mathbb{C}$; bekanntlich hat diese

Gleichung die Lösungen $e^{2\pi ij/k}(j = 0, 1, \ldots, k-1)$ in $\mathbb{C}$. Bei geeigneter Nummerierung ist also $e^{i\varphi_j} = e^{2\pi ij/k}$, d.h. $r_j = e^{2\pi ij/k}\rho(A)$ für $j = 1, \ldots, k$. Die übrigen Aussagen folgen leicht.



□

Satz 27.8. *Sei $A \in \mathbb{R}^{n\times n}$ unzerlegbar mit $A \geq 0$. Ferner habe A genau k komplexe Eigenwerte vom Betrag $\rho(A)$. Dann existiert eine Permutationsmatrix $P \in \mathbb{R}^{n\times n}$ derart, dass PAP^{-1} die folgende Form hat:*

$$PAP^{-1} = \begin{pmatrix} 0 & A_{12} & 0 & \ldots & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & 0 & \ldots & \ldots & 0 & A_{k-1,k} \\ A_{k1} & 0 & \ldots & \ldots & 0 & 0 \end{pmatrix}$$

Beweis. Nach den vorigen Sätzen und Beweisen existieren eine Diagonalmatrix $D \in \mathbb{C}^{n\times n}$ mit Koeffizienten vom Betrag 1 auf der Hauptdiagonalen, so dass gilt:

$$A = e^{2\pi i/k} DAD^{-1}.$$

Für jede Permutationsmatrix $P \in \mathbb{R}^{n\times n}$ gilt also:

$$PAP^{-1} = e^{2\pi i/k}(PDP^{-1})(PAP^{-1})(PDP^{-1})^{-1};$$

dabei ist PDP^{-1} wieder eine Diagonalmatrix. Indem wir A notfalls durch PAP^{-1} ersetzen können wir annehmen, dass D die folgende Form hat:

$$D = \begin{pmatrix} e^{i\delta_1} 1_{n_1} & & 0 \\ & \ddots & \\ 0 & & e^{i\delta_m} 1_{n_m} \end{pmatrix} \qquad (\delta_1, \ldots, \delta_m \in [0, 2\pi[\text{ paarweise verschieden}).$$

Da wir D noch mit einem Skalar vom Betrag 1 multiplizieren können, dürfen wir $\delta_1 = 0$ annehmen. Wir zerlegen A entsprechend:

$$A = \begin{pmatrix} A_{11} & \ldots & A_{1m} \\ \vdots & & \vdots \\ A_{m1} & \ldots & A_{mm} \end{pmatrix}.$$

Aus $A = e^{2\pi i/k} DAD^{-1}$ ergibt sich:

$$A_{jl} = e^{i(2\pi/k+\delta_j-\delta_l)} A_{jl} \qquad (j,l = 1,\dots,m).$$

Im Fall $A_{jl} \neq 0$ ist also $e^{i(2\pi/k+\delta_j-\delta_l)} = 1$. Da $\delta_1,\dots,\delta_m \in [0,2\pi[$ paarweise verschieden sind, existiert zu jedem $j \in \{1,\dots,m\}$ höchstens ein $l \in \{1,\dots,m\}$ mit $A_{jl} \neq 0$ (und $e^{i(2\pi/k+\delta_j-\delta_l)} = 1$). Da A unzerlegbar ist, existiert zu jedem $j \in \{1,\dots,m\}$ genau ein $l \in \{1,\dots,m\}$ mit $A_{jl} \neq 0$ (und $e^{i\delta_l} = e^{i(2\pi/k+\delta_j)}$). Durch Permutation der Blöcke $e^{i\delta_1}1_{n_1},\dots,e^{i\delta_m}1_{n_m}$ in D können wir erreichen:

$$\begin{aligned}
&\delta_1 = 0 \qquad \text{(wie bereits vereinbart)}\\
&A_{12} \neq 0 \text{ und } \delta_2 = \frac{2\pi}{k}\\
&A_{23} \neq 0 \text{ und } \delta_3 = 2\frac{2\pi}{k}\\
&\dots\\
&A_{k-1,k} \neq 0 \text{ und } \delta_k = (k-1)\frac{2\pi}{k}.
\end{aligned}$$

Sei $l \in \{1,\dots,m\}$ mit $A_{kl} \neq 0$. Dann ist $e^{i\delta_l} = e^{i(2\pi/k+\delta_k)} = 1 = e^{i\delta_1}$, also $l = 1$. Es ist also $m \geq k$, und A hat die Form

$$A = \left(\begin{array}{cccccc|c}
0 & A_{12} & 0 & \dots & \dots & 0 & \\
\vdots & \ddots & \ddots & \ddots & & \vdots & \\
\vdots & & \ddots & \ddots & \ddots & \vdots & \\
\vdots & & & \ddots & \ddots & 0 & 0\\
0 & 0 & \dots & \dots & 0 & A_{k-1,k} & \\
A_{k1} & 0 & \dots & \dots & 0 & 0 & \\
\hline
 & & & \star & & & \star
\end{array}\right)$$

Da A unzerlegbar und die linke obere Teilmatrix quadratisch ist, folgt die Behauptung. □

Bemerkung 27.8. Im Fall $a_{jj} \neq 0$ für ein $j \in \{1,\dots,n\}$ ist also $k = 1$.

Satz 27.9. *Sei $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ unzerlegbar mit $A \geq 0$. Für $x = (x_1,\dots,x_n)^T \in \mathbb{R}^{n\times 1}$ mit $x > 0$ gilt dann:*

$$\min\left\{\sum_{k=1}^n a_{jk}x_k/x_j : j = 1,\dots,n\right\} \leq \rho(A) \leq \max\left\{\sum_{k=1}^n a_{jk}x_k/x_j : j = 1,\dots,n\right\}.$$

Insbesondere ist

$$\min\left\{\sum_{k=1}^n a_{jk} : j = 1,\dots,n\right\} \leq \rho(A) \leq \max\left\{\sum_{k=1}^n a_{jk} : j = 1,\dots,n\right\}.$$

Beweis. Wir wählen $z = (z_1, \ldots, z_n)^T \in \mathbb{R}^{n\times 1}$ mit $z > 0$ und $A^T z = \rho(A)z$. Für $j = 1, \ldots, n$ setzen wir

$$y_j := \sum_{k=1}^{n} a_{jk}x_k \text{ und } t_j := y_j/x_j.$$

Dann gilt:

$$\begin{aligned}\sum_{j=1}^{n}(t_j - \rho(A))x_j z_j &= \sum_{j=1}^{n} y_j z_j - \sum_{j,k=1}^{n} a_{kj}z_k x_j \\ &= \sum_{j,k=1}^{n} a_{jk}x_k z_j - \sum_{j,k=1}^{n} a_{kj}x_j z_k = 0.\end{aligned}$$

Wegen $x > 0$ und $z > 0$ existieren $j, k \in \{1, \ldots, n\}$ mit $t_j - \rho(A) \geq 0 \geq t_k - \rho(A)$. Daher ist

$$\min\{t_1, \ldots, t_n\} \leq t_k \leq \rho(A) \leq t_j \leq \max\{t_1, \ldots, t_n\}.$$

Die letzte Aussage folgt, indem wir $x := (1, \ldots, 1)^T$ setzen. □

Bemerkung 27.9. Im Folgenden übertragen wir einige Sätze auf zerlegbare Matrizen.

Satz 27.10. *Für $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ mit $A \geq 0$ gilt:*

(i) *$\rho(A)$ ist Eigenwert von A.*

(ii) *Ist $r \in \mathbb{C}$ Eigenwert von A mit $|r| = \rho(A)$, so existieren ein $\epsilon \in \mathbb{C}$ und ein $m \in \mathbb{N}$ mit $r = \epsilon\rho(A), \epsilon^m = 1$ und $m \leq n$.*

(iii) *Es existiert ein Eigenvektor $z \in \mathbb{R}^{n\times 1}$ von A zum Eigenwert $\rho(A)$ mit $z \geq 0$.*

(iv) *Für alle $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^{n\times 1}$ mit $x > 0$ gilt:*

$$\rho(A) \leq \max\left\{\sum_{k=1}^{n} a_{jk}x_k/x_j : j = 1, \ldots, n\right\};$$

insbesondere ist $\rho(A) \leq \max\left\{\sum_{k=1}^{n} a_{jk} : j = 1, \ldots, n\right\}$.

Beweis. Für unzerlegbare Matrizen haben wir die Aussagen bereits bewiesen. Sei also A zerlegbar. Wir können annehmen, dass A die folgende Form hat:

$$A = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix} \text{ mit } B = (b_{ij}) \in \mathbb{R}^{k\times k} \text{ und } 0 < k < n.$$

Induktiv können wir voraussetzen, dass die Aussagen für B und $D = (d_{ij})$ bereits bewiesen sind.

(i) Nach Induktion sind $\rho(B)$ und $\rho(D)$ Eigenwerte von B bzw. D. Daher ist $\rho(A) = \max\{\rho(B), \rho(D)\}$ Eigenwert von B oder D, also auch von A.

(ii) Sei $r \in \mathbb{C}$ Eigenwert von A mit $|r| = \rho(A)$. Dann ist r Eigenwert von B oder D. Wir können annehmen, dass r Eigenwert von B ist. Dann ist $\rho(A) = |r| \leq \rho(B) \leq \rho(A)$, d.h. $|r| = \rho(B) = \rho(A)$. Nach Induktion existieren ein $\epsilon \in \mathbb{C}$ und ein $m \in \mathbb{N}$ mit $r = \epsilon\rho(B), \epsilon^m = 1$ und $m \leq k$.

(iii) Wir suchen ein $z \in \mathbb{R}^{n\times 1}$ mit $0 \neq z \geq 0$ und $Az = \rho(A)z$. Der Ansatz $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$ liefert die Gleichung $\begin{pmatrix} Bz_1 \\ Cz_1 + Dz_2 \end{pmatrix} = \rho(A)\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$.

Fall 1: $\rho(A) = \rho(D)$.
Nach Induktion existiert ein $z_2 \in \mathbb{R}^{(n-k)\times 1}$ mit $0 \neq z_2 \geq 0$ und $Dz_2 = \rho(D)z_2 = \rho(A)z_1$. Wir können also $z_1 := 0$ setzen.

Fall 2: $\rho(A) = \rho(B) > \rho(D)$.
Nach Induktion existiert ein $z_1 \in \mathbb{R}^{k\times 1}$ mit $0 \neq z_1 \geq 0$ und $Bz_1 = \rho(B)z_1 = \rho(A)z_1$. Wir brauchen noch ein $z_2 \in \mathbb{R}^{(n-k)\times 1}$ mit $z_2 \geq 0$ und $Cz_1 + Dz_2 = \rho(A)z_2$, d.h.

$$\left(1_n - \frac{1}{\rho(A)}D\right)z_2 = \frac{1}{\rho(A)}Cz_1 \geq 0.$$

Nach Satz 27.1 ist $1_n - \frac{1}{\rho(A)}D$ invertierbar mit

$$\left(1_n - \frac{1}{\rho(A)}D\right)^{-1} = \sum_{j=0}^{\infty} \rho(A)^{-j}D^j \geq 0.$$

Wir setzen also

$$z_2 := \left(1_n - \frac{1}{\rho(A)}D\right)^{-1} \frac{1}{\rho(A)}Cz_1 \geq 0.$$

(iv) Wir schreiben $C = (c_{ij})$ und $x = \begin{pmatrix} y \\ z \end{pmatrix}$ mit $0 < y \in \mathbb{R}^{k\times 1}, 0 < z \in \mathbb{R}^{(n-k)\times 1}$. Nach Induktion gilt dann:

$$\begin{aligned}
\rho(B) &\leq \max\left\{\sum_{l=1}^{k} b_{jl}y_l/y_j : j = 1, \ldots, k\right\} \\
&= \max\left\{\sum_{l=1}^{n} a_{jl}x_l/x_j : j = 1, \ldots, k\right\} \\
&\leq \max\left\{\sum_{l=1}^{n} a_{jl}x_l/x_j : j = 1, \ldots, n\right\}
\end{aligned}$$

und

$$\begin{aligned}\rho(D) &\leq \max\left\{\sum_{l=1}^{n-k} d_{jl}z_l/z_j : j = 1,\dots,n-k\right\}\\ &\leq \max\left\{\sum_{l=1}^{n-k} d_{jl}z_l/z_j + \sum_{l=1}^{k} c_{jl}y_l/z_j : j = 1,\dots,n-k\right\}\\ &= \max\left\{\sum_{l=1}^{n} a_{jl}x_l/x_j : j = k+1,\dots,n\right\}\\ &\leq \max\left\{\sum_{l=1}^{n} a_{jl}x_l/x_j : j = 1,\dots,n\right\}.\end{aligned}$$

Wegen $\rho(A) = \max\{\rho(B), \rho(D)\}$ folgt die erste Behauptung. Die zweite Behauptung ergibt sich wieder, indem man $x := (1,\dots,1)^T$ setzt.

□

28 Einige Anwendungen

Das Leontieff-Modell (V. Leontieff, Nobelpreis Wirtschaftswissenschaften 1973)

Ein Konzern besitzt n Fabriken (z.B. Bergwerk, Kraftwerk, Automobilfabrik, d.h. $n = 3$).

Um Kohle im Wert von 1$ zu produzieren, wird benötigt:
Kohle im Wert von 0,1$
Strom im Wert von 0,3$
Autos im Wert von 0,1$

Um Strom im Wert von 1$ zu produzieren, wird benötigt:
Kohle im Wert von 0,25$
Strom im Wert von 0,4$
Autos im Wert von 0,15$

Um Autos im Wert von 1$ zu produzieren, wird benötigt:
Kohle im Wert von 0,2$
Strom im Wert von 0,5$
Autos im Wert von 0,1$

Außerdem hat der Markt einen externen Bedarf pro Woche von
Kohle im Wert von 50000$
Strom im Wert von 75000$
Autos im Wert von 125000$.

Wie viel Kohle, Strom, Autos müssen produziert werden, um sowohl den externen als auch den internen Bedarf zu befriedigen? (Ohne Überschüsse!)

Verbrauchsmatrix $A = \begin{pmatrix} 0,1 & 0,25 & 0,2 \\ 0,3 & 0,4 & 0,5 \\ 0,1 & 0,15 & 0,1 \end{pmatrix} \geq 0$

Bedarfsvektor $y = \begin{pmatrix} 50000 \\ 75000 \\ 125000 \end{pmatrix} \geq 0$

Gesucht: ***Produktionsvektor*** $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \geq 0$

Ansatz: $\underbrace{y}_{\text{externer Bedarf}} = x - \underbrace{Ax}_{\text{interner Verbrauch}} = (1_n - A)x$

Das System hat eine nichtnegative Lösung $x = (1_n - A)^{-1}y = \sum_{k=0}^{\infty} A^k y$, falls $\rho(A) < 1$ gilt. Dies ist sicher dann der Fall, falls jede Spaltensumme von A kleiner als 1 ist, d.h. falls jede Fabrik rentabel arbeitet (Satz 27.9).
Allgemeiner hat das System für jedes $y \geq 0$ eine nichtnegative Lösung, wenn ein $z > 0$ mit $Az < z$ existiert; dies bedeutet, dass eine Produktion existiert, bei der jede Fabrik beschäftigt ist und von jedem Produkt eine positive Menge für den Markt übrig bleibt.
In unserem Beispiel ist

$$(1_3 - A)^{-1} = \frac{1}{127}\begin{pmatrix} 186 & 102 & 98 \\ 128 & 316 & 204 \\ 42 & 64 & 186 \end{pmatrix}$$

und

$$x = (1_3 - A)^{-1}y = \frac{1}{127}\begin{pmatrix} 22.900.000 \\ 55.600.000 \\ 30.150.000 \end{pmatrix} \approx \begin{pmatrix} 229.921 \\ 437.795 \\ 237.401 \end{pmatrix}.$$

Die Fabriken müssen also produzieren: Kohle für 229.921 \$
Strom für 437.795 \$
Autos für 237.401 \$.

Das Leslie-Modell in der Populationsdynamik

Eine Population wird in n Altersklassen eingeteilt. Für $i = 1, \ldots, n-1$ sei $s_i > 0$ die Überlebensrate der Klasse i, d.h. von x_i Individuen der Klasse i erreichen $s_i x_i$ Individuen die Altersklasse $i + 1$. Die Überlebensrate der Klasse n sei 0.
Für $i = 1, \ldots, n$ sei f_i die Fruchtbarkeitsrate der Klasse i, d.h. x_i Individuen in Klasse i

haben durchschnittlich $f_i x_i$ Nachkommen.

$$\textbf{\textit{Leslie-Matrix}} \qquad A = \begin{pmatrix} f_1 & f_2 & \dots & f_{n-1} & f_n \\ s_1 & 0 & \dots & \dots & 0 \\ & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & \vdots \\ 0 & & & s_{n-1} & 0 \end{pmatrix} \geq 0.$$

Am Anfang seien jeweils x_i Individuen in der Altersklasse i vorhanden.

$$\textbf{\textit{Populationsvektor}} \qquad x = (x_1, \dots, x_n)^T \geq 0$$

Nach einer Zeiteinheit hat man den Populationsvektor $x' = Ax$, nach zwei Zeiteinheiten den Populationsvektor $x'' = Ax' = A^2x$, usw. Man interessiert sich für das Verhalten der Population nach langer Zeit, d.h. für

$$\lim_{k\to\infty} A^k x \text{ oder besser } \lim_{k\to\infty} A^k.$$

Man kann $f_n > 0$ annehmen; denn sonst ist

$$A = \begin{pmatrix} B & 0 \\ s & 0 \end{pmatrix} \text{ mit } B \in \mathbb{R}^{(n-1)\times(n-1)} \text{ und } A^k = \begin{pmatrix} B^k & 0 \\ sB^{k-1} & 0 \end{pmatrix} \text{ für } k \in \mathbb{N}.$$

Es genügt also, $\lim_{k\to\infty} B^k$ zu berechnen.
Daher sei im Folgenden $f_n > 0, s_1 > 0, \dots, s_{n-1} > 0$. Man kann sich überlegen, dass A dann unzerlegbar ist (Übungsaufgabe). Man kann ferner beweisen, dass A genau dann nur einen Eigenwert vom Betrag $\rho(A)$ hat, wenn die Indizes $i \in \{1, \dots, n\}$ mit $f_i \neq 0$ teilerfremd sind. Dies setzen wir im Folgenden voraus, definieren $r := \rho(A)$ und berechnen $\lim_{k\to\infty} r^{-k}A^k$.
Nach Satz 27.5 existiert ein $T \in \mathrm{GL}(n, \mathbb{C})$ mit

$$T^{-1}(r^{-1}A)T = \begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix} \text{ und } \rho(B) < 1.$$

Daher ist

$$T^{-1}(\lim_{k\to\infty} r^{-k}A^k)T = \lim_{k\to\infty} (T^{-1}r^{-1}AT)^k = \lim_{k\to\infty} \begin{pmatrix} 1 & 0 \\ 0 & B^k \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Wir schreiben

$$T = \begin{pmatrix} t_1 & & \\ \vdots & \star & \\ t_n & & \end{pmatrix}, T^{-1} = \begin{pmatrix} u_1 & \dots & u_n \\ & \star & \\ & & \end{pmatrix}.$$

Dann ist

$$P := \lim_{k\to\infty} r^{-k}A^k = T\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}T^{-1} = \dots = \begin{pmatrix} t_1u_1 & \dots & t_1u_n \\ \vdots & & \vdots \\ t_nu_1 & \dots & t_nu_n \end{pmatrix}.$$

Sei $z = (z_1, \ldots, z_n)^T \in \mathbb{R}^{n\times 1}$ mit $z > 0$ und $Az = rz$, d.h.

$$z = Pz = \begin{pmatrix} t_1(u|z) \\ \vdots \\ t_n(u|z) \end{pmatrix} \text{ mit } u := (u_1, \ldots, u_n)^T;$$

dabei ist $(\cdot|\cdot)$ das Standardskalarprodukt. Wir können T so normieren, dass $(u|z) = 1$ gilt. Dann ist $z_j = t_j$ für $j = 1, \ldots, n$.
Sei $y = (y_1, \ldots, y_n)^T \in \mathbb{R}^{n\times 1}$ mit $y > 0$ und $A^T y = ry$. Dann ist

$$y = P^T y = \begin{pmatrix} u_1(y|z) \\ \vdots \\ u_n(y|z) \end{pmatrix}.$$

Wir können y so normieren, dass $(y|z) = 1$ gilt. Dann ist $y_j = u_j$ für $j = 1, \ldots, n$. Wir sehen also:

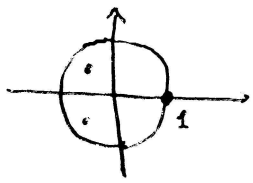
$$\lim_{k\to\infty} r^{-k} A^k = \begin{pmatrix} z_1 y_1 & \ldots & z_1 y_n \\ \vdots & & \vdots \\ z_n y_1 & \ldots & z_n y_n \end{pmatrix}$$

mit $y = (y_1, \ldots, y_n)^T, z = (z_1, \ldots, z_n)^T \in \mathbb{R}^{n\times 1}, y > 0, z > 0, (y|z) = 1, Az = rz$ und $A^T y = ry$. Daher gilt:

$$\lim_{k\to\infty} A^k = \begin{cases} 0 & \text{falls } \rho(A) < 1 \\ P & \text{falls } \rho(A) = 1 \\ \text{existiert nicht} & \text{falls } \rho(A) > 1. \end{cases}$$

Beispiel 28.1. $n = 3$.

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



Die Eigenwerte sind $1, -\frac{1}{2} \pm \frac{i}{2}$, d.h. $r = \rho(A) = 1$.
Eigenvektoren von A und A^T zum Eigenwert 1 sind $z = (6, 3, 1)^T$ bzw. $y = \frac{1}{15}(1, 2, 3)^T$. Beachte: $(y|z) = 1$. Wir erhalten:

$$\lim_{k\to\infty} A^k = P = \begin{pmatrix} \frac{2}{5} & \frac{4}{5} & \frac{6}{5} \\ \frac{1}{5} & \frac{2}{5} & \frac{3}{5} \\ \frac{1}{15} & \frac{2}{15} & \frac{1}{5} \end{pmatrix}.$$

Hat man am Anfang den Populationsvektor $x = (10, 10, 10)^T$, so hat man nach langer Zeit den Populationsvektor $Px = (24, 12, 4)^T$.

Markoff-Prozesse

Wir betrachten ein "System", das sich in einem von n Zuständen befinden kann. In einer Zeiteinheit gehe das System mit Wahrscheinlichkeit $a_{ij} \in [0,1]$ vom Zustand i in den Zustand j über. Die ***Übergangsmatrix*** $A := (a_{ij}) \in \mathbb{R}^{n\times n}$ ist dann nichtnegativ mit

$$\sum_{j=1}^{n} a_{ij} = 1 \qquad \text{für } i = 1, \dots, n.$$

Derartige Matrizen nennt man ***stochastisch***. Ggf. ist $(1, \dots, 1)^T$ Eigenvektor von A zum Eigenwert 1; insbesondere ist $\rho(A) \geq 1$. Nach Satz 27.10 ist andererseits

$$\rho(A) \leq \max\left\{\sum_{k=1}^{n} a_{jk} : j = 1, \dots, n\right\} = 1,$$

d.h. insgesamt ist $\boxed{\rho(A) = 1.}$

Für zwei stochastische Matrizen $A = (a_{ij}), B = (b_{ij}) \in \mathbb{R}^{n\times n}$ ist auch AB stochastisch; denn für $i = 1, \dots, n$ gilt:

$$\sum_{j=1}^{n}\left(\sum_{k=1}^{n} a_{ik}b_{kj}\right) = \sum_{k=1}^{n} a_{ik}\sum_{j=1}^{n} b_{kj} = \sum_{k=1}^{n} a_{ik} = 1.$$

Die Matrix A^2 beschreibt die Übergangswahrscheinlichkeit nach zwei Zeiteinheiten, usw. Man interessiert sich für die Entwicklung des Systems nach langer Zeit, d.h. für $\lim_{k\to\infty} A^k$.

Ist $A \in \mathbb{R}^{n\times n}$ stochastisch und existiert $P := \lim_{k\to\infty} A^k$, so gilt: $AP = P$ (und analog $PA = P$); denn

$$A \lim_{k\to\infty} A^k = \lim_{k\to\infty} A^{k+1} = \lim_{k\to\infty} A^k = P.$$

Wegen $AP = P$ liegen die Spalten von P im Eigenraum von A zum Eigenwert 1. Ferner gilt:

$$P^2 = (\lim_{k\to\infty} A^k)P = \lim_{k\to\infty}(A^kP) = \lim_{k\to\infty} P = P,$$

d.h. $P^2 = P$. Daher ist das Minimalpolynom μ_P ein Teiler von $X^2 - X = X(X-1)$. Daher hat P höchstens die Eigenwerte 0 und 1. Da μ_P in paarweise verschiedene Linearfaktoren zerfällt, ist P diagonalisierbar. Außerdem ist P wieder stochastisch.

Satz 28.2. *Sei $A \in \mathbb{R}^{n\times n}$ stochastisch. Ferner sei 1 der einzige Eigenwert von A in $\mathbb{C}$ vom Betrag 1. Dann existiert $P := \lim_{k\to\infty} A^k$.*

Beweis. Sei $S \in \mathrm{GL}(n, \mathbb{C})$ derart, dass $J := S^{-1}AS$ Jordan-Normalform hat. Dann ist $\rho(J) = \rho(A) = 1$, und 1 ist der einzige Eigenwert von J vom Betrag 1. Wir schreiben

$$J = \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_t \end{pmatrix} \text{ und } J_s = \begin{pmatrix} a_s & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & a_s \end{pmatrix} (s = 1, \dots, t).$$

Im Fall $|a_s| = \rho(J_s) < 1$ ist $\lim_{k\to\infty} J_s^k = 0$.
Für $B = (b_{ij}) \in \mathbb{R}^{n\times n}$ sei $\|B\| := \sum_{i,j=1}^n |b_{ij}|$. Für $C = (c_{ij}) \in \mathbb{R}^{n\times n}$ ist dann $\|BC\| \le \|B\| \cdot \|C\|$; denn

$$\begin{aligned}\|BC\| &\le \sum_{j,k,l=1}^n |b_{jl}| \cdot |c_{lk}| = \sum_{j,l=1}^n |b_{jl}| \sum_{k=1}^n |c_{lk}| \\ &\le \sum_{j,l=1}^n |b_{jl}| \sum_{k,l=1}^n |c_{lk}| = \|B\| \cdot \|C\|.\end{aligned}$$

Daher gilt für $k \in \mathbb{N}$:

$$\|J^k\| = \|S^{-1}A^kS\| \le \|S^{-1}\| \cdot \underbrace{\|A^k\|}_{=n} \cdot \|S\| = \underbrace{\|S\|^{-1}n\|S\|}_{\text{konstant}}.$$

Wir nehmen an, dass J einen Jordan-Block

$$J_s = \begin{pmatrix} 1 & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & 1 \end{pmatrix} \in \mathbb{R}^{m\times m} \qquad \text{mit } m > 1$$

enthält. Dann gilt für $k \in \mathbb{N}$:

$$(J_s)^k = \begin{pmatrix} 1 & & & 0 \\ k & \ddots & & \\ & \ddots & \ddots & \\ \star & & k & 1 \end{pmatrix},$$

d.h. $\|J^k\| > k$. Widerspruch.
Dies zeigt, dass die einzigen Jordan-Blöcke von J zum Eigenwert 1 das Format 1×1 haben müssen. Bei passender Nummerierung der Jordan-Blöcke ist also

$$J = \begin{pmatrix} 1_m & 0 \\ 0 & K \end{pmatrix} \text{ mit } \rho(K) < 1 \text{ und } \lim_{k\to\infty} J^k = \begin{pmatrix} 1_m & 0 \\ 0 & 0 \end{pmatrix},$$

d.h. $\lim_{k\to\infty} A^k = S \begin{pmatrix} 1_m & 0 \\ 0 & 0 \end{pmatrix} S^{-1}$. □

Index