Mathematik für Physiker
Lineare Algebra

Prof. Dr. Markus Klein
Anton Kirch, Dr. Elke Rosenberger
Universität Potsdam

1. Juli 2013

Inhaltsverzeichnis

Kapitel 1

Basis, Dimension und lineare Abhängigkeit

1.1 Der Begriff des Vektorraumes

Als ersten Schritt definieren wir das **geordnete n-Tupel** mit n$\in \mathbb{N}$

$$\mathbb{R}^n = \{x = (x_1, ..., x_n)\,;\, x_1, ..., x_n \in \mathbb{R}\}$$

und die **komponentenweise Addition**

$$x + y := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \qquad \text{kurz} \quad (x+y)_j := x_j + y_j \ .$$

$\mathbb{R} \ni x_j$ heißt j-te Komonente von $x \in \mathbb{R}^n$.
Bemerkung: $\mathbb{R}^n$ besteht üblicherweise aus Spalten. Wird $x \in \mathbb{R}^n$ als Zeile geschrieben, so fassen wir das immer als eine Schreibweise für eine Spalte auf.

Weiterhin vereinbaren wir die **komponentenweise Multiplikation mit Skalaren** $\lambda \in \mathbb{R}$**:**

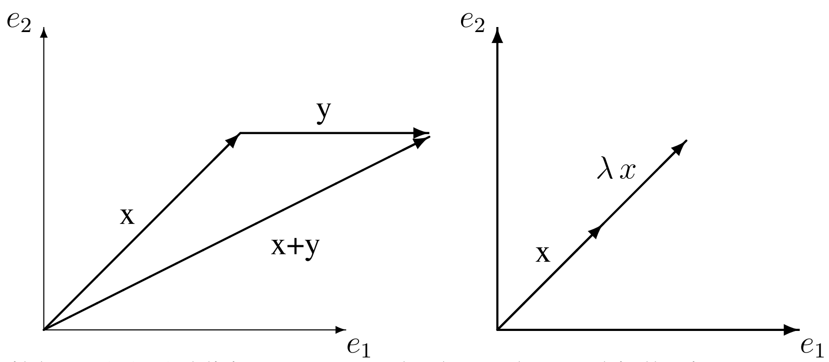


Abbildung 1.1: Addition von x und y bzw. der Multiplkation von x mit λ

$$\lambda x := \lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix} , \qquad \text{kurz} \quad (\lambda x)_j := \lambda x_j \; .$$

Das motiviert die folgende

DEFINITION 1.1 *Sei $\mathbb{K}$ ein Körper. Eine Menge V heißt ein $\mathbb{K}$-**Vektorraum** bezüglich der Addition $+ : V \times V \longmapsto V$ und der Multiplikation $\cdot : \mathbb{K} \times V \longmapsto V$:$\Leftrightarrow$*

V1) Das Tupel $(V, +)$ ist eine abelsche Gruppe. Insbesondere ist 0 das neutrale Element und $-x$ das Inverse zu x.

V2) $\forall v, w \in V,\ \mu, \lambda \in \mathbb{K}$ gilt:

(a) $(\lambda + \mu)v = \lambda v + \mu v$

(b) $\lambda(v + w) = \lambda v + \lambda w$

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

(d) $1v = v$

Das Quadrupel $(V, \mathbb{K}, +, \cdot)$ bezeichnet diesen $\mathbb{K}$-Vektorraum.

Beispiel: 1) $\mathbb{C}^n = \mathbb{C} \times ... \times \mathbb{C}$ ist ein $\mathbb{C}$-Vektorraum.
2) $\mathbb{Z}_p^n = \mathbb{Z}_p \times \mathbb{Z}_p \ldots \times \mathbb{Z}_p$ ist ein Vektorraum über dem Körper $\mathbb{Z}_p$, falls p prim ist.

Bemerkung: Allgemein ist $\mathbb{K}^n$ ist ein Vektorraum über den Körper $\mathbb{K}$ bezüglich der komponentenweisen Addition und Multiplikation mit Skalaren (d.h. Elementen aus K).

NOTIZ 1.2 *1)* $0 \cdot v = \lambda \cdot 0 = 0 \in V$
($0 \in K$ und $0 \in V$ wird mit dem selben Symbol bezeichnet!)

2) $(-1)v = -v$
denn: $(-1)v + v \overset{(d)}{=} (-1)v + 1 \cdot v \overset{(1)}{=} (-1 + 1)v = 0 \cdot v = 0$

3) $\lambda v = 0 \Rightarrow (v = 0$ *oder* $\lambda = 0)$
denn: Sei $\lambda \neq 0$, dann impliziert $\lambda v = 0 : v = \lambda^{-1}(\lambda v) = \lambda^{-1} 0 = 0$

NOTIZ 1.3 *Sei X eine Menge und $\mathbb{K}$ ein Körper, dann ist die Menge* $\mathrm{Abb}\,(X, \mathbb{K}) := \{f : X \to \mathbb{K}\}$ *ein $\mathbb{K}$-Vektorraum bzgl. der punktweisen Addition und Multiplikation mit Skalaren, die für $f, g \in$* $\mathrm{Abb}\,(X, \mathbb{K})$ *und $\lambda \in \mathbb{K}$ definiert ist durch*

$$\begin{aligned} (f + g)(x) &:= f(x) + g(x) \\ (\lambda f)(x) &:= \lambda \cdot f(x) \end{aligned}$$

*Ist speziell $X = \{1, ..., n\}$, dann führt das auf den **Standardvektorraum** $\mathbb{K}^n$.*

DEFINITION 1.4 *Sei V ein $\mathbb{K}$-Vektorraum. Eine Teilmenge $W \subset V$ heißt* ***Untervektorraum*** *von V $:\Leftrightarrow$ $W \neq \emptyset$ und die Verknüpfungen $+, \cdot$ lassen sich auf W beschränken, d.h. für alle $x, y \in W$ und $\lambda \in \mathbb{K}$ sind $x + y \in W$ und $\lambda \cdot x \in W$.*

Bemerkung: $(W, \mathbb{K}, +, \cdot)$ ist selbst ein $\mathbb{K}$-Vektorraum, falls W ein Untervektorraum ist.

Beispiel 1) Für $V = \mathbb{R}^3$ ist die Ebene $W = \mathbb{R}^2$ ein Untervektorraum.
2) Der Dualraum V^* eines Vektorraums V (siehe Definition 1.6, Notiz 1.7) ist ein Untervektorraum von $\mathrm{Abb}\,(V, \mathbb{K})$.

SATZ 1.5 *Sei V ein $\mathbb{K}$-Vektorraum und seien W_i, $i \in I$, Untervektorräume, dann ist*

$$W := \bigcap_{i \in I} W_i \subset V \tag{1.1}$$

ein Untervektorraum.

Beweis von Satz 1.5. Seien $v, w \in W$ und $\lambda \in \mathbb{K}$. Dann gilt

$$\begin{aligned} v, w \in W_i\,,\ (i \in I) \Rightarrow v + w \in W_i\,,\ (i \in I) \Rightarrow v + w \in W \\ w \in W_i\,,\ (i \in I) \Rightarrow \lambda \cdot w \in W_i\,,\ (i \in I) \Rightarrow \lambda \cdot w \in W \end{aligned}$$

Also lassen sich die Operationen +,$\cdot$ auf W beschränken. Wegen $0 \in W_i$ ist $0 \in W$, also $W \neq \emptyset$. □

DEFINITION 1.6 *Sei V eine $\mathbb{K}$-Vektorraum. Dann heißt $f \in \mathrm{Abb}(V, \mathbb{K})$* ***linear*** *oder* ***Linearform*** *$:\Leftrightarrow$ für alle $v_1, v_2, v \in V$ und $\lambda \in \mathbb{K}$ gilt*

$$f(v_1 + v_2) = f(v_1) + f(v_2) \tag{1.2}$$

$$f(\lambda v) = \lambda f(v) \tag{1.3}$$

Die Menge $V^* := \{f \in \mathrm{Abb}(V, \mathbb{K})\,;\, f\, linear\}$ heißt ***Dualraum*** *von V.*

NOTIZ 1.7 *Der Dualraum V^* eines Vektorraumes V ist selbst ein Vektorraum.*
Da $V^* \subset \mathrm{Abb}\,(V, \mathbb{K})$ genügt es hierfür, zu zeigen dass V^* ein Untervektorraum von $\mathrm{Abb}\,(V, \mathbb{K})$ ist: Für alle $f, g \in V^*$ und $u, v \in V$ gilt:

$$\begin{aligned} (f + g)(u + v) = f(u + v) + g(u + v) = f(u) + f(v) + g(u) + g(v) = (f + g)(u) + (f + g)v \\ (f + g)(\lambda v) = f(\lambda v) + g(\lambda v) = \lambda(f + g)(v) \end{aligned}$$

Also ist $(f + g) \in V^*$ (also linear). Analog folgt, dass $(\lambda f) \in V^*$ ist. Damit ist V^* ein Untervektorraum von $\mathrm{Abb}\,(V, \mathbb{K})$.

1.2 Basis und Dimension

DEFINITION 1.8 *Sei V ein $\mathbb{K}$-Vektorraum. Ein n-Tupel $(a_1, ..., a_n)$, $a_1, ..., a_n \in V$ heißt eine (endliche)* ***Basis*** *von V :⇔ Für alle $v \in V$ gibt es eindeutige $\lambda_1, \dots \lambda_n \in \mathbb{K}$, so dass gilt*

$$v = \sum_{j=1}^{n} \lambda_j a_j \ .$$

Wir sagen, v hat eine ***eindeutige Darstellung*** *als sogenannte* ***L****inearkombination der Basiselemente $a_1, ..., a_n$.*
Die $\lambda_1, ..., \lambda_n$ heißen ***Koordinaten*** *von v in der Basis $\mathcal{A} = (a_1, ..., a_n)$.*

Beispiel: Die Vektoren $e_i = (0, ..., 1, ...0)$, wobei sich der Eintrag 1 an der i-ten Stelle befindet, bilden die sogenannte **kanonische Basis** des $\mathbb{K}^n$. Die kanonische Basis wird allgemein auch als Standardbasis bezeichnet.

$$\mathbb{K}^n \ni x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 e_1 + ... + x_n e_n$$

DEFINITION 1.9 *Ein $\mathbb{K}$-Vektorraum V heißt* ***endlichdimensional*** :⇔ *V ist 0-dimensional (das heißt $V = \{0\}$) oder V besitzt eine endliche Basis.*

Ist V nicht endlichdimensional, dann ist V unendlichdimensional.

SATZ 1.10 *In einem endlichdimensionalen $\mathbb{K}$-Vektorraum V hat jede Basis $\mathcal{A} = (a_1, ..., a_n)$ dieselbe Länge n. Die Zahl der Basiselemente $n =: \dim V$ heißt* ***Dimension*** *von V.*

Bevor wir Satz 1.10 beweisen, führen wir zunächst noch einige Definitionen ein.

Die Definition einer Basis enthält zwei Teile. Für $v \in V$ gilt:

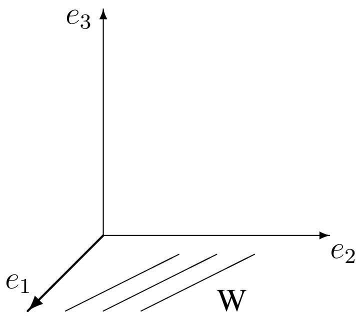
1. die Existenz einer Linearkombination $v = \sum x_j \, a_j$.

2. die Eindeutigkeit dieser Darstellung.

1. führt zu der

DEFINITION 1.11 *Sei V ein $\mathbb{K}$-Vektorraum und $M \subset V$. Die Menge aller Linearkombinationen von Elementen aus M heißt* ***lineare Hülle*** *von M, geschrieben als $\operatorname{Span} M$ (oder LH(M) oder $\langle\langle M \rangle\rangle$ oder ...). Wir sagen: $\operatorname{Span} M$ wir von M aufgespannt.*

Beispiel:



$M = \{e_1, e_2\} \subset \mathbb{R}^3$ spannen im $\mathbb{R}^3$ die Ebene

$$W = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} \; ; \; x_1, x_2 \in \mathbb{R} \right\} \text{ auf.}$$

Offensichtlich ist $\operatorname{Span} M$ ein Untervektorraum von V, genauer

NOTIZ 1.12 $\operatorname{Span} M$ *ist der kleinste Untervektorraum, der M enthält, d.h. es gilt*

$$\operatorname{Span} M = \bigcap_{i \in I} W_i \; ,$$

wobei $W_i\,, i \in I$ Untervektorräume von V mit $M \subset W_i$ sind. Die lineare Hülle von M liegt also im Schnitt aller Untervektorräume von V, die M enthalten.

DEFINITION 1.13 *Ein Tupel von Vektoren $(a_1, ..., a_r)$ in einem $\mathbb{K}$-Vektorraum V heißt* ***linear unabhängig*** *:⇔ Der Nullvektor $0 \in V$ ist eine eindeutige Linearkombination der a_i , $1 \leq i \leq r$, das heisst*

$$\forall \lambda_1, ...\lambda_r \in K, \; j = 1, ..., n \;\; : \quad \left(\sum_{j=1}^{r} \lambda_j a_j = 0 \; \Rightarrow \; \lambda_j = 0 \, , \; 1 \leq j \leq r \right) .$$

NOTIZ 1.14 *Das Tupel $(a_1, ..., a_r)$ in V ist linear unabhängig ⇒ jedes $v \in V$ hat entweder keine oder eine eindeutige Darstellung als Linearkombination der a_j, denn*

$$v = \sum_{i=1}^{r} x_i a_i = \sum \tilde{x}_i a_i \Rightarrow \sum (x_i - \tilde{x}_i) a_i = 0 \overset{a_i \text{ lin. unabh.}}{\Rightarrow} x_i - \tilde{x}_i = 0 \; .$$

NOTIZ 1.15 *Ein Tupel $\mathcal{A} = (a_1, ..., a_n)$ in einem Vektorraum V ist Basis von V ⇔*

1. $\operatorname{Span} \{a_1, ..., a_n\} = V$.
2. *Das Tupel $(a_1, ..., a_n)$ ist linear unabhängig.*

Beweis zu Satz 1.10. Sei $\mathcal{A} = (a_1, ..., a_n)$ Basis von V. Dann ist zu zeigen:
Kein Tupel $(b_1, ..., b_m)$ mit $m \neq n$ bildet eine Basis von V.
Wir behandeln (mit Widerspruchsbeweis) den Fall $m > n$, der Fall $m < n$ folgt dann durch Vertauschung der Rolle von $\mathcal{A}$ und $\mathcal{B}$.
Widerspruchsannahme: Sei $\mathcal{B} = (b_1, ..., b_m)$ eine weitere Basis von V mit $m > n$.

Da $\mathcal{A}$ eine Basis ist, existieren eindeutig bestimmte $c_{ij} \in \mathbb{K}$, $j = 1, ..., m\,,\ i = 1, \ldots n$, so dass $b_j = \sum_{i=1}^n c_{ij}\, a_i$,. Für $x_j \in K\ ,\ j = 1, \ldots m$ gilt damit:

$$\sum_{j=1}^m x_j\, b_j = \sum_{j=1}^m \sum_{i=1}^n x_j\, c_{ij}\, a_i = \sum_{i=1}^n \underbrace{\left(\sum_{j=1}^m c_{ij}\, x_j \right)}_{=:y_i} a_i = \sum_{i=1}^n y_i a_i$$

Wir setzen nun diesen Ausdruck = 0 und betrachten die Konsequenz:

$$\begin{array}{lcl}
\sum_{j=1}^m x_j\, b_j = 0 & \Longleftrightarrow & \sum_{i=1}^n y_i\, a_i = 0 \\
\Updownarrow\ b_j \text{ lin. unabh.} & & \Updownarrow\quad \mathcal{A} \text{ Basis} \\
x_j = 0, \quad j = 1, ..., m & & \sum_{j=1}^m c_{ij}\, x_j = y_i = 0 \quad \forall i \in \{1, ..., n\} \\
 & & \Updownarrow\quad \text{nach Definition} \\
 & & \sum_{j=1}^m c_{ij}\, x_j = 0
\end{array}$$

Das homogene Gleichungssystem $\sum_{j=1}^m c_{ij} x_j = 0$ mit $i = 1, \ldots, n$ hat also nur die triviale Lösung $x_1 = x_2 = ... = x_m = 0$. Das ist aber ein Widerspruch zu dem folgenden

LEMMA 1.16 (**Fundamental-Lemma**) *Ein homogenes lineares Gleichungssystem*

$$\sum_{j=1}^m c_{ij} x_j = 0 \qquad \textit{mit} \quad i = 1, \ldots, n \tag{1.4}$$

besitzt in jedem Körper $\mathbb{K}$ *eine nicht-triviale Lösung* $x = (x_1, \ldots x_m) \neq 0$*, falls die Anzahl der Unbekannten* m *größer ist als die Anzahl der Gleichungen* n*.*

Damit ist Satz 1.10 gezeigt. □

Bemerkung: Das Fundamental-Lemma (und sein Beweis) ist genauso wichtig wie der Satz über die Wohldefiniertheit der Dimension als Länge einer Basis.

Beweis des Fundamental-Lemmas. Induktion nach $n \in \mathbb{N}$:
IV (Induktionsverankerung, also A(n=1)): Beh.: $c_1 x_1 + \ldots + c_m x_m = 0$ hat eine Lösung $x \neq 0$ für alle $m > 1$.

Beweis von A(n=1) wieder durch Induktion nach $m \geq 2$:
IV: Beh. $c_1 x_1 + c_2 x_2 = 0$ hat eine Lösung $x \neq 0$.
Bew: Für $c_1 = c_2 = 0$ okay. Sei o.B.d.A. (“ohne Beschränkung der Allgemeinheit”) $c_1 \leq 0$, dann folgt $x_1 = -\frac{c_2}{c_1} x_2$, was immer eine Lösung $x = (x_1, x_2) \neq 0$ liefert.
IS (Induktionsschritt, also $\forall m \geq 2 : A(m) \Rightarrow A(m+1)$):

Beh: $c_1x_1 + \dots c_mx_m + c_{m+1}x_{m+1} = 0$ hat Lsg. $x \neq 0$.
Bew.: Da A(m) wahr, hat $c_1x_1 + \dots c_mx_m = 0$ eine Lösung $x' = (x_1, \dots x_m) \neq 0$, also löst $x = (x', 0) \neq 0$ die obige Gleichung.

IS (also $\forall n \in \mathbb{N}^* : A(n-1) \Rightarrow A(n)$): Sei in (1.4) mindestens ein $c_{ij} \neq 0$ (sonst wäre jedes $x \in \mathbb{R}^m$ eine Lösung und wir wären fertig). Nach umnummerieren: $x_j \mapsto x_1\,,\ x_1 \mapsto x_j$ und Vertauschen der Gleichungen können wir o.B.d.A annehmen $c_{11} \neq 0$. Durch äquivalente Umformungen erhalten wir aus (1.4)

$$\begin{array}{lllll} c_{11}x_1 & + \dots & & +c_{1m}x_m & = 0 \\ 0 \cdot x_1 & +\left(c_{22} - \frac{c_{21}}{c_11}\right)x_2 & + \dots & & = 0 \\ 0 \cdot x_1 & +\left(c_{32} - \frac{c_{31}}{c_11}\right)x_2 & + \dots & & = 0 \\ & \dots & & & = 0 \\ 0 \cdot x_1 & +\left(c_{n2} - \frac{c_{n1}}{c_11}\right)x_2 & + \dots & +\left(c_{nm} - \frac{c_{n1}}{c_11}\right)x_m & = 0 \end{array} \tag{1.5}$$

Dann löst $x = (x_1 \dots . x_m)$ das Gleichungssystem (1.4) genau dann, wenn x das Gleichungssystem (1.5) löst. Aber nach Induktionsvoraussetzung gibt es $\tilde{x} = (x_2, \dots x_m) \neq 0$, das das homogene Gleichungssystem

$$\begin{array}{llll} \left(c_{22} - \frac{c_{21}}{c_11}\right)x_2 & + \dots & & = 0 \\ \left(c_{32} - \frac{c_{31}}{c_11}\right)x_2 & + & \dots & = 0 \\ & & \dots & = 0 \\ \left(c_{n2} - \frac{c_{n1}}{c_11}\right)x_2 & + \dots & +\left(c_{nm} - \frac{c_{n1}}{c_11}\right)x_m & = 0 \end{array} \tag{1.6}$$

löst. Damit löst $x = (x_1 = \frac{1}{c_{11}}(c_{12}x_2 + \dots c_{1m}x_m), x_2, \dots x_m) \neq 0$ das Gleichungssystem (1.5) und somit auch (1.4). $\square$

NOTIZ 1.17 **(Gaußsches Eliminationsverfahren)** *Eine Fortsetzung des Eliminationsverfahrens liefert (gegebenenfalls nach Umnummerieren und Vertauschen der Gleichungen !)*

$$\begin{array}{cccccc} c_{11}x_1 & + \dots & & & +c_{1m}x_m & = 0 \\ 0 & +\tilde{c}_{22}x_2 & + \dots & & & = 0 \\ 0 & +0 & +\tilde{c}_{33}x_3 \dots & & & = 0 \\ & \dots & \dots & & & = 0 \\ 0 & +0 & + \dots & +0 & +\tilde{c}_{nm}x_m & = 0 \end{array} \tag{1.7}$$

Dabei können auf der Diagonalen durchaus Nullen stehen! Jede Lösung $(x_1, \dots x_m)$ *der Zeile* n *liefert durch Einsetzen eine (oder* ∞ *viele) Lösungen von* (1.7)*. Alle Lösungen können aus* (1.7) *abgelesen werden, durch sukzessives Einsetzen "von unten nach oben".*

Bsp.:

$$\begin{array}{llll} x_1 & +2x_2 & +3x_3 & = 0 \\ x_1 & +x_2 & +x_3 & = 0 \end{array} \sim \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \end{pmatrix}$$

Die untere Zeile des letzten Systems ist ausgeschrieben $-x_2 - 2x_3 = 0$ *und ist äquivalent zu* $x_3 = t$ *und* $x_2 = -2t$ *für beliebiges* $t \in \mathbb{K}$*. Eingesetzt in die obere Zeile liefert das* $x_1 - 4t + 3t = 0$ *und damit* $x_1 = t$*. Damit ist die Lösung*

$$x = \begin{pmatrix} t \\ -2t \\ t \end{pmatrix} = t \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \quad , \quad t \in \mathbb{K}\,.$$

LEMMA 1.18 *Sei V ein $\mathbb{K}$-Vektorraum, dann gilt:*

1) $(a_1, ..., a_n)$ *sind linear abhängig in* $V \Leftrightarrow$
   Mindestens ein a_j *ist eine Linearkombination der übrigen.*

2) $(a_1, ..., a_n)$ *ist linear unabhängig und* $(a_1, ..., a_n, a_{n+1})$ *linear abhängig* $\Rightarrow$
   a_{n+1} *ist Linearkombination der anderen Vektoren* $(a_1, ..., a_n)$

Beweis zu Lemma 1.18. ad 1)
"$\Rightarrow$"

$$\sum_{j=0}^{n} \lambda_j a_j = 0 \text{ mit einem } \lambda_j \neq 0 \;\Rightarrow\; a_j = -\frac{1}{\lambda_j} \sum_{i \neq j} \lambda_i a_i$$

"$\Leftarrow$"

$$a_j = \sum_{i \neq j} \mu_i a_i \;\Rightarrow\; a_j - \sum_{i \neq j} \mu_i a_i = 0 \quad \text{nicht-triviale Linearkomb.}$$

ad 2) Nach Voraussetzung gibt es $\lambda_1, ..., \lambda_{n+1} \in \mathbb{K}$ mit $\sum_{j=1}^{n+1} \lambda_j a_j = 0$. Hier sind nicht alle $\lambda_j = 0 \Rightarrow \lambda_{n+1} \neq 0$ (sonst wäre $\sum_{i=1}^{n} \lambda_i a_i = 0$ nichttriviale Linearkombination, und dann wären $(a_1, ..., a_n)$ linear abhängig)

$$\Rightarrow a_{n+1} = -\frac{1}{\lambda_{n+1}} \sum_{i=1}^{n} \lambda_i a_n \,.$$

$\square$

DEFINITION 1.19 *Sei V ein $\mathbb{K}$-Vektorraum und $A = \{a_1, ..., a_n\} \subset V$ eine endliche Menge. Eine linear unabhängige Teilmenge $F \subset A$ heißt* ***maximal*** *:$\Leftrightarrow$ Jedes $x \in A$ ist eine Linearkombination von Elementen aus F.*

SATZ 1.20 *Jede linear unabhängige Teilmenge $A' \subset A$ liegt in einer maximalen linear unabhängigen Teilmenge $F \subset A$ und es gilt:* $\operatorname{Span} F = \operatorname{Span} A$.

Beweis zu Satz 1.20. Enthält $A \setminus A'$ einen Vektor x, der keine Linearkombination von A' ist?
Falls nein: A' ist maximal und damit $F = A'$ für eine maximale Menge F.
Falls ja: bilde $A' \cup \{x\} =: A''$. Dann ist A'' linear unabhängig nach Lemma 1.18 (sonst wäre x eine Linearkombination aus A', im Widerspruch zur Voraussetzung). Nun weiter wie vorne für A'' iterativ $\overset{A\,endlich}{\Rightarrow}$ Stop falls die Antwort "nein" lautet, also auf maximaler Menge F. $\square$

Bemerkung F ist nicht eindeutig, aber $\#F = \dim \operatorname{Span} A =: \operatorname{Rang} A$.

SATZ 1.21 (BASISERGÄNZUNGSSATZ) *Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und sei $A' = \{a_1, ..., a_m\}$ linear unabhängig in V, dann gibt es eine Basis von V, die A' enthält.*

Beweis. Wähle eine Basis $(a_{m+1}, ..., a_{m+n})$ von V mit $n = \dim V$.
Sei $A := \{a_1, ..., a_m; a_{m+1}, ..., a_{m+n}\}$, dann folgt aus Satz1.20, dass es eine maximale linear unabhängige Menge $F \supset A'$ von A gibt und F ist die gesuchte Basis, denn $\operatorname{Span} F = \operatorname{Span} A = V$. $\square$

1.3 Übungen

Sei W ein Untervektorraum von V. Zeige, dass gilt:

1. $dim\ W \leq dim\ V$
2. Falls V endlich dimensional ist: $dim\ W = dim\ V \Rightarrow W = V$

Kapitel 2

Euklidische und unitäre Vektorräume

2.1 Norm und Skalarprodukt

Die Längen- und Winkelmessung in einem Vektorraum V erfordert eine Zusatzstruktur, wir betrachten hier nur die Fälle des reellen und des komplexen Vektorraums, also $\mathbb{K} = \mathbb{R}$ bzw. $\mathbb{K} = \mathbb{C}$.

DEFINITION 2.1 *Sei V ein $\mathbb{K}$-Vektorraum. Dann heißt die Abbildung $\|\cdot\| : V \longmapsto \mathbb{R}$* ***Norm****:$\Leftrightarrow$*

1) $\|x\| \geq 0$ und es gilt $\|x\| = 0 \Rightarrow x = 0$ für alle $x \in V$.

2) $\|\lambda x\| = |\lambda|\|x\|$ für alle $\lambda \in K$ und $x \in V$.

3) $\|x + y\| \leq \|x\| + \|y\|$ für alle $x, y \in V$ (Dreiecksungleichung).

$(V, \|\cdot\|)$ heißt ***normierter Vektorraum****.*

Beispiel 1) Sei $V = \mathbb{R}^n$ oder $V = \mathbb{C}^n$ mit $x = (x_1, \ldots x_n) \in V$, dann sind Normen gegeben durch

$$\|x\|_1 := \sum_{j=1}^{n} |x_j|$$
$$\|x\|_\infty := \max_{1\leq j\leq n} |x_j|$$
$$\|x\|_2 := \sqrt{\sum_{j=1}^{n} |x_j|^2}$$

Bei $\|\cdot\|_1$ und $\|\cdot\|_\infty$ sind die Eigenschaften offensichtlich, bei $\|\cdot\|_2$ bedarf die Dreiecksungleichung eines Beweises.
2) Sei $V = \ell^1 := \{(a_k)_{k\in\mathbb{N}} \,|\, a_k \in \mathbb{R}$ und $\sum_{k\in\mathbb{N}} |a_k| < \infty\}$ der Raum der absolut-summierbaren rellen Folgen. Dann ist für $a = (a_k)$ durch $\|a\|_1 := \sum_{k\in\mathbb{N}} |a_k|$ eine Norm gegeben.

DEFINITION 2.2 *Sei V ein $\mathbb{R}$-Vektorraum. Dann heißt die Abbildung $\langle\cdot,\cdot\rangle : V \times V \longmapsto \mathbb{R}$* ***Skalarprodukt***:$\Leftrightarrow$ *$\langle\cdot,\cdot\rangle$ ist*

1) *bilinear, d.h. für alle $x \in V$ sind die Abbildungen $\langle x,\cdot\rangle$ und $\langle\cdot,x\rangle : V \to \mathbb{R}$ linear.*

2) *symmetrisch, d.h. $\langle x,y\rangle = \langle y,x\rangle$ für alle $x,y \in V$.*

3) *positiv definit, d.h. $\langle x,x\rangle > 0$ falls $x \neq 0$.*

$(V,\langle\cdot,\cdot\rangle)$ heißt ***euklidischer Raum***.

Beispiel

1) $\mathbb{R}^n$ mit Standardskalarprodukt $\langle x,y\rangle := \sum x_i\, y_i$, (alternative Schreibweisen: $\langle x,y\rangle = x\cdot y = (x,y)$)

2) $V = \mathcal{C}([a,b],\mathbb{R})$ mit Skalarprodukt $\langle f,g\rangle := \int_a^b f(x)\, g(x)\, dx$.

SATZ 2.3 *Ein eukidischer Vektorraum $(V,\langle\cdot,\cdot\rangle)$ ist ein normierter Vektorraum mit der durch das Skalarprodukt induzierten Norm*

$$\|x\| := \sqrt{\langle x,x\rangle}$$

SATZ 2.4 (CAUCHY-SCHWARZ-UNGLEICHUNG) *In jedem euklidischen Vektorraum $(V,\langle\cdot,\cdot\rangle)$ mit induzierter Norm $\|\cdot\|$ gilt die Ungleichung:*

$$|\langle x,y\rangle| \leq \|x\|\ \|y\|\ , \qquad x,y \in V$$

Beweis von Satz 2.4. Sei $y \in V$ und $e := \frac{y}{\|y\|}$. Dann gilt $\|e\|^2 = \langle e,e\rangle = \frac{\langle e,e\rangle}{\|y\|^2} = 1$ und damit

$$\begin{aligned}
0 \leq \|x - \langle e,x\rangle\, e\|^2 &= \langle x - \langle e,x\rangle\, e, x - \langle e,x\rangle\, e\rangle \\
\overset{bilinear}{=} \langle x,x\rangle - \langle e,x\rangle\langle e,x\rangle - \langle e,x\rangle\langle x,e\rangle + \langle e,x\rangle^2 &\overset{symmetrisch}{=} \|x\|^2 - \langle e,x\rangle^2 \\
\Rightarrow \left|\left\langle \frac{y}{\|y\|}, x\right\rangle\right| \leq \|x\| &\Rightarrow |\langle y,x\rangle| \leq \|x\|\|y\|
\end{aligned}$$

□

Beweis von Satz 2.3. Die Eigenschaften 1) und 2) sind klar, wir zeigen die Dreiecksungleichung 3):

$$\begin{aligned}
(\|x\| + \|y\|)^2 &= \|x\|^2 + \|y\|^2 + 2\underbrace{\|x\|\|y\|}_{\geq|\langle x,y\rangle|} \\
&\geq \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\
&\quad = \langle x+y, x+y\rangle = \|x+y\|^2
\end{aligned}$$

□

Bemerkung: Wegen Satz 2.3 ist jeder euklidische Vektorraum $(V, \langle \cdot, \cdot \rangle)$ ein metrischer Raum.

DEFINITION 2.5 *Ist $(V, \langle\ ,\ \rangle)$ ein euklidischer Vektorraum, dann definieren wir zu $x, y \in V$ den* ***Öffnungswinkel*** *$\alpha(x, y)$ durch*

$$\cos\,\alpha(x, y) := \frac{\langle x, y\rangle}{\|x\|\|y\|} \quad (\in [-1, 1] \text{ nach Satz 2.4})$$

Insbesondere ist $\alpha(x, y) \in [0, \pi]$ wohldefiniert.

DEFINITION 2.6 *Seien V, W $\mathbb{C}$-Vektorräume, dann heißt die Abbildung*
$F : V \to W$ ***semilinear*** *oder* ***antilinear****: $\Leftrightarrow \forall x, y \in V$*

SL 1) $F(x + y) = F(x) + F(y)$

SL 2) $F(\lambda x) = \overline{\lambda} F(x)$

DEFINITION 2.7 *Seien V, W $\mathbb{C}$-Vektorräume, dann heißt eine Abbildung $f : V \times W \to \mathbb{C}$* ***Sesquilinearform*** *:$\Leftrightarrow$*

1) *$f(v, \cdot)$ ist linear für alle $v \in V$.*

2) *$f(\cdot, w)$ ist antilinear für alle $w \in W$.*

Eine Sesquilinearform $f : V \times V \to \mathbb{C}$ heißt ***hermitesch*** *oder* ***symmetrisch*** *:$\Leftrightarrow$*

$$f(x, y) = \overline{f(y, x)}\,, \qquad x, y \in V$$

Bemerkung: f hermitisch $\Rightarrow \quad f(x, x) = \overline{f(x, x)} \quad \in \mathbb{R}$

DEFINITION 2.8 *Eine hermitesche Form $f : V \times V \to \mathbb{C}$ heißt* ***positiv definit*** *:$\Leftrightarrow$*

$$f(x, x) > 0 \quad \forall x \in V, x \neq 0\,.$$

Eine positiv definite hermitische Sesquilinearform $f(\cdot, \cdot) =: \langle \cdot, \cdot \rangle$ auf einem $\mathbb{C}$-Vektorraum V heißt ***komplexes Skalarprodukt*** *auf V und $(V, \langle \cdot, \cdot \rangle)$ heißt* ***unitärer Vektorraum****.*

NOTIZ 2.9 *Im einzelnen gilt für ein komplexes Skalarprodukt $\langle \cdot, \cdot \rangle$ in V:*

1) $\langle \lambda(x + y), z\rangle = \overline{\lambda}\,\langle x, z\rangle + \overline{\lambda}\,\langle y, z\rangle$ *für alle $\lambda \in \mathbb{C}$ und $x, y, z \in V$.*

2) $\langle x, y\rangle = \overline{\langle y, x\rangle}$

3) $\langle x, x\rangle > 0$ *für $x \neq 0$*

Beispiel: 1) Sei $V = \mathbb{C}^n$, dann ist hier ist das Standardskalarprodukt als $\langle x, y\rangle = \sum_{i=1}^n \overline{x_i}\, y_i$ definiert.
2) $V = C([a,b], \mathbb{C})$ mit Skalarprodukt $\langle f, g\rangle := \int_a^b \overline{f}(x)\, g(x)\, dx$.

SATZ 2.10 (CAUCHY-SCHWARZ-UNGLEICHUNG) *Sei $(V, \langle\cdot,\cdot\rangle)$ ein unitärer Vektorraum über $\mathbb{C}$, dann ist durch $\|x\| := \sqrt{\langle x, x\rangle}$ eine Norm auf V definiert (die durch das Skalarprodukt induzierte Norm) und es gilt*

$$|\langle x, y\rangle| \leq \|x\|\|y\|\ , \qquad x, y \in V$$

Beweis von Satz 2.10. Übung □

Bemerkung: Jede durch ein Skalarprodukt induzierte Norm $\|x\| = \sqrt{\langle x, x\rangle}$ erfüllt die Parallelogrammregel

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

In $C([a,b], \mathbb{C})$ mit $\|f\|_\infty := \sup|f(x)|$, $x \in [a,b]$ gilt die Parallelogrammregel im allgemeinen nicht:
Es kann gelten $1 = \|f\|_\infty = \|g\|_\infty = \|f+g\|_\infty = \|f-g\|_\infty$, dann ist aber $1 + 1 \neq 2(1+1)$.

2.2 Orthonormierung

DEFINITION 2.11 *Sei $(V, \langle\cdot,\cdot\rangle)$ ein $\mathbb{K}$-Vektorraum mit Skalarprodukt, wobei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$.*

1) *$x, y \in V$ heißen **orthogonal** oder **senkrecht** zueinander (wir schreiben kurz $x \perp y$) $:\Leftrightarrow$ $\langle x, y\rangle = 0$.*

2) *Für $M \subset V$ heißt $M^\perp := \{x \in V; \langle x, y\rangle = 0 \quad \forall y \in M\}$ das **orthogonale Komplement** von M.*

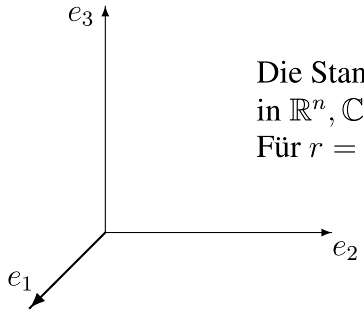
NOTIZ 2.12 *Es gilt $M^\perp = \bigcap_{y\in M} f_y^{-1}(0)$, mit $f_y(\cdot) = \langle y, \cdot\rangle : V \to \mathbb{C}$ stetig.*
$M^\perp$ ist ein abgeschlossener Untervektorraum von V, denn
1) $0 \in M^\perp$, da $0 \perp M$
2) für $x, y \in M^\perp$ ist $x + y \in M^\perp$, denn $\forall u \in M,\ x, y \in V$ gilt
$x, y \in M^\perp \Rightarrow \langle x+y, u\rangle = \langle x, u\rangle + \langle y, u\rangle = 0$
3) für $x \in M^\perp$ und $\lambda \in \mathbb{K}$ ist $\lambda x \in M^\perp$, denn $\langle \lambda x, u\rangle = \overline{\lambda}\,\langle x, u\rangle = 0$ für alle $u \in M$.
4) Die Urbilder abgeschlossener Mengen (hier die Menge $\{0\}$) unter stetigen Abbildungen (hier f_y) sind wiederum abgeschlossen und der Schnitt abzählbarer Mengen ist abzählbar.

DEFINITION 2.13 *Ein r-Tupel $(v_1, ..., v_r)$ von Vektoren in einem Vektorraum $(V, \langle\cdot,\cdot\rangle)$ heißt **orthonormal** bzw. **Orthonormalsystem** (abgekürzt ONS) $:\Leftrightarrow$*

$$\langle v_i, v_j\rangle = \delta_{ij} = \begin{cases} 1 \text{ für } i = j \\ 0 \text{ für } i \neq j \end{cases} \quad , \qquad \text{für alle } i, j = 1, ..., r \tag{2.1}$$

Das heißt $\|v_i\| = 1$ *und* $v_i \perp v_j$ *für alle* $i \neq j, 1 \leq i, j \leq r$.
Ist ein Orthonormalsystem eine Basis von V*, dann heisst es* ***Orthonormalbasis*** *von* $(V, \langle \, , \, \rangle)$.

Beispiel:



Die Standardvektoren $(e_1, ..., e_r)$
in $\mathbb{R}^n, \mathbb{C}^n$ bilden ein Orthonormalsystem.
Für $r = n$ bilden sie sogar eine Orthonormalbasis.

LEMMA 2.14 *Ein Orthonormalsystem* $(v_1, ..., v_r)$ *ist immer linear unabhängig.*

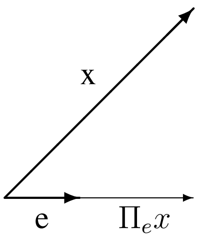
Beweis von Lemma 2.14. Es gelte $\lambda_1 v_1 + ... + \lambda_r v_r = 0$. Auf diese Gleichung wenden wir die Funktion $\langle v_j, \, . \rangle : V \to \mathbb{K}$ an und erhalten $0 + \lambda_j + 0 = 0$ und damit $\lambda_j = 0$.

□

SATZ 2.15 *Ist* $(v_1, ..., v_r)$ *eine Orthonormalbasis von V, so gilt die folgende Entwicklungsformel:*

$$x = \sum_{j=1}^{n} v_j \langle v_j, x \rangle \tag{2.2}$$

Heuristik



$\Pi_e x$ ist die orthogonale Projektion
des Vektors x auf Vektor e.

$$\langle e, e \rangle = 1 \tag{2.3}$$
$$\Pi_e x = \langle e, x \rangle e \tag{2.4}$$
$$\|x - \Pi_e x\| = \text{Min } \|x - y\|, \quad y \in \text{Span } e \tag{2.5}$$

Beweis von Satz 2.15. Da $(v_1, ..., v_r)$ eine Basis ist, gilt $x = \sum_i \lambda_i v_i$ mit eindeutig bestimmten $\lambda_i \in K \Rightarrow \langle v_j, x \rangle = \sum \langle v_j, \lambda_i v_i \rangle = \lambda_j$.

□

SATZ 2.16 *Sei* $(v_1, ..., v_r)$ *ein Orthonormalsystem in V und* $U := \mathrm{Span}\,(v_1, ..., v_r)$ *der von* $v_1, ..., v_r$ *aufgespannte Unterraum. Dann lässt sich jedes* $x \in V$ *für* $u \in U$, $w \in U^\perp$ *auf genau eine Weise als Summe schreiben:*

$$x = u + w \quad \textit{mit} \quad u = \sum_{j=1}^{r} v_j \langle v_j, x\rangle \quad \textit{und} \quad w = x - u$$

Beweis zu Satz 2.16. 1) Eindeutigkeit: Wir nehmen an, es gelte $x = u + w = u' + w'$, dann folgt $\underbrace{(u' - u)}_{\in U} + \underbrace{(w' - w)}_{\in U^\perp} = 0$ und damit $0 = \|(u' - u) + (w' - w)\|^2 = \|u' - u\|^2 + \|w' - w\|^2$, wobei wir für die letzte Gleichung verwendet haben, dass wegen der Definition von U und $U^\perp$ gilt $\langle u - u', w - w'\rangle = 0$. Aus der positiven Definitheit der Norm folgt dann $u' - u = 0$ und $w' - w = 0$.
2) Existenz der Zerlegung: $u = \sum_{j=1}^{r} v_j \langle v_j, x\rangle$ ist in U wegen der Definition von U.
Zu zeigen bleibt also noch, dass $x - u =: w \in U^\perp$. Hierfür genügt es, zu zeigen, dass für $j = 1, ..., r$ gilt $\langle w, v_j\rangle = 0$, da $(v_1, \ldots v_n)$ eine Basis von U ist. Es gilt aber mit $\mu_i = \langle v_i, x\rangle$

$$\langle w, v_j\rangle = \langle x, v_j\rangle - \langle u, v_j\rangle = \langle x, v_j\rangle - \sum_{i=1}^{r} \langle \mu_i\, v_i, v_j\rangle = \langle x, v_j\rangle - \sum_{i=1}^{r} \overline{\mu_i}\, \delta_{ij} = \langle x, v_j\rangle - \langle x, v_j\rangle = 0$$

$\square$

Diese Situation wollen wir noch genauer analysieren:

DEFINITION 2.17 *Sei* V *ein Vektorraum und seien* U, W *Untervektorräume von* V.

1) *Die Menge* $U + W := \{u + w,\ u \in U,\ w \in W\}$ *heißt Summe von* U, W.
   $U + W$ *ist ein Untervektorraum von* V *und es gilt* $U + W = \mathrm{Span}\,(U \cup W)$.

2) *Ist* $U \cap W = \{0\}$, *so heißt die* ***Summe direkt***, *wir schreiben dann* $U \oplus W$.

LEMMA 2.18 *Sind* U *und* W *Unterräume eines Vektorraumes* V, *dann ist* $U + W$ *eine direkte Summe genau dann, wenn die Darstellung* $x = u + w$ *mit* $u \in U$ *und* $w \in W$ *eindeutig ist für alle* $x \in U + W$.

Beweis von Lemma 2.18. ”$\Rightarrow$” (wir verwenden Kontraposition, also $(a \Rightarrow b) \Leftrightarrow (\neg b \Rightarrow \neg a)$
Wir nehmen an, dass es für ein $x \in U + W$ zwei verschiedene Summenzerlegungen gibt, also $u + w = x = u' + w'$ mit $u, u' \in U$ und $w, w' \in W$ und $u - u' \neq 0 \neq w - w'$. Dann folgt aber, dass $u - \tilde{u} = \tilde{w} - w$ sowohl in U als auch in W sein muss, also in $U \cap W$ und damit ist $U \cap W \neq \{0\}$ und also die Zerlegung nicht direkt.
”$\Leftarrow$” (wir verwenden wieder Kontraposition):
Ist die Summe nicht direkt, dann gibt es ein $x \in U \cap W$ mit $x \neq 0$ und $x = x + 0 = 0 + x = 0,5\,x + 0,5\,x$ sind verschiedene Summenzerlegungen von x.

$\square$

Fazit: Die Zerlegung von $V = U \oplus U^{\perp}$, $U = \mathrm{Span}\,(v_1, ..., v_r)$ ist direkt.
Orthogonale Summen sind immer direkt.

Das Gram-Schmidt'sche Orthonormierungsverfahren in $(V, \langle\ ,\ \rangle)$
Gegeben sei ein linear unabhängiges Tupel $(y_1, ..., y_r)$ in V.
Gesucht ist ein ONS in V, so dass $\mathrm{Span}\,(y_1, ..., y_k) = \mathrm{Span}\,(v_1, ..., v_k)$ für alle $k = 1, ..., r$

1) Wir setzen $v_1 := \frac{y_1}{\|y_1\|}$, damit ist $U_1 := \mathrm{Span}\ v_1 = \mathrm{Span}\ y_1$ und $\|v_1\| = 1$.

2) Wir setzen $w_2 := y_2 - v_1 \langle v_1, y_2 \rangle =: y_2 - \Pi_{U_1}\, y_2$ (wir ziehen also von y_2 den Anteil ab, der in Richtung von v_1 weist) und $v_2 := \frac{w_2}{\|w_2\|}$. Dann ist wieder $\mathrm{Span}(v_1, v_2) = \mathrm{Span}(y_1, y_2) =: U_2$ und (v_1, v_2) sind orthonormal.

3) Das weitere Verfahren ist rekursiv: wir setzen für $j = 3, \ldots r$: $U_j := \mathrm{Span}(v_1, \ldots v_j)$ und

$$w_{j+1} := y_{j+1} - \sum_{k=1}^{j} v_k \langle v_k, y_{j+1} \rangle = y_{j+1} - \Pi_{U_j} y_{j+1} \qquad \text{und} \qquad v_{j+1} := \frac{w_{j+1}}{\|w_{j+1}\|}\ .$$

Damit ist $(v_1, \ldots v_j)$ orthonormal (also normiert und orthogonal) und $\mathrm{Span}\,(y_1, ..., y_j) = \mathrm{Span}\,(v_1, ..., v_j)$.

Kapitel 3

Lineare Abbildungen und Matrizen

3.1 Lineare Abbildungen

DEFINITION 3.1 *Für $\mathbb{K}$-Vektorräume V, W heißt eine Abbildung $f : V \to W$* ***linear*** *(oder* ***Homomorphismus*** *oder* ***strukturerhaltend****)* $:\Leftrightarrow \forall x, y \in V, \lambda \in \mathbb{K}$:

$$f(x+y) = f(x) + f(y) \tag{3.1}$$

$$f(\lambda x) = \lambda f(x) \tag{3.2}$$

Die Menge der linearen Abbildung von V nach W bezeichnen wir mit $\mathrm{Hom}\,(V, W)$ *oder* $\mathrm{Hom}_{\mathbb{K}}(V, W)$ *oder* $\mathcal{L}(V, W)$

Eine lineare Abbildung f wird eindeutig durch seine Werte auf einer Basis bestimmt.

SATZ 3.2 *Seien V und W $\mathbb{K}$-Vektorräume und sei $\mathcal{A} := (a_1, ..., a_n)$ eine Basis von V. Dann gibt es zu jedem n-Tupel $(w_1, ...w_n)$ in W genau eine lineare Abbildung $f : V \to W$ mit $f(a_i) = w_i,\ i = 1, ..., n$.*

Beweis. Eindeutigkeit: Seien $f, g \in \mathrm{Hom}\,(V, W)$ mit $g(a_i) = w_i = f(a_i)$, dann gilt

$$\begin{aligned} f(v) = f(x_1a_1 + ... + x_na_n) &= \sum x_i f(a_i) = \sum x_i g(a_i) \\ &= g(x_1a_1 + ... + x_na_n) = g(v) \end{aligned}$$

Existenz: Für $v = \sum x_i a_i$ setze $f(v) := \sum x_i w_i$. Da die Koordinaten $(x_i)_{i=1,...,n}$ bezüglich der Basis $\mathcal{A}$ eindeutig bestimmt sind, ist $f : V \to W$ wohldefiniert. Offensichtlich ist f linear, denn für $x = \sum x_i a_i,\ y = \sum y_i a_i$ und $\lambda \in \mathbb{K}$ gilt:

$$f(x+y) = \sum (x_i + y_i) w_i = f(x) + f(y)$$

$$\lambda f(x) = \lambda \sum x_i w_i = \sum (\lambda x_j) w_j = f((\lambda x_1)a_1) + ... + (\lambda x_n)a_n) = f(\lambda x)$$

□

NOTIZ 3.3 *1) Sind* V, W, X $\mathbb{K}$*-Vektorräume und* f, g *linear:*$V \xrightarrow{f} W \xrightarrow{g} X$*, dann ist* $g \circ f \in \mathrm{Hom}\,(V, X)$*, denn* $f(g(\lambda v)) = f(\lambda g(v)) = \lambda f(g(v))$ *und* $f(g(v+w)) = f(g(v)+g(w)) = f(g(v)) + f(g(w))$ *für alle* $v, w \in V$ *und* $\lambda \in \mathbb{K}$*.*
Die Identitätsabbildung $Id_v : V \ni v \longmapsto v \in V$ *ist linear*
2) $\mathrm{Hom}\,(V, W)$ *ist* $\mathbb{K}$*-Vektorraum bzgl. punktweiser Addition und Multiplikation mit Skalaren:*

$$(f+g)(x) = f(x) + g(x); \quad (\lambda f)(x) = \lambda f(x)$$

Beispiele zu Satz 3.2
1) ***Koordinatenabbildung*** $\phi_{\mathcal{A}} : V \longrightarrow \mathbb{K}^n$:
Sei $\mathcal{A} = (a_1, ... a_n)$ eine Basis von V und $e_i\,,\ i = 1, \dots n$ die Standardbasis in $\mathbb{K}^n$, dann gibt es genau eine Abbildung $\phi_{\mathcal{A}}$ mit $\phi_{\mathcal{A}}(a_i) = e_i,\ i = 1, ..., n$.
2) ***Duale Basis:*** Für $j = 1, \dots n$ gibt es genau eine Linearform $a^j : V \longrightarrow \mathbb{K}$ mit $a^j(a_i) = \delta_{ij},\ i = 1, ..., n$. Für $x = \sum x_i a_i$ gilt dann $a^j(x) = \sum x_i a^j(a_i) = x_j$. x_j ist die j-te Koordinate von $x \in V$ bezüglich der Basis $\mathcal{A}$. Das Tupel $\mathcal{A}' := \{a^1, \dots a^n\}$ heisst duale Basis zu $\mathcal{A}$.

DEFINITION UND SATZ 3.4 *Für* $f \in \mathrm{Hom}\,(V, W)$ *sind*

$$\begin{aligned} \mathrm{Bild}\, f &:= \{y \in W;\ \textit{Es gibt}\, x \in V\, mit\, f(x) = y\} && (3.3) \\ \mathrm{Kern}\, f &:= \{x \in V;\ f(x) = 0\} && (3.4) \end{aligned}$$

Untervektorräume von W *bzw.* V*.*

NOTIZ 3.5 $f \in \mathrm{Hom}\,(V, W)$ *injektiv* $\Leftrightarrow$ $\mathrm{Kern}\, f = \{0\}$, *denn* $f(x) = f(y) \ \Leftrightarrow\ f(x-y) = 0$

Notation:
$f \in \mathrm{Hom}\,(V, W)$ heißt

$$\begin{aligned} \text{Isomorphismus} &:\Leftrightarrow && f \text{ ist bijektiv} \\ \text{Endomorphismus} &:\Leftrightarrow && V = W \\ \text{Automorphismus} &:\Leftrightarrow && V = W\, und\, f\, bijektiv \\ \text{Epimorphismus} &:\Leftrightarrow && f \text{ ist surjektiv} \\ \text{Monomorphismus} &:\Leftrightarrow && f \text{ ist injektiv} \end{aligned}$$

NOTIZ 3.6 *Sei* $f \in \mathrm{Hom}\,(V, W)$ *ein (*$\mathbb{K}$*-Vektorraum-)Isomorphismus, dann ist die mengentheoretische Umkehrabbildung* $f^{-1} : W \longrightarrow V$ *auch linear. Es gilt also* $f^{-1} \in \mathrm{Hom}\,(W, V)$ *und ist damit ist* f^{-1} *auch ein Isomorphismus.*

SATZ 3.7 *Seien* V, W *endlichdimensionale* $\mathbb{K}$*-Vektorräume. Dann ist* $f \in \mathrm{Hom}(V, W)$ *genau dann ein Isomorphismus, wenn für jede Basis* $(a_1, ..., a_n)$ *von* V *das Tupel* $(f(a_1), ..., f(a_n))$ *eine Basis von W ist,.*

Beweis. "$\Rightarrow$": Sei $y \in W$ und $x := f^{-1}(y)$ das eindeutig bestimmte Urbild von y. Sei $x = \sum x_i a_i$ die Darstellung von x in der Basis $\mathcal{A} = (a_1, ..., a_n)$.
Dann ist $y = f(\sum x_i a_i) = \sum x_i f(a_i)$ und damit hat y also eine eindeutige Darstellung bzgl. der

Basis $(f(a_1), \ldots f(a_n))$.
"$\Leftarrow$": f ist surjektiv, denn $y = \sum y_i f(a_i) \in W$ ist das Bild von $x = \sum y_i a_i \in V$
f ist injektiv, denn für $\sum x_i a_i = x \in \mathrm{Kern}\, f$ gilt nach Definition $f(x) = \sum x_i f(a_i) = 0$ und aus der linearen Unabhängigkeit von $(f(a_1), \ldots f(a_n))$ folgt $x_1 = ... = x_n = 0$ und damit $x = 0$. Mit Notiz 3.5 folgt die Behauptung.

$\square$

NOTIZ 3.8 *1) Zwei endlichdimensionale $\mathbb{K}$-Vektorräume V, W sind isomorph (das heißt es gibt einen Isomorphismus $f : V \longrightarrow W$), genau dann wenn gilt* $\dim V = \dim W$*:*
"$\Rightarrow$": Eine Basis $(a_1, ..., a_n)$ von V liefert nach Satz 3.7 eine Basis $(f(a_1), ..., f(a_n))$ von W und damit gilt $\dim V = \dim W$
"$\Leftarrow$": Die Basen $(a_1, ..., a_n)$ von V und $(b_1, ..., b_n)$ von W induzieren eine lineare Abbildung $f : V \to W$ mit $f(a_i) = b_i,\ i = 1, , ..., n.$ und f ist ein Isomorphismus, da eine Basis in eine Basis übergeht (Satz 3.7)..
2) Insbesondere ist ein n-dimensionaler Vektorraum V isomorph zu $\mathbb{K}^n$, der Isomorphismus ist durch die Koordinatenabbildung ϕ (vgl. Beispiel zu Satz 3.2, 1) gegeben, die ja nach Definition eine Basis in eine Basis abbildet (vgl. Satz 3.7).

SATZ 3.9 (DIMENSIONSFORMEL) *Sei $f \in \mathrm{Hom}(V, W)$ und V endlichdimensional. Dann gilt:*

$$\dim \mathrm{Kern}\, f + \dim \mathrm{Bild}\, f = \dim V$$

Beweis. Sei $\dim V = n$ und $\dim \mathrm{Kern}\, f = r \leq n$ und $(a_1, ..., a_r)$ eine Basis von $\mathrm{Kern}\, f \subset V$. Wir ergänzen diese Basis von $\mathrm{Kern}\, f \subset V$ zu einer Basis $(a_1, ..., a_r, a_{r+1}, ..., a_n)$ von V (Basisergänzungssatz). Dann gilt für alle $v \in V$

$$f(v) = f(\lambda_1 a_1 + ... + \lambda_n a_n) = \sum_{k=1}^{r} \lambda_k f(a_k) + \sum_{k=r+1}^{n} \lambda_k f(a_k) = 0 + \lambda_{r+1} \underbrace{f(a_{r+1})}_{=:w_1} + ... + \lambda_n \underbrace{f(a_n)}_{=:w_{n-r}} .$$

Also gilt $\mathrm{Bild}\, f = \mathrm{Span}(w_1, ..., w_{n-r}) \subset W$.
Beh.: Das Tupel $(w_1, ..., w_{n-r})$ ist linear unabhängig.

> Bew.: Sei $\sum_{k=1}^{n-r} \mu_k w_k = 0$, dann folgt aus der Definition der w_k und der Linearität von f sofort $f(\sum_{k=r+1}^{n} \mu_k a_k) = 0$ und damit $\sum_{k=r+1}^{n} \mu_k a_k \in \mathrm{Kern}\, f$. Da aber $(a_1, \ldots, a_r)$ nach Voraussetzung eine Basis von $\mathrm{Kern}\, f$ ist, gibt es $\lambda_1, \ldots \lambda_r \in \mathbb{K}$ so dass gilt
>
> $$\sum_{k=r+1}^{n} \mu_k a_k - \sum_{k=1}^{r} \lambda_k a_k = 0$$
>
> und da nach Konstruktion das Tupel $(a_1, \ldots a_n)$ linear unabhängig ist folgt $\mu_j = 0$ und $\lambda_k = 0$ für alle $j = r + 1, \ldots n$ und $k = 1, \ldots r$.

Damit folgt, dass $(w_1, ..., w_{n-r})$ eine Basis von $\mathrm{Bild}\ f$ ist und $\dim \mathrm{Kern}\, f + \dim \mathrm{Bild}\, f = r + (n - r) = n = \dim V$

$\square$

NOTIZ 3.10 *Eine lineare Abbildung* $f : V \longrightarrow W$ *zwischen zwei Vektorräumen gleicher Dimension* $n = \dim V = \dim W$ *ist genau dann surjektiv, wenn sie injektiv ist:*
f *injektiv* $\Leftrightarrow \ \mathrm{Kern}\, f = 0 \ \Leftrightarrow \ \dim \mathrm{Kern}\, f = 0 \ \Leftrightarrow \ \dim \mathrm{Bild}\, f = n = dimW$
$\Leftrightarrow \ \mathrm{Bild}\, f = W \ \Leftrightarrow \ f$ *ist surjektiv.*

Beispiel: Sei $f : V \longrightarrow \mathbb{K}$ linear (also $f \in V^* = \mathrm{Hom}(V, K)$ eine Linearform). Dann ist für $f \neq 0$ aber $\mathrm{Bild}\, f = \mathbb{K}$, also $\dim \mathrm{Bild}\, f = 1$ und damit $\dim \mathrm{Kern}\, f = \dim V - 1$. Also ist der Unterraum $\mathrm{Kern}\, f$ eine sogenannte *Hyperebene* in V durch 0, d.h. ein Unterraum der Kodimension 1.
Die *Kodimension* eines Unterraumes $W \subset V$ ist gerade $\mathrm{codim}\, W := \dim V - \dim W$.
Für $V = \mathbb{R}^3$ wäre ein ein Unterraum der Kodimension 1 also eine echte Ebene, daher der Begriff Hyperebene für allgemeine Unterräume der Kodimension 1.

3.1.1 Matrizen und lineare Abbildungen

DEFINITION 3.11 *Eine* $m \times n$*-Matrix über einem Körper* $\mathbb{K}$ *ist eine Abbildung*

$$a : \{1, ..., m\} \times \{1, ..., n\} \to K\,, \quad (i, j) \mapsto a_{ij}\,,$$

welche als Box geschrieben werden kann:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} =: \big(a_{ij}\big)_{\binom{1\le i\le m}{1\le j\le n}} =: A$$

Wir bezeichnen die Menge aller $m \times n$*-Matrizen als* $\mathcal{M}(m \times n; \mathbb{K})$.

$$\begin{pmatrix} & a_{1j} & \\ & \vdots & \\ a_{i1} & \cdots & a_{in} \\ & \vdots & \\ & a_{mj} & \end{pmatrix} \leftarrow a_{i\bullet} \in \mathbb{K}^n \text{ bezeichnet die i-te Zeile}$$

$$\uparrow a_{\bullet j} \in \mathbb{K}^m \text{ bezeichnet die j-te Spalte}$$

Bemerkung
erster Index = Zeilenindex
zweiter Index = Spaltenindex

NOTIZ 3.12 $\mathcal{M}(m \times n; \mathbb{K})$ *ist ein* $\mathbb{K}$*-Vektorraum mit punktweiser Addition und Multiplikation mit Skalaren.*

$$\big(a_{ij}\big) + \big(b_{ij}\big) = \big(a_{ij} + b_{ij}\big)\,, \qquad \lambda\big(a_{ij}\big) = \big(\lambda a_{ij}\big)$$

THEOREM 3.13 *Jede Matrix $A \in \mathcal{M}(m\times n;\mathbb{K})$ bestimmt eindeutig eine lineare Abbildung $j(A) \in \mathrm{Hom}\,(\mathbb{K}^n,\mathbb{K}^m)$ mit der **SBE**-Regel:*

$$\textit{Die }\textbf{S}\textit{palten } a_{\bullet j} \textit{ von } A \textit{ sind die }\textbf{B}\textit{ilder der }\textbf{E}\textit{inheitsvektoren } e_j = (0,\ldots,\underbrace{1}_{j},\ldots 0) \in \mathbb{K}^n \tag{3.5}$$

Insbesondere ist die Abbildung

$$j:\,\mathcal{M}(m\times n;\mathbb{K}) \longrightarrow \mathrm{Hom}(\mathbb{K}^n,\mathbb{K}^m)\;, \quad A\mapsto j(A)$$

ein Isomorphismus und zu jeder linearen Abbildung $f \in \mathrm{Hom}(K^n,K^m)$ existiert eine korrespondierende Matrix $j^{-1}(f)$ mit (3.5).

Beweis. Die lineare Abbildung $j(A)$ zu einer Matrix A ist durch die Werte auf der Basis $\mathcal{E} = (e_1,...,e_n)$ in $\mathbb{K}^n$ bestimmt (Satz 3.2). Um zu gegebener linearer Abbildung f die Matrix $j^{-1}(f)$ zu bestimmen, schreibt man die Werte $f(e_1),...,f(e_n)$ als Spalten einer Matrix.
Es bleibt noch zu zeigen, dass die Abbildung j linear ist: Es gilt für alle $k = 1.\ldots n$:

$$j(A+B)(e_k) = (A+B)_{\cdot k} = A_{\cdot k} + B_{\cdot k} = j(A)e_k + j(B)e_k$$

und damit $j(A+B) = j(A)+j(B)$. Analog kann man zeigen $j(\lambda A) = \lambda j(A)$.

□

Beispiel zu Theorem 3.13: Wir fassen $\mathbb{C}$ als $\mathbb{R}$-Vektorraum auf mit Basis $(a_1,a_2) = (1,i)$: Dann betrachten wir die Koordinatenabbildung $\phi:\mathbb{C}\to\mathbb{R}^2$ mit $\phi(1) = e_1 = (1,0)$ und $\phi(i) = e_2 = (0,1)$ (vgl. Notiz 3.8, 2 und das Beispiel zu Satz 3.2). Die $\mathbb{R}$-Vektorräume $\mathbb{R}^2$ und $\mathbb{C}$ sind also isomorph.
Die reell-lineare Abbildung $f:\mathbb{C}\to\mathbb{C}$, sei gegeben ist durch $f(z) = \bar{z}$ mit $\bar{z} = x-iy$ für $z = x+iy$. Dann können wir zu der Abbildung $F := \phi\circ f\circ\phi^{-1}:\mathbb{R}^2\to\mathbb{R}^2$ nach Theorem 3.13 eine Matix $A := j^{-1}(F) \in \mathcal{M}(2\times 2,\mathbb{R})$ assoziieren und es gilt

$$\left.\begin{array}{r} F(e_1) = \phi\circ f(a_1) = \phi(1) = e_1 = (1,0) \\ F(e_2) = \phi\circ f(a_2) = \phi(-i) = -e_2 = (0,-1) \end{array}\right\} \;\Rightarrow\; A := j^{-1}(F) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\;.$$

NOTIZ 3.14 *1) Es gilt $j(A)(e_i) = a_{\bullet i} = \sum_{k=1}^m a_{ki}e_k$.*
2) Mit 1) folgt für beliebiges $x\in\mathbb{K}^n$

$$\begin{aligned} j(A)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} &= j(A)\Big(\sum_{i=1}^m x_ie_i\Big) = \sum_i x_i j(A)(e_i) \\ &= \sum_{i=1}^n x_i\sum_k a_{ki}\,e_k = \begin{pmatrix} a_{11}x_1+ & \cdots & +a_{1n}x_n \\ \vdots & & \vdots \\ a_{m1}x_1+ & \cdots & +a_{mn}x_n \end{pmatrix} \end{aligned}$$

Nun definieren wir das Produkt einer Matrix mit einem Spaltenvektor, um die Abbildung j loszuwerden:

DEFINITION 3.15 *Für eine Matrix $A \in \mathcal{M}(m \times n; \mathbb{K})$ und einen Vektor $x \in \mathbb{K}^n$ - diesen Raum betrachten wir als Raum von Spaltenvektoren $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ - definieren wir das Produkt durch*

$$A \cdot x := j(A)(x)$$

Bemerkung
Die Definition 3.15 erlaubt uns, die Matrix $A \in \mathcal{M}(m \times n; \mathbb{K})$ direkt mit der linearen Abbildung $j(A) \in \operatorname{Hom}(\mathbb{K}^n, \mathbb{K}^m)$ zu identifizieren. Das Produkt einer Matix mit einem Vektor verstehen wir als Anwendung der Matrix (= linearen Abbildung) auf den Vektor. Wir verwenden das Wort Matrix jetzt explizit doppeldeutig. Um diesen Schritt vollkommen rechtfertigen zu können ist es nötig, eine Matrixmultiplikation einzuführen, die der Komposition der zugehörigen linearen Abbildungen entspricht. Wir betrachten

$$\mathbb{K}^n \xrightarrow{j(A)} \mathbb{K}^m \xrightarrow{j(B)} \mathbb{K}^\ell .$$

Es gilt

$$\begin{aligned} j(B) \circ j(A)(e_i) &= j(B) \sum_{k=1}^{m} a_{ki} e_k = \sum_k a_{ki} j(B)(e_k) \qquad (3.6) \\ &= \sum_{k=1}^{m} \sum_{r=1}^{\ell} a_{ki} b_{rk} e_r = \sum_{r=1}^{\ell} \underbrace{\left(\sum_{k=1}^{m} b_{rk} a_{ki} \right)}_{=: j(B \cdot A)_{ri} \text{ gewünscht}} e_r \end{aligned}$$

Daher definieren wir das Matizenprodukt durch:

DEFINITION 3.16 *Seien $A \in \mathcal{M}(m \times n; \mathbb{K})$ und $B \in \mathcal{M}(\ell \times m; \mathbb{K})$, dann wird die Matrix mit den Elementen*

$$\big(B \cdot A\big)_{ri} := \sum_{n=1}^{m} b_{rk} a_{ki} \qquad (3.7)$$

*als das **Produkt** $B \cdot A$ von B und A bezeichnet.*

THEOREM 3.17 *Seien $A \in \mathcal{M}(m \times n; \mathbb{K})$ und $B \in \mathcal{M}(\ell \times m; \mathbb{K})$ und $B \cdot A$ wie in Definition 3.16. Dann gilt*

$$j(B) \circ j(A) = j(B \cdot A) . \qquad (3.8)$$

Insbesondere ist das Produkt zweier Matrizen assoziativ, aber nicht kommutativ.

Beweis. Die Gleichung (3.8) folgt sofort aus (3.6). Das Produkt zweier Matrizen ist assoziativ, denn die Komposition von Abbildungen ist assoziativ:

$$\begin{aligned} A \cdot (B \cdot C) &= j^{-1}\left(j(A) \circ j(B \cdot C)\right) = j^{-1}\left(j(A) \circ (j(B) \circ j(C))\right) \\ &= j^{-1}\left((j(A) \circ j(B)) \circ j(C)\right) = j^{-1}\left(j(A \cdot B) \circ j(C)\right) \\ &= (A \cdot B) \cdot C \end{aligned}$$

Da die Komposition von Abbidungen nicht kommutativ ist, ist folgt auch diese Eigenschaft für das Matrixprodukt.

□

Die Matrix einer linearen Abbildung

Seien V und W $\mathbb{K}$-Vektorräume mit den Basen $\mathcal{A} = (a_1, ..., a_n)$ und $\mathcal{B} = (b_1, ..., b_m)$. Sei $f \in \operatorname{Hom}(V, W)$ und sei

$$\phi_{\mathcal{A}} : V \to \mathbb{K}^n\,, \qquad v = \sum_{i=1}^{n} x_i a_i \;\mapsto\; \phi_{\mathcal{A}}(v) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \tag{3.9}$$

die Koordinatenabbildung (ein Isomorphismus) zur Basis $\mathcal{A}$ (vgl. Beispiel zu Satz 3.2 und Notiz 3.8). Analog definieren wir die Koordinatenabbildung $\phi_{\mathcal{B}} : W \to \mathbb{K}^m$.
Dann induziert die lineare Abbildung f zunächst eine lineare Abbildung $j\big(\phi_{\mathcal{B}} \circ f \circ \phi_{\mathcal{A}}\big)$ von $\mathbb{K}^n$ nach $\mathbb{K}^m$ und mit der oben beschriebenen Identifizierung wiederum eine Matrix $A = \phi_{\mathcal{B}}^{\mathcal{A}}(f) \in \mathcal{M}(m \times n, \mathbb{K})$. Das Diagramm

$$\begin{array}{ccc} V & \xrightarrow{\;f\;} & W \\ {\scriptstyle \phi_{\mathcal{A}}}\big\downarrow & & \big\downarrow{\scriptstyle \phi_{\mathcal{B}}} \\ \mathbb{K}^n & \xrightarrow[A]{} & \mathbb{K}^m \end{array} \tag{3.10}$$

ist kommutativ. Es gilt für $\phi_{\mathcal{B}}^{\mathcal{A}}(f)$

$$\begin{aligned} \text{k-te Spalte von } \phi_{\mathcal{B}}^{\mathcal{A}}(f) &= (\phi_{\mathcal{B}} \circ f \circ \phi_{\mathcal{A}}^{-1})(e_k) = \phi_{\mathcal{B}}\left(f(a_k)\right) \\ &= \phi_{\mathcal{B}}(a_{1k} b_1 + ... + a_{mk} b_m) = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{pmatrix} \end{aligned}$$

Damit gilt das folgende

THEOREM 3.18 *Seien V und W Vektorräume über $\mathbb{K}$ mit Basen $\mathcal{A} = (a_1, ..., a_n)$ und $\mathcal{B} = (b_1, ..., b_m)$. Seien $\phi_{\mathcal{A}}$ und $\phi_{\mathcal{B}}$ die zugehörigen Koordinatenabbildungen (siehe* (3.9)*). Dann ist die Abbildung*

$$\phi_{\mathcal{B}}^{\mathcal{A}} \;:\; \operatorname{Hom}_{\mathbb{K}}(V, W) \to \mathcal{M}(m \times n, \mathbb{K})\,, \quad \phi_{\mathcal{B}}^{\mathcal{A}}(f) = \phi_{\mathcal{B}} \circ f \circ \phi_{\mathcal{A}}^{-1}$$

ein Isomorphismus. Für $f \in \operatorname{Hom}(V, W)$ sind die Spalten der Matrix $A := \phi_{\mathcal{B}}^{\mathcal{A}}(f)$ gerade die Koordinaten (bzgl. $\mathcal{B}$) der Bilder (unter f) der Basisvektoren aus $\mathcal{A}$.
Das Diagramm (3.10) *ist dann kommutativ und es gilt*

$$\phi_{\mathcal{B}}^{\mathcal{A}}(f) = \big(a_{ik}\big) \quad \textit{mit Spalten } a_{\bullet k} = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{pmatrix} \quad \textit{gegeben durch} \quad f(a_k) = \sum_{i=1}^{m} a_{ik} b_i\ .$$

Merkregel: SKBB = Die **S**palten sind die **K**oordinaten der **B**ilder der **B**asisvektoren.

Beweis von Theorem 3.18. Da die Koordinatenabbildungen bijektiv sind, ist auch $\phi^{\mathcal{A}}_{\mathcal{B}}$ bijektiv und

$$\left(\phi^{\mathcal{A}}_{\mathcal{B}}\right)^{-1}(A) = \phi_{\mathcal{B}}^{-1} \circ A \circ \phi_{\mathcal{A}}$$

Die Linearität von $\phi^{\mathcal{A}}_{\mathcal{B}}$ folgt sofort aus der Linearität der Koordinatenabbildungen und der Tatsache, dass die Komposition von linearen Abbildungen linear ist (vgl. Notiz 3.3). □

Beispiel: siehe das Beispiel zu Theorem 3.13 (Komplex-Konjuation).

DEFINITION 3.19 *Sei* $A \in \mathcal{M}(n \times n; \mathbb{K})$. *Dann ist* $A \in \mathrm{GL}(n, \mathbb{K}) \quad :\Longleftrightarrow$ *Es existiert* $B \in \mathcal{M}(n \times n, \mathbb{K})$ *mit*

$$A \cdot B = B \cdot A = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} =: \mathbf{1}$$

Wir bezeichnen dann B *mit* A^{-1}.

SATZ 3.20 *Sei* $A \in \mathrm{GL}(n, \mathbb{K})$, *dann ist* A^{-1} *eindeutig und* A *ist ein Isomorphismus auf* $\mathbb{K}^n$. *Weiterhin ist* $\big(\mathrm{GL}(n, \mathbb{K}), \cdot\big)$ *eine Gruppe (die "General Linaear Group").*

Beweis. selber! □

SATZ 3.21 *Seien* U, V *und* W $\mathbb{K}$*-Vektorräume mit den Basen* $\mathcal{A}, \mathcal{B}$ *und* $\mathcal{C}$ *und* $f \in \mathrm{Hom}(U, V)$, $g \in \mathrm{Hom}(V, W)$. *Dann ist mit* $A = \phi^{\mathcal{A}}_{\mathcal{B}}$ *und* $B = \phi^{\mathcal{B}}_{\mathcal{C}}$ *das Diagramm*

$$\begin{tikzcd} U \arrow[r, "f"] \arrow[d, "\phi_{\mathcal{A}}"'] & V \arrow[r, "g"] \arrow[d, "\phi_{\mathcal{B}}"] & W \arrow[d, "\phi_{\mathcal{C}}"] \\ \mathbb{K}^n \arrow[r, "A"'] & \mathbb{K}^m \arrow[r, "B"'] & \mathbb{K}^\ell \end{tikzcd} \tag{3.11}$$

kommutativ und es gilt $\phi^{\mathcal{A}}_{\mathcal{C}}(g \circ f) = \phi^{\mathcal{B}}_{\mathcal{C}}(g)\ \phi^{\mathcal{A}}_{\mathcal{B}}(f)$.

Beweis . Man liest aus dem kommutativen Diagramm ab:

$$\begin{aligned} \phi^{\mathcal{A}}_{\mathcal{C}}(g \circ f) &= \phi_{\mathcal{C}} \circ g \circ f \circ \phi_{\mathcal{A}}^{-1} = \left(\phi_{\mathcal{C}} \circ g \circ \phi_{\mathcal{B}}^{-1}\right) \circ \left(\phi_{\mathcal{B}} \circ f \circ \phi_{\mathcal{A}}^{-1}\right) \\ &= \phi^{\mathcal{B}}_{\mathcal{C}}(g)\ \phi^{\mathcal{A}}_{\mathcal{C}}(f) \end{aligned}$$

□

Koordinatenwechsel

Sei V ein $\mathbb{K}$-Vektorraume mit den Basen $\mathcal{A}$ und $\mathcal{A}'$ und W ein $\mathbb{K}$-Vektorraum mit den Basen $\mathcal{B}$ und $\mathcal{B}'$. Wir betrachten die kommutativen Diagramme

$$\begin{tikzcd} V \arrow[r, "\mathrm{id}"] \arrow[d, "\phi_{\mathcal{A}}"'] & V \arrow[d, "\phi_{\mathcal{A}'}"] \\ \mathbb{K}^n \arrow[r, "\phi^{\mathcal{A}}_{\mathcal{A}'}(\mathrm{id})"'] & \mathbb{K}^n \end{tikzcd} \qquad \begin{tikzcd} W \arrow[r, "\mathrm{id}"] \arrow[d, "\phi_{\mathcal{B}}"'] & W \arrow[d, "\phi_{\mathcal{B}'}"] \\ \mathbb{K}^m \arrow[r, "\phi^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id})"'] & \mathbb{K}^m \end{tikzcd} \tag{3.12}$$

Dann gilt

$$\phi^{\mathcal{A}}_{\mathcal{A}'}(\mathrm{id}) = \phi_{\mathcal{A}'} \circ \phi_{\mathcal{A}}^{-1} \quad \text{und} \quad \phi^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id}) = \phi_{\mathcal{B}'} \circ \phi_{\mathcal{B}}^{-1}$$

SATZ 3.22 *Sei $f \in \mathrm{Hom}(V, W)$ und seien $\mathcal{A}$ und $\mathcal{A}'$ Basen von V und $\mathcal{B}$ und $\mathcal{B}'$ Basen von W. Dann gilt:*

$$\phi^{\mathcal{A}}_{\mathcal{B}}(f) = \phi^{\mathcal{A}}_{\mathcal{B}}(\mathrm{id}_W \circ f \circ \mathrm{id}_V) = \phi^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{id}_W) \circ \phi^{\mathcal{A}'}_{\mathcal{B}'}(f) \circ \phi^{\mathcal{A}}_{\mathcal{A}'}(\mathrm{id}_V)$$

entsprechend dem kommutativen Diagramm

$$\begin{tikzcd} & V \arrow[r, "f"] \arrow[d, "\phi_{\mathcal{A}'}"] \arrow[ddl, "\phi_{\mathcal{A}}"'] & W \arrow[d, "\phi_{\mathcal{B}'}"'] \arrow[ddr, "\phi_{\mathcal{B}}"] & \\ & \mathbb{K}^n \arrow[r, "\phi^{\mathcal{A}'}_{\mathcal{B}'}"] \arrow[dl] & \mathbb{K}^m \arrow[dr] & \\ \mathbb{K}^n \arrow[rrr, "\phi^{\mathcal{A}}_{\mathcal{B}}(f)"] & & & \mathbb{K}^m \end{tikzcd}$$

3.1.2 Der Rang einer Matrix

DEFINITION 3.23 *Für $f \in \mathrm{Hom}(V, W)$ ist der **Rang** von f definiert als die Dimension des Bildes von f, d.h. es gilt* $\mathrm{Rang}\, f = \dim \mathrm{Bild}\, f$.

*Für $A \in \mathcal{M}(m \times n, \mathbb{K})$ ist der **Zellenrang** bzw. **Spaltenrang** definiert als die maximale Anzahl linear unabhängiger Zeilen bzw. Spalten von A. Der Rang einer Matrix ist der Rang der assoziierten linearen Abbildung, also* $\operatorname{Rang} A = \dim \operatorname{Bild} A$.

NOTIZ 3.24 *Da für $A \in \mathcal{M}(m \times n, \mathbb{K})$ gilt* $\operatorname{Bild} A = \operatorname{Span} \underbrace{\{a_{\bullet 1}, ..., a_{\bullet n}\}}_{\text{Spaltenvektoren von } A}$ *folgt*

$$\begin{aligned} \operatorname{Rang} A &= \dim \operatorname{Bild} A = \textit{maximale Anzahl linear unabhängiger Spaltenvektoren} \\ &= \textit{Spaltenrang } A \end{aligned} \tag{3.13}$$

Nach Definition gilt

$$\textit{Zeilenrang } A := \dim \operatorname{Span} \underbrace{\{a_{1\bullet}, ..., a_{n\bullet}\}}_{\text{Zeilenvektoren von } A} . \tag{3.14}$$

DEFINITION 3.25 *Zwei Matrizen $A, B \in \mathcal{M}(m \times n; \mathbb{K})$ heißen **äquivalent**, geschrieben $A \sim B :\Leftrightarrow$ Es gibt invertierbare Matrizen $P \in \mathrm{GL}(n, \mathbb{K})$ und $Q \in \mathrm{GL}(m, \mathbb{K})$, sodass das Diagramm*

$$\begin{tikzcd} \mathbb{K}^n \arrow[r, "A"] \arrow[d, shift right=2, "Q^{-1}"'] & \mathbb{K}^m \arrow[d, shift left=2, "P^{-1}"] \\ \mathbb{K}^n \arrow[r, "B"] \arrow[u, shift right=2, "Q"'] & \mathbb{K}^m \arrow[u, shift left=2, "P"] \end{tikzcd}$$

kommutativ ist. Das heißt, es gilt: $B = P^{-1} \cdot A \cdot Q$.

NOTIZ 3.26 *1) $\sim$ ist die einfachste und gröbste Äquivalenzrelation für Matrizen.*
2) P und Q können wir jeweils auffassen als Matrix eines Basiswechsels:
Bezeichnet $\mathcal{E} = (e_1, \ldots e_n)$ die Standardbasis in $\mathbb{K}^n$, dann ist $Pe_i = p_{\bullet i}$ der i-te Spaltenvektor von P. Da P invertierbar ist, gilt $\dim \operatorname{Bild} P = n$, die Spalten von P sind also linear unabhängig und $\mathcal{P} := (p_{\bullet 1}, \ldots, p_{\bullet n})$ ist eine andere Basis von $\mathbb{K}^n$. Mit der SKBB-Regel und da $Pe_j = P_{\bullet j} = P_{1j}e_1 + ... + P_{nj}e_n$ gilt, folgt $P = \phi^{\mathcal{E}}_{\mathcal{P}}(\mathrm{id}_{\mathbb{K}^n})$. Damit folgt:

$$A \sim B \;\Leftrightarrow\; \textit{Es gibt Basen } \mathcal{P} \textit{ von } \mathbb{K}^n \textit{ und } \mathcal{Q} \textit{ von } \mathbb{K}^m \textit{ mit } B = \phi^{\mathcal{P}}_{\mathcal{Q}}(A) .$$

SATZ 3.27 *Sei $A \in \mathcal{M}(m \times n; \mathbb{K})$ und $\operatorname{Rang} A = r$. Dann ist $r \leq \min\{m, n\}$ und mit $\mathbf{1}_r \in \mathcal{M}(r \times r; K)$ gilt:*

$$A \sim \begin{pmatrix} 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix} = \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix} =: E_r$$

Ist $A \in \mathcal{M}(n \times n, \mathbb{K})$, dann gilt: $A \in \mathrm{GL}(n, \mathbb{K}) \;\Leftrightarrow\; \operatorname{Rang} A = n$.

Beweis. Es genügt zu zeigen: In geeigneten Basen $\mathcal{A}, \mathcal{B}$ von $\mathbb{K}^n$ bzw. $\mathbb{K}^m$ gilt:

$$E_r = \phi^{\mathcal{A}}_{\mathcal{B}}(A) = \phi^{\mathcal{E}}_{\mathcal{B}}(\mathrm{id})\ A\ \phi^{\mathcal{A}}_{\mathcal{E}}(\mathrm{id})$$

Mit der Dimensionsformel gilt $\dim \mathrm{Kern}\, A = n - \mathrm{Rang}\, A = n - r$. Mit dem Basisergänzungssatz ergänzen wir eine Basis $(a_{r+1}, ..., a_n)$ von $\mathrm{Kern}\, A$ zu einer Basis $(a_1, ..., a_n)$ von $\mathbb{K}^n$. Dann ist $(Aa_1, ..., Aa_r) =: (b_1, ..., b_r)$ eine Basis von $\mathrm{Bild}\, A$, die wir zu einer Basis $(b_1, ..., b_m)$ von $\mathbb{K}^m$ ergänzen. Daraus folgt $r \leq \min\{n, m\}$ und

$$Aa_j = \begin{cases} b_j, \ j \leq r \\ 0, \ sonst \end{cases} \qquad \text{also: } \phi^{\mathcal{A}}_{\mathcal{B}}(A) = \begin{pmatrix} 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix}$$

Falls $A \in \mathrm{GL}(n, \mathbb{K})$, dann existiert nach Definition nach Definition eine inverse Matrix A^{-1}, die nach Satz 3.20 eindeutig ist. Damit ist A ein Isomorphismus auf $\mathbb{K}^n$ und bildet nach Satz 3.7 eine Basis in eine Basis ab. Also bilden die Spalten von A eine Basis von $\mathbb{K}^n$ und sind insbesondere linear unabhängig. Hieraus folgt sofort, dass der Spaltenrang und damit der Rang von A gleich n ist. Ist andererseits Rang $A = n$, dann bilden die Spalten von A eine Basis von $\mathbb{K}^n$ und nach Satz 3.7 ist A ein Isomorphismus. Nach Definition existiert dann eine inverse Matrix A^{-1}.

□

KOROLLAR 3.28 (RANGSATZ) *Für $A, B \in \mathcal{M}(m \times n; \mathbb{K})$ gilt:*

$$A \sim B \Leftrightarrow \mathrm{Rang}\ A = \mathrm{Rang}\ B\ .$$

Beweis. "⇒": Aus dem Diagramm liest man ab: $\mathrm{Bild}\ B \cong \mathrm{Bild}\ A$.
"⇐": Sei $r := \mathrm{Rang}\, A = \mathrm{Rang}\, B$, dann gilt $A \sim E_r$ und $E_r \sim B$, also $A \sim B$. □

Bemerkung: Bezüglich der Äquivalenzrelation $\sim$ haben wir damit die Matrizen in $\mathcal{M}(m \times n; \mathbb{K})$ klassifiziert:
1) Durch die Angabe einer **Normalform** E_r, denn alle $B \in [A]$ sind äquivalent zu E_r, wobei $r = \mathrm{Rang}\ A$.
2) Durch die Angabe eines **charakteristischen Datums**

$$c : \mathcal{M}(m \times n; \mathbb{K}) \longrightarrow \mathbb{N} \text{ mit: } A \sim B \Leftrightarrow c(A) = c(B)$$

$c = \mathrm{Rang}$ ist ausreichend.

DEFINITION 3.29 *Die zu $A \in \mathcal{M}(m \times n; \mathbb{K})$ **transponierte Matrix** $A^t \in \mathcal{M}(n \times m; \mathbb{K})$ entsteht durch das Vertauschen von Zeilen und Spalten, d.h. es gilt $a^t_{ij} := a_{ji}$ für alle $j = 1 \ldots m$ und $i = 1, \ldots n$.*

Offensichtlich gilt:

$$\mathrm{Rang}\ E_r^t = \begin{pmatrix} 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix} = r = \mathrm{Rang}\ E_r \tag{3.15}$$

LEMMA 3.30 *1. Für $A \in \mathcal{M}(m \times n; \mathbb{K}), B \in \mathcal{M}(n \times l; \mathbb{K})$ gilt $(A \cdot B)^t = B^t \cdot A^t$.*

2. Für Isomorphismen $A \in \mathrm{GL}(n\ \mathbb{K})$ gilt $(A^t)^{-1} = (A^{-1})^t$. Insbesondere ist A^t wiederum ein Isomorphismus.

Beweis. ad 1)

$$\left[(AB)^t\right]_{ji} = (AB)_{ij} = \sum_r A_{jr}B_{ri} = \sum_r (A^t)_{rj}(B^t)_{jr} = (B^t \cdot A^t)_{ij}$$

ad 2) Es gilt $A^t \cdot (A^{-1})^t = (A^{-1} \cdot A)^t = \mathbf{1}^t = \mathbf{1}$ und analog $(A^{-1})^t \cdot A^t = \mathbf{1}$. Damit folgt $(A^{-1})^t = (A^t)^{-1}$.

□

Damit erhalten wir aus dem Rangsatz das

KOROLLAR 3.31 *Für $A \in \mathcal{M}(m \times n; \mathbb{K})$ gilt Zeilenrang A =Spaltenrang $A = \mathrm{Rang}\, A$.*

Beweis. Sei (mit Notiz 3.24) $r = \mathrm{Rang}\, A =$ Spaltenrang$A$, dann gibt es Isomorphismen P, Q in $\mathbb{K}^m$ bzw. $\mathbb{K}^n$ mit $A = PE_rQ$. Daraus folgt mit eqrefranger und Lemma 3.30

$$\begin{aligned} \text{Zeilenrang}\, A &= \mathrm{Rang}\, A^t = \mathrm{Rang}(PE_rQ)^t = \mathrm{Rang}\ Q^tE_r^tP^t \\ &= \mathrm{Rang}\ E_r^t = \mathrm{Rang}\ E_r = \text{Spaltenrang}\, A \end{aligned}$$

□

3.2 Lineare Gleichungssysteme

3.2.1 Lösbarkeit linearer Gleichungssysteme

DEFINITION 3.32 *Für $A \in \mathcal{M}(m \times n; \mathbb{K}),\ b \in \mathbb{K}^m$ heißt*

$$Ax = b \tag{3.16}$$

***lineares Gleichungssystem** mit m Gleichungen und n Unbekannten $(x_1, ..., x_n) = x \in \mathbb{K}^n$. b heißt **Inhomogenität** des Gleichungssystems. Für $b = 0$ heißt das Gleichungssystem **homogen**. Die **Lösungsmenge** ist gegeben durch Lös$(A, b) := \{x \in \mathbb{R}^n; Ax = b\}$.*

DEFINITION 3.33 *Sei V ein Vektorraum und W ein Untervektorraum von V. Dann heisst $M \subset V$ **affiner Teilraum** von V zu W :⇔ $\exists x_0 \in V$ mit $M = x_0 + W$.*

NOTIZ 3.34 *Ist $x_0 \in \mathbb{K}^n$ eine Lösung von* (3.16)*, gilt also $Ax_0 = b$, so gilt*

$$\textit{Lös}(A, b) = x_0 + \mathrm{Kern}\ A := \{x_0 + x \,|\, x \in \mathrm{Kern}\ A\}\ . \tag{3.17}$$

Damit ist Lös(A, b) entweder leer oder ein affiner Teilraum von $\mathbb{K}^n$ zum Unterraum Kern A.
Um (3.17) *einzusehen, betrachten wir ein beliebiges $x_0 \in$ Lös(A, b). Dann gilt Lös$(A, b) \supset x_0 +$ Kern A, denn*

$$\forall y \in \mathrm{Kern}\ A: \quad A(x_0 + y) = Ax_0 + Ay = b + 0\ .$$

Es gilt aber auch Lös$(A, b) \subset x_0 +$ Kern A, denn

$$\begin{aligned} \forall v \in \textit{Lös}(A, b): \qquad A(v - x_0) &= Av - Ax_0 = b - b = 0 \\ &\Rightarrow v - x_0 \in Kern\ A \\ &\Rightarrow v \in [x_0 + Kern\ A] \end{aligned}$$

SATZ 3.35 *Das lineare Gleichungssystem $Ax = b$ (siehe* (3.16)*) ist genau dann lösbar, wenn* Rang $A =$ Rang $(A|b)$ *gilt, wobei $(A|b) \in \mathcal{M}(m \times (n+1); \mathbb{K})$ die sogenannte **erweiterte Matrix** ist, die gegeben ist durch*

$$(A|b) := \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_n \end{pmatrix}\ .$$

Beweis:. "$\Rightarrow$":
$Ax = b$ lösbar $\Leftrightarrow b \in$ Bild $A =$ Span$(a_{\cdot 1}, ..., a_{\cdot n}) \Rightarrow$ Span$(a_{\cdot 1}, ..., a_{\cdot n}) =$ Span$(a_{\cdot 1}, ..., a_{\cdot n}, b)$

"$\Leftarrow$":
Sei Span$(a_{\cdot 1}, ..., a_{\cdot n}, b) =: W$, dann ist W ein Untervektorraum von $\mathbb{K}^m$ und Bild $A \subset W$. Nach Voraussetzung gilt aber dim $W =$ dim Bild A und damit $\Rightarrow$ Bild $A = W$. Da nach Definition $b \in W$ folgt sofort $b \in$ Bild A und damit ist $Ax = b$ lösbar. □

NOTIZ 3.36 *Ist x_0 eine Lösung von* (3.16) *und $(v_1, ..., v_r)$ eine Basis von* Kern A, *so gilt nach Notiz 3.34*

$$\textit{Lös}(A, b) = \{x_0 + \lambda_1 v_1 + ... + \lambda_r v_r;\ \lambda_i \in \mathbb{K}\}, \tag{3.18}$$

*(dies ist die **Parameterdarstellung** von $x_0 +$ Kern A). Dabei ist nach der Dimensionsformel $r =$* dim Kern $A = n -$ Rang A.

NOTIZ 3.37 *Ein lösbares Gleichungssystem $Ax = b$ ist genau dann* **eindeutig lösbar**, *falls* Kern $A = \{0\}$ *ist (d.h. falls* Rang $A = n$ *gilt). Dies folgt sofort aus Notiz 3.34, da dann für jede Lösung gilt $x = x_0 + 0$, falls x_0 eine Lösung ist.*

Erinnerung: Ein homogenes lineares Gleichungssystem mit mehr Unbekannten als Gleichungen ist niemals eindeutig lösbar (Fundamental-Lemma, lieferte den Begriff der Basis).

3.2.2 Gaußverfahren

DEFINITION 3.38 *Es gibt drei Typen* ***elementarer Zeilenumformungen*** *von $A \in \mathcal{M}(m \times n, \mathbb{K})$:*

Typ 1) Vertauschen zweier Zeilen.

Typ 2) Multiplikation einer Zeile mit $\lambda \neq 0,\ \lambda \in K$.

Typ 3) Addition eines beliebigen Vielfachen einer Zeile zu einer anderen Zeilen.

Analog sind ***elementare Spaltenumformungen*** *definiert.*

SATZ 3.39 *Seien $A, A' \in \mathcal{M}(m \times n; \mathbb{K})$ und $b, b' \in \mathbb{K}^m$. Lässt sich $(A|b)$ durch elementare Zeilenumformungen auf $(A'|b')$ abbilden (wir schreiben dann kurz $(A|b) \sim (A'|b')$), so gilt Lös$(A|b) =$ Lös$(A'|b')$.*

Beweis:. Wir betrachten einzeln die Fälle, dass $(A'|b')$ aus $(A|b)$ durch elementare Zeilenumformungen vom Typ k für $k = 1, 2$ oder 3 hervorgeht.

Typ 1) klar!

Typ 2) $\forall \lambda \neq 0:\ a_1x_1 + ... + a_nx_n = b \Leftrightarrow \lambda a_1x_1 + ... + \lambda a_nx_n = \lambda b$

Typ 3)

$$\left.\begin{array}{ll} (\text{Zeile } i) & a_{i1}x_1 + ... + a_{in}x_n = b_i \\ (\text{Zeile } j) & a_{j1}x_1 + ... + a_{jn}x_n = b_j \end{array}\right\} \overset{\forall \lambda \in \mathbb{K}}{\Rightarrow} \left.\begin{array}{lr} (\text{Zeile } i) & a_{i1}x_1 + ... + a_{in}x_n = b_i \\ (\text{Zeile } j) & a_{j1}x_1 + ... + a_{jn}x_n = b_j \\ (\lambda \cdot \text{Zeile } i) & \lambda a_{i1}x_1 + ... + \lambda a_{in}x_n = \lambda b_i \end{array}\right\}$$

$$\Rightarrow \left.\begin{array}{cr} (\text{Zeile } i) & a_{i1}x_1 + ... + a_{in}x_n = b_i \\ (\text{Zeile } j + \lambda \cdot \text{Zeile } i) & (a_{j1} + \lambda a_{i1})x_1 + \ldots + (a_{jn} + \lambda a_{in})x_n = b_j + \lambda b_i \end{array}\right\}$$

Andersherum gilt analog $(i) \wedge [(j) + \lambda(i)] \Rightarrow (i) \wedge [(j) + \lambda(i) - \lambda(i)] \Rightarrow (i) \wedge (j)$.

□

NOTIZ 3.40 *1. Durch $(A|b) \sim (A'|b')$ wird eine Äquivalenzrelation auf $\mathcal{M}(m \times n; \mathbb{K}) \times \mathbb{K}^m$ erklärt.*

2. Jede elementare Zeilenumformung entspricht einer Multiplikation von $A \in \mathcal{M}(m \times n; \mathbb{K})$ von links mit einer ***invertierbaren Matrix*** *$P \in \mathrm{GL}(m; \mathbb{K})$, einer sogenannten* ***Elementarmatrix****.*

1) **Permutationsmatrix**

$$P_i^j \text{ vertauscht Zeile i und j: } \begin{pmatrix} \vdots \\ a_i \\ a_j \\ \vdots \end{pmatrix} \mapsto \begin{pmatrix} \vdots \\ a_j \\ a_i \\ \vdots \end{pmatrix}$$

2) $\forall \lambda \in K: \; T_i(\lambda): \; e_k \mapsto \begin{cases} e_k & k \neq i \\ \lambda e_k & k = i \end{cases}$

$$\Rightarrow T_i(\lambda) = \begin{pmatrix} 1 & & & \vdots & & & \\ & \ddots & & \vdots & 0 & & \\ & & 1 & \vdots & & & \\ \cdots & \cdots & \cdots & \lambda & \cdots & \cdots & \cdots \\ & & & \vdots & 1 & & \\ & & 0 & \vdots & & \ddots & \\ & & & \vdots & & & 1 \end{pmatrix}$$

3) $N_i^j(\lambda): \begin{pmatrix} \vdots \\ a_i \\ a_j \\ \vdots \end{pmatrix} \mapsto \begin{pmatrix} \vdots \\ a_i + \lambda a_j \\ a_i \\ \vdots \end{pmatrix}$

Bemerkung: Für $A \in \mathcal{M}(m \times n; \mathbb{K}), \; B \in \mathrm{GL}(m; \mathbb{K})$ gilt $\text{Lös}(A|b) = \text{Lös}(BA|Bb)$ (allgemein gilt $B(A|b) = (BA|Bb)$). Als Spezialfall ergibt sich wieder die Invarianz des Lösungsraums unter elementaren Zeilenumformungen.

Warnung: Elementare Spaltenumformungen ändern den Lösungsraum und zwar durch Umnummerieren der Unbekannten $x_1, ..., x_n$. Der Rang bleibt dabei invariant.

LEMMA 3.41 *Jede Matrix $A \in \mathcal{M}(m \times n; \mathbb{K})$ kann durch endlich viele elementare Zeilenumformungen in eine Matrix in* ***Zeilenstufenform*** *B überführt werden, wobei*

$$B = \begin{pmatrix} 0\ldots & \underline{|b_{1j_1}} & * & \cdots & & \\ 0\ldots & 0 & \underline{|b_{2j_2}} & * & \cdots & \\ 0\ldots & 0 & 0 & \underline{|b_{3j_3}} & * & \cdots \\ 0\ldots & 0 & 0 & 0 & \ddots & \cdots \\ 0\ldots & 0 & 0 & 0 & 0 & \underline{|b_{rj_r}} \\ 0\ldots & 0 & 0 & 0 & 0 & 0 \\ & & & \vdots & & \\ 0\ldots & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.19}$$

mit $b_{1j_1}, ..., b_{rj_r} \neq 0$. Man liest ab: $\text{Rang } A = \text{Rang } B = r$ *=Zeilenrang B*

Beweisskizze von Lemma 3.41. 1. Schritt

- Falls $a_{11} \neq 0$, dann ist $a_{11} = b_{11}$. Addiere für $j = 2, \ldots m$ das $\left(-\frac{a_{j1}}{a_{11}}\right)$-fache von Zeile 1 zu Zeile j. Das liefert

$$\left(\begin{array}{c|c} a_{11} & \\ * & \end{array} \right) \sim \left(\begin{array}{c|ccc} a_{11} & & & * \\ 0 & a_{22} - \frac{a_{21}}{a_{11}} & & \\ \vdots & & & \ddots \\ 0 & & * & \end{array} \right) \tag{3.20}$$

- Falls $a_{11} = 0$ und ein Element der ersten Spalte $a_{j1} \neq 0$: Zeilenvertauschung, dann weiter wie oben. Es ist dann $b_{11} = a_{j1}$.

- Falls $a_{j1} = 0$ für alle j (also die gesamte erste Spalte $a_{\cdot 1} = 0$), dann weiter bis zur ersten Spalte $a_{\cdot j_1}$, die nicht null ist. Dann genauso verfahren wie oben, wobei 1 durch j_1 ersetzt wird. j_1 ist dabei der Index der ersten Spalte, die $\neq 0$ ist. Dieses Verfahren liefert b_{1j_1}.

2. Schritt
Iteration des Verfahrens für die Submatrix A' in

$$\tilde{A} = \left(\begin{array}{c|c} b_{1j_1} & * \\ 0 & A' \end{array} \right) \tag{3.21}$$

STOP auf Zeilenstufenform B von A.

□

NOTIZ 3.42 *Das Gleichungssystem $Ax = b$ ist genau dann unlösbar, wenn die Zeilenstufenform der erweiterten Matrix $(A|b)$ von der Gestalt $\left(\begin{array}{c|c} * & b' \\ 0 & b'' \end{array} \right)$ mit $b'' \neq 0$ ist. Dies folgt aus Satz 3.35, der besagt, dass $Ax = b$ genau dann unlösbar, wenn* $\operatorname{Rang} A < \operatorname{Rang}(A|b)$ *gilt (der Rang der erweiterten Matrix kann nie echt kleiner sein als der Rang der ursprünglichen Matrix). Dies ist aber genau dann der Fall, wenn $b'' \neq 0$ gilt.*

Bemerkung: Falls $Ax = b$ lösbar, liefern die Zeilenstufenform und eventuelle Spaltenvertauschungen

$$(A|b) \sim \left(\begin{array}{ccc|c|c} \tilde{a}_{11} & \cdots & & & \\ 0 & \ddots & & \tilde{S} & \tilde{b} \\ 0 & 0 & \tilde{a}_{rr} & & \\ \hline & 0 & & 0 & 0 \end{array} \right) \tag{3.22}$$

mit $\tilde{a}_{jj} \neq 0, j = 1, ..., r,\ r = \operatorname{Rang} A$ und weiter

$$\sim \left(\begin{array}{c|c|c} \mathbf{1} & S & b' \\ \hline 0 & 0 & 0 \end{array} \right) =: B \tag{3.23}$$

Jetzt kann man die Lösung ablesen:

$$\text{Lös}\, B = w_0 + \operatorname{Kern} B \quad \text{mit} \quad w_0 = \begin{pmatrix} b' \\ 0 \end{pmatrix}, \quad \text{denn} \quad \begin{pmatrix} \mathbf{1} & S \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} b' \\ 0 \end{pmatrix} = \begin{pmatrix} b' \\ 0 \end{pmatrix}.$$

NOTIZ 3.43 *Eine Basis* $(w_1, ..., w_n)$ *von* $\mathrm{Kern}\begin{pmatrix} \mathbf{1} & S \\ 0 & 0 \end{pmatrix}$ *ergibt sich mit dem Ansatz:*

$$(w_1, ..., w_n) = \begin{pmatrix} y_{11} & & y_{1k} \\ \vdots & & \vdots \\ y_{r1} & & y_{rk} \\ 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}, \qquad k = n - r\,. \tag{3.24}$$

Für jede Spalte ergibt sich:

$$\begin{pmatrix} \mathbf{1} & S \\ 0 & 0 \end{pmatrix} w_j = 0 \quad \Leftrightarrow \quad \begin{pmatrix} y_{1j} \\ \vdots \\ y_{rj} \end{pmatrix} + \begin{pmatrix} s_{1j} \\ \vdots \\ s_{rj} \end{pmatrix} = 0\,.$$

Also kann man die Lösung in der Zeilenstufenform ablesen.

3.3 Übungen

1. Beweisen Sie Satz 3.4!

2. Beweisen Sie Notiz 3.19!

Kapitel 4

Determinanten

4.1 Geometrische Motivation im $\mathbb{R}^2$ und $\mathbb{R}^3$

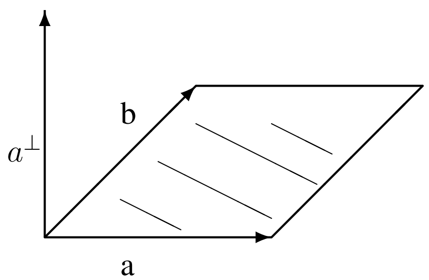
Determinanten kann man algebraisch über das Lösen von Gleichungssystemen motivieren. In ihrer Anwendung sind sie sehr bedeutsam für die Geometrie.

4.1.1 Determinante in $\mathbb{R}^2$

Sei $a = (a_1, a_2) \in \mathbb{R}^2$ ein beliebiger Vektor, dann steht der Vektor $a^\perp = (-a_2, a_1)$ bzgl. des Standard-Skalarproduktes $\langle x, y \rangle := x_1 y_1 + x_2 y_2$ senkrecht auf a und hat dieselbe Länge, d.h. $a^\perp \perp a$ und $\|a\| = \|a^\perp\|$. Die Fläche des von a und $a^\perp$ aufgespannten Quadrates ist gegeben durch $\|a\|^2 = \|a\| \cdot \|a^\perp\|$.

Für einen beliebigen Vektor $b \in \mathbb{R}^2$ ergibt sich die Fläche als das Produkt aus der Länge von a mit der Länge von dem Anteil von b, der senkrecht auf a steht:

$$\text{Fläche } (a, b) = \|a\| \cdot \left| \langle b, \frac{a^\perp}{\|a^\perp\|} \rangle \right| = |\langle b, a^\perp \rangle| = \left| \left\langle \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \begin{pmatrix} -a_2 \\ a_1 \end{pmatrix} \right\rangle \right| = |a_1 b_2 - a_2 b_1|$$



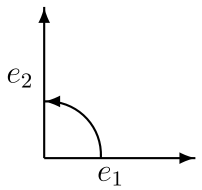
Das führt uns auf die folgende

DEFINITION 4.1 *Für $a, b \in \mathbb{R}^2$ heißt*

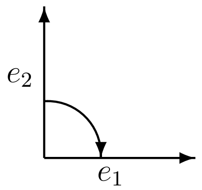
$$\det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = \det(a, b) := a_1 b_2 - a_2 b_1 \tag{4.1}$$

***Determinante** von a und b bzw. Determinante der Matrix A mit den Spaltenvektoren a und b.*

Mit dieser Definition gilt:



$$\det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 = \det(e_1, e_2)$$



$$\det \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = -1 = \det(e_2, e_1)$$

Die Determinante enthält also Informationen über die aufgespannte Fläche und die Orientierung der Vektoren.

NOTIZ 4.2 *Wie leicht zu sehen ist, hat die oben definierte Abbildung* $\det : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ *die Eigenschaften:*

D1) $\det$ *ist linear in jedem Eingang, d.h.* $\det(\lambda a, b) = \lambda \det(a, b) = \det(a, \lambda b)$ *und* $\det(a + b, c) = \det(a, c) + \det(b, c)$ *(Scherungsinvarianz)* .

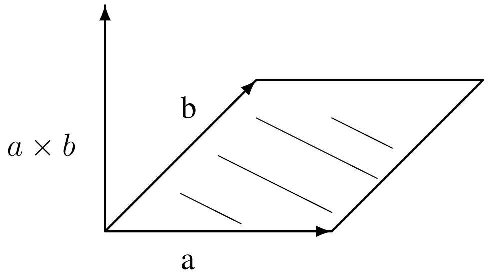
D2) *det ist alternierend, d.h.* $\det(a, b) = -\det(b, a)$.

D3) $\det$ *ist normiert:* $\det \mathbf{1} = 1$.

Wir werden später zeigen, dass die Eigenschaften $D1, D2, D3$ *die Abbildung* $\det$ *sogar schon eindeutig festlegen.*

4.1.2 Das Kreuzprodukt in $\mathbb{R}^3$

Wenden wir uns im Folgenden dem Kreuzprodukt im $\mathbb{R}^3$ als sehr wichtigem Hilfsmittel zu.



Zu zwei Vektoren $a, b \in \mathbb{R}^3$ suchen wir einen Vektor $a \times b \in \mathbb{R}^3$ mit den Eigenschaften:

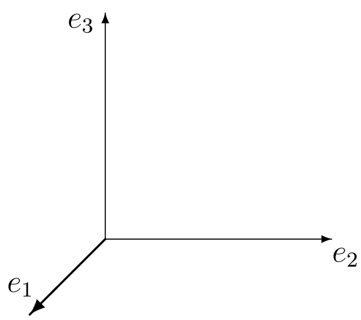
K1 $\|a \times b\|$ entspricht dem Betrag der von a und b aufgespannten Fläche

K2 $a \times b \perp \operatorname{Span}(a, b)$

K3 Die drei Vektoren $(a, b, a \times b)$ bilden ein Rechtssystem.

Dass drei Vektoren (a, b, c) ein Rechtssystem beschreiben ist physikalisch mittels unserer Hände definiert: ihre Richtung lässt sich durch Daumen $(= a)$, Zeigefinger$(= b)$ und Mittelfinger$(= c)$ der rechten Hand (in dieser Reihenfolge) darstellen.

Wegen K1 ist die Abbildung $\times : \mathbb{R}^3 \times \mathbb{R}^3 \mapsto \mathbb{R}^3$, $(a, b) \mapsto a \times b$, linear in jedem Eingang. Damit ist sie eindeutig bestimmt durch Werte auf einer Basis. Benutzen wir als die Standardbasis:



$$\left.\begin{array}{l} e_1 \times e_2 = e_3 = -e_2 \times e_1 \\ e_2 \times e_3 = e_1 = -e_3 \times e_2 \\ e_3 \times e_1 = e_2 = -e_1 \times e_3 \end{array}\right\} *$$

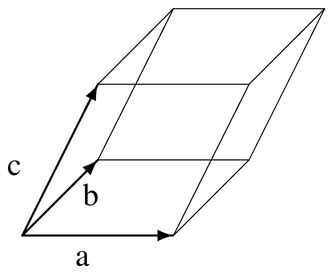
Mit $*$ folgt aus der Linearität von $\times$ für $a, b \in \mathbb{R}^3$

$$\begin{aligned} a \times b &= (a_2 b_3 - a_3 b_2)\, e_1 + (a_3 b_1 - a_1 b_3)\, e_2 + (a_1 b_2 - a_2 b_1)\, e_3 \\ &= \det \begin{pmatrix} a_2 & b_2 \\ a_3 & b_3 \end{pmatrix} e_1 + \det \begin{pmatrix} a_3 & b_3 \\ a_1 & b_1 \end{pmatrix} e_2 + \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} e_3 \,. \end{aligned} \tag{4.2}$$

SATZ 4.3 *Das durch* (4.2) *definierte Kreuzprodukt* $a \times b \in \mathbb{R}^3$ *hat alle gewünschten Eigenschaften K1, K2 und K3.*

4.1.3 Determinante in $\mathbb{R}^3$

Wenden wir uns mit dem uns nun zur Verfügung stehenden Kreuzprodukt der geometrischen Veranschaulichung der Determinante im $\mathbb{R}^3$ zu.



Parallelepiped, welches durch $a, b, c \in \mathbb{R}^3$ aufgespannt wird

$$\begin{aligned} &|Volumen(a,b,c)| \\ &= |Fläche(a,b)| \times \ Höhe \\ &= |a \times b| \cdot \left| \left\langle \tfrac{a \times b}{|a \times b|}, c \right\rangle \right| = |\langle a \times b, c\rangle| \end{aligned}$$

DEFINITION 4.4 *Für $a, b, c \in \mathbb{R}^3$ heißt*

$$\det(a,b,c) := \langle a \times b, c \rangle$$

Determinante *oder* ***Spatprodukt*** *der Vektoren a, b und c bzw. Determinante der Matrix A mit den Spalten a, b und c.*

NOTIZ 4.5

1. *Die Determinante in $\mathbb{R}^3$ hat auch die Eigenschaften D1, D2 und D3 der Determinante in $\mathbb{R}^2$. wobei D2 hier bedeutet, dass die Determinante bei Vertauschung zweier Vektoren (bzw. Spalten der Matrix) das Vorzeichen ändert (z.B. $\det(a,b,c) = -\det(b,a,c)$).*

2. *Die Determinante dreier Vektoren $a, b, c \in \mathbb{R}^3$ beschreibt das orientierte Volumen des von den Vektoren aufgespannten Parallelepipeds bzw. Spates.*

3. *Es gilt $\det(a,b,c) > 0$, wenn die Vektoren (a,b,c) ein Rechtssystem bilden, bilden sie ein Linkssystem (werden also durch Daumen, Zeigefinger und Mittelfinger der linken Hand dargestellt), so gilt $\det(a,b,c) < 0$.*

4. *Zyklische Vertauschung der Vektoren ändert den Wert der Determinante nicht:*

$$\det(a,b,c) = \langle a \times b, c\rangle = \langle b \times c, a\rangle = \langle c \times a, b\rangle$$

5. *Unter Verwendung der Determinante in $\mathbb{R}^2$ und (4.2) ergibt sich die Formel:*

$$\begin{aligned} \det(a,b,c) &= \sum c_i (a \times b)_i \\ &= c_1 \det \begin{pmatrix} a_2 & b_2 \\ a_3 & b_3 \end{pmatrix} + c_2 \det \begin{pmatrix} a_3 & b_3 \\ a_1 & b_1 \end{pmatrix} + c_3 \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \end{aligned}$$

Dieser Zusammenhang rechtfertigt die Merkregel

$$a \times b = \det \begin{pmatrix} a_1 & b_1 & e_1 \\ a_2 & b_2 & e_2 \\ a_3 & b_3 & e_3 \end{pmatrix},$$

wobei die rechte Seite nur symbolisch zu lesen ist, da $e_i \in \mathbb{R}^3$, $i = 1, 2, 3$, die Einheitsvektoren in $\mathbb{R}^3$ sind, hier als Eingänge in der Matrix also statt Zahlen Vektoren stehen.

Im folgenden soll die Determinante auch für n Vektoren aus $\mathbb{K}^n$ bzw. für quadratische Matrizen über $\mathbb{K}$ definiert werden.

Ziel: Suche für einen Körper $\mathbb{K}$ eine Abbildung $\det : \mathbb{K}^n \times ... \times \mathbb{K}^n \mapsto \mathbb{K}$ mit den Eigenschaften

D1) $\det : K^n \times ... \times K^n \mapsto K$ ist multilinear.

D2) $\det$ ist alternierend.

D3) $\det$ ist normiert: $\det \mathbf{1} = \det(e_1, ..., e_n) = 1$.

Dazu müssen wir als Hilfsmittel die Vorzeichen gut abzählen können.

4.2 Permutation und Signum

DEFINITION 4.6 *1. Die **symmetrische Gruppe** auf der Menge $X = \{1, ..., n\}$ ist gegeben durch*

$$\mathcal{S}_n := \{\pi : \{1, ..., n\} \to \{1, ..., n\} \,|\, \pi \text{ ist bijektiv } \}\,.$$

*2. Seien $(G, \cdot)$ und $(\tilde{G}, \tilde{\cdot})$ Gruppen. Eine Abbildung $\phi : G \mapsto \tilde{G}$ heißt **Gruppenhomomorphismus** :⇔*

$$\phi(g_1 \cdot g_2) = \phi(g_1) \,\tilde{\cdot}\, \phi(g_2)\,, \qquad g_1, g_2 \in G \tag{4.3}$$

NOTIZ 4.7 *1. Ist e das neutrale Element in G, dann gilt $\phi(e) \,\tilde{\cdot}\, \phi(g) = \phi(e \cdot g) = \phi(g)$, also ist $\phi(e) =: \tilde{e}$ das neutrale Element in $\tilde{G}$.*

2. Ein Gruppenhomomorphismus ist ein bijektiver Homomorphismus

DEFINITION 4.8 *Sei $(G, \circ)$ eine Gruppe und X eine Menge. Eine **Operation von** G **auf** X **von links** ist eine Abbildung*

$$\Phi : G \times X \to X\,, \quad (g, x) \mapsto \Phi(g, x) =: \Phi_g(x) =: g \cdot x$$

mit den Eigenschaften

1) $e \cdot x = x$ für alle $x \in X$, wobei e das neutrale Element von G ist.

2) $(a \circ b) \cdot x = a \cdot (b \cdot x)$ für alle $x \in X$ und $a, b \in G$.

SATZ 4.9 *Auf der Menge $X := \{f : \mathbb{R}^n \to \mathbb{R}\}$ operiert $\mathcal{S}_n$ von links durch*

$$(\pi \cdot f)(x_1, ..., x_n) := f(x_{\pi(1)}, ..., x_{\pi(n)}), \qquad \pi \in \mathcal{S}_n, f \in X, x = (x_1, \dots x_n) \in \mathbb{R}^n \ .$$

Beweis:. 1) klar
2) Für $\pi, \sigma \in \mathcal{S}_n$ gilt

$$\begin{aligned}(\pi \cdot (\sigma \cdot f))(x_1, ..., x_n) &= (\sigma \cdot f)(\underbrace{x_{\pi_1}}_{y_1}, ..., \underbrace{x_{\pi_n}}_{y_n}) = f(y_{\sigma_1}, ..., y_{\sigma_n}) \\ &= f(x_{\pi(\sigma(1))}, ..., x_{\pi(\sigma(n))}) = ((\pi \circ \sigma) \cdot f)(x_1, ..., x_n)\end{aligned}$$

□

LEMMA 4.10 *Für die Funktion $\Delta : \mathbb{R}^n \to \mathbb{R}$, die definiert ist durch $\Delta(x_1, ..., x_n) = \prod_{i<j}(x_j - x_i)$, gilt $\tau \cdot \Delta = -\Delta$ für jede Transposition $\tau \in \mathcal{S}_n$.*

Beweis:. Sei τ die Transposition, welche r und s vertauscht (sei oBdA $s < r$):

$$\tau(j) := \begin{cases} j & j \neq r, s \\ s & j = r \\ r & j = s \end{cases} \qquad \text{also} \quad \tau = \begin{pmatrix} 1 & \cdots & s & \cdots & r & \cdots & n \\ 1 & \cdots & r & \cdots & s & \cdots & n \end{pmatrix} .$$

Dann gilt unter Aktion von τ:

$$(x_r - x_s) \overset{\tau}{\mapsto} (x_s - x_r) = -(x_r - x_s)$$

Alle anderen Linearfaktoren sind entweder invariant unter τ, oder teilen sich in invariante Paare auf:

$$\text{für} \begin{cases} j > r & \text{ist} \quad (x_j - x_s)(x_j - x_r) \\ s < j < r & \text{ist} \quad (x_s - x_j)(x_j - x_r) \\ j < s & \text{ist} \quad (x_r - x_j)(x_s - x_j) \end{cases} \quad \text{invariant unter } \tau \ .$$

□

LEMMA 4.11 *Jede Permutation in $\mathcal{S}_n$ mit $n \geq 2$ ist eine Komposition von Transpositionen:*

$$\forall \pi \in \mathcal{S}_n, n \geq 2 \ \exists k \in \mathbb{N} \ \exists \tau_1, \dots \tau_k \ : \quad \pi = \tau_1 \circ ... \circ \tau_k \ . \tag{4.4}$$

Die Zerlegung und die Anzahl k sind nicht eindeutig bestimmt.

Beweis:. Induktion nach $n = 2, 3, ...$

I.V $n = 2 :$ $\mathbf{1}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = \tau$, klar! ($\mathbf{1} = \tau^2$)

I.S. $n \Rightarrow (n+1)$

Jede Permutation der Form $\pi = \begin{pmatrix} \cdots & n+1 \\ * & n+1 \end{pmatrix}$ (die also $n+1$ invariant lässt) kann als Element von $\mathcal{S}_n$ aufgefasst werden.

Sei $\sigma \in \mathcal{S}_{n+1}$ und τ die Transposition $(n+1) \longleftrightarrow \sigma^{-1}(n+1)$. Dann gilt

$$(n+1) \stackrel{\tau}{\longmapsto} \sigma^{-1}(n+1) \stackrel{\sigma}{\longmapsto} (n+1)$$

$$\text{und damit} \quad \sigma \circ \tau = \begin{pmatrix} \cdots & n+1 \\ * & n+1 \end{pmatrix} \in \mathcal{S}_n$$

$$\Rightarrow \sigma \circ \tau = \tau_1 \circ ... \circ \tau_{k-1} \Rightarrow \sigma = \tau_1 \circ ... \circ \tau_{k-1} \circ \tau^{-1}$$

□

KOROLLAR 4.12 *Ist Δ gegeben wie in Lemma 4.10, dann gilt für jede Permutation $\pi = \tau_1 \circ ... \circ \tau_k \in \mathcal{S}_n$:*

$$\pi \cdot \Delta = (\tau_1 \circ ... \circ \tau_k) \cdot \Delta = (-1)^k \Delta \ , \tag{4.5}$$

*Insbesondere ist das **Signum** einer Permutation $\pi \in \mathcal{S}_n$*

$$\mathrm{sign}\ \pi = \mathrm{sign}(\tau_1 \circ ... \circ \tau_k) := (-1)^k$$

unabhängig von Wahl der Zerlegung von π in Transpositionen und damit wohldefiniert.

Beweis:. Gleichung (4.5) folgt sofort aus Lemma 4.10. Falls $\pi = \tilde{\tau}_1 \circ ... \circ \tilde{\tau}_\ell$ mit den Transpositionen $\tilde{\tau}_j\ ,\ j = 1, \ldots \ell$, so gilt nach (4.5) $(-1)^k \Delta = \Pi \cdot \Delta = (-1)^\ell \Delta$. Die Auswertung auf $(x_1, ..., x_n) \in \mathbb{R}^n$ mit $\Delta(x_1, ..., x_n) \neq 0$ liefert $(-1)^k = (-1)^\ell$ (damit auch $k = \ell$ mod 2).

□

SATZ 4.13 *Die Abbildung* $\mathrm{sign} : \mathcal{S}_n \to \{-1, +1\}$ *ist ein Gruppenhomomorphismus von $(\mathcal{S}_n, \circ)$ in die multiplikative Gruppe $(\{\pm 1\}, \cdot) \cong (\mathbb{Z}_2, \cdot)$.*

Beweis:. Seien $\pi = \tau_1 \circ ... \circ \tau_k,\ \sigma = \tilde{\tau}_1 \circ ... \circ \tilde{\tau}_r\ \in \mathcal{S}_n$, dann gilt

$$\begin{aligned} \mathrm{sign}\ (\pi \circ \tau) &= \mathrm{sign}\ (\tau_1 \circ ... \circ \tau_k \circ \tilde{\tau}_1 \circ ... \circ \tilde{\tau}_r) \\ &= (-1)^{r+k} = (-1)^r \cdot (-1)^k = \mathrm{sign}\ \pi \cdot \mathrm{sign}\ \sigma \end{aligned}$$

□

4.3 Exkurs: Ein wenig Gruppen-ABC

DEFINITION 4.14 *Sei $(G, \cdot)$ eine Gruppe, dann heißt $H \subset G$ **Untergruppe** von G :⇔*

1. *$H \neq \emptyset$.*

2. *$a \cdot b \in H$ für alle $a, b \in H$.*

3. *$a^{-1} \in H$ für alle $a \in H$.*

NOTIZ 4.15 *$(H, \cdot)$ mit der von G vererbten Verknüpfung ist dann selbst eine Gruppe.*

DEFINITION 4.16 1. *Eine Untergruppe H von G heißt **normal** bzw. **Normalteiler** von G :⇔*

$$ghg^{-1} \in H\,, \quad \textit{wir schreiben kurz } H \lhd G\,, \qquad g \in G\,,\ h \in H \tag{4.6}$$

2. *Das **Zentrum** einer Gruppe G ist die Menge $\mathcal{Z}(G) := \{h \in G \,|\, g \cdot h = h \cdot g \; \forall g \in G\}$.*

Bemerkung:

1) Das Zentrum einer Gruppe ist normal, d.h. $g \cdot h \cdot g^{-1} = h \in \mathcal{Z}(G)$ für alle $h \in \mathcal{Z}(G), g \in G$.

2) Auf einem Normalteiler $H \lhd G$ operiert G durch Konjugation $\alpha_g :\ h \mapsto g \cdot h \cdot g^{-1} \in H$

SATZ 4.17 *Sei $\phi : G \longmapsto \tilde{G}$ ein Gruppenhomomorphismus, dann ist* $\mathrm{Bild}\ \phi = \phi(G)$ *eine Untergruppe von $\tilde{G}$ und* $\mathrm{Kern}\ \phi = \{g \in G \,|\, \phi(g) = \tilde{e} \in \tilde{G}\}$ *ein Normalteiler von G.*

Beweis. 1) z.z.: $\mathrm{Bild}\ \phi$ ist Untergruppe von $\tilde{G}$:

Seien $a, b \in G$ und $\tilde{a} = \phi(a),\ \tilde{b} = \phi(b) \in \mathrm{Bild}\ \phi$, dann folgt aus der Homomorphismus-Eigenschaft $\tilde{a} \,\tilde{\cdot}\, \tilde{b} = \phi(a) \,\tilde{\cdot}\, \phi(b) = \phi(a \cdot b) \in \mathrm{Bild}\ \phi$.

Sei $a \in G$ mit inversem Element $a^{-1} \in G$, dann gilt $\phi(a) \,\tilde{\cdot}\, \phi(a^{-1}) = \phi(a \cdot a^{-1}) = \phi(e) = \tilde{e}$ und damit $\phi(a^{-1}) = (\phi(a))^{-1} \in \mathrm{Bild}\ \phi$.

2) z.z.: $\mathrm{Kern}\ \phi$ ist Normalteiler von G:

- z.z.: $\mathrm{Kern}\ \phi$ ist Untergruppe von G:

  Seien $a, b \in \mathrm{Kern}\ \phi$, also $\phi(a) = \phi(b) = \tilde{e}$. dann folgt $\tilde{e} = \phi(a) \,\tilde{\cdot}\, \phi(b) = \phi(a \cdot b)$ und damit ist $a \cdot b \in \mathrm{Kern}\ \phi$.

  Sei $a \in \mathrm{Kern}\ \phi$, also $\phi(a) = \tilde{e}$, dann gilt $\tilde{e} = \phi(a) \,\tilde{\cdot}\, \phi(a^{-1}) = \tilde{e} \,\tilde{\cdot}\, \phi(a^{-1}) = \phi(a^{-1})$ und damit $a^{-1} \in \mathrm{Kern}\ \phi$.

- z.z.: $\mathrm{Kern}\ \phi$ ist Normalteiler von G:

  Für $h \in \mathrm{Kern}\ \phi$ und $g \in G$ gilt $\phi(g \cdot h \cdot g^{-1}) = \phi(g) \,\tilde{\cdot}\, \phi(h) \,\tilde{\cdot}\, \phi(g^{-1}) = \phi(g) \,\tilde{\cdot}\, \tilde{e} \,\tilde{\cdot}\, \phi(g^{-1}) = \phi(g) \,\tilde{\cdot}\, \phi(g^{-1}) = \phi(g \cdot g^{-1}) = \phi(e) = \tilde{e}$ und damit $g \cdot h \cdot g^{-1} \in \mathrm{Kern}\ \phi$

□

DEFINITION 4.18 *Sei $(G, \cdot)$ eine Gruppe und $H \subset G$ eine Untergruppe von G.*

1. *Die **Linksnebenklassen** von H in G sind die Mengen*

$$aH := \{g = a \cdot h \,|\, h \in H\}\,, \qquad a \in G\,,$$

*also die **Äquivalenzklassen** $[a]$ unter der Äquivalenzrelation $g \sim a \;:\Leftrightarrow\; \exists h \in H :\ g = a \cdot h$.*

2. *Die **Rechtsnebenklassen** von H in G sind die Mengen*

$$Ha = \{g = h \cdot a \,|\, h \in H\}\,, \qquad a \in G\,, \tag{4.7}$$

also die Äquivalenzklassen unter der Äquivalenzrelation $g \sim a \;:\Leftrightarrow\; \exists h \in H:\; g = h \cdot a$.

NOTIZ 4.19 *Ist H normal, dann sind die Linksnebenklassen gleich den Rechtsnebenklassen:*

$$\begin{aligned} H \lhd G \quad &\Rightarrow \quad \textit{Für } g \in aH \textit{ gilt:} \quad g = ah = \underbrace{aha^{-1}}_{\in H} a = \tilde{h}a \in Ha \\ &\Rightarrow \quad aH = Ha \quad \textit{(Linksnebenklasse = Rechtsnebenklasse)} \end{aligned}$$

SATZ 4.20 *Sei $(G, \cdot)$ eine Gruppe und $H \lhd G$, dann induziert die Gruppenstruktur von G eine Gruppenstruktur auf dem Quotientenraum $G/H := \{aH \,|\, a \in G\}$ vermöge*

$$aH \,\hat{\cdot}\, bH = (a \cdot b)H\;. \tag{4.8}$$

*Die **Restklassenabbildung** $\Pi : \;\; G \mapsto G/H,\; a \mapsto aH$ ist ein surjektiver Gruppenhomomorphismus, die sogenannte **kanonische Projektion**, mit $\mathrm{Kern}\ \Pi = H$. Die Gruppe $(G/H, \hat{\cdot})$ heißt **Quotientengruppe von G nach dem Normalteiler H**.*

Beweis. Zunächst ist zu zeigen, dass die in (4.8) definierte Verknüpfung $\hat{\cdot}$ wohldefiniert, also repräsentantenunabhängig, ist. Seien also $a' = a \cdot h_1$ und $b' = b \cdot h_2$ andere Repräsentanten von aH und bH, dann gilt

$$a'H \,\hat{\cdot}\, b'H = (a' \cdot b')H = (a \cdot h_1 \cdot b \cdot h_2)H = \big(a \cdot b \cdot \underbrace{b^{-1} \cdot h_1 \cdot b}_{\in H} \cdot h_2\big)H = (a \cdot b)H = aH \hat{\cdot} bH\;.$$

$(G/H, \hat{\cdot})$ ist eine Gruppe, denn $\hat{\cdot}$ ist abgeschlossen nach Definition und

- Das neutrale Element ist eH, denn $aH \,\hat{\cdot}\, eH = (a \cdot e)H = aH$.
- Das inverse Element zu aH ist $a^{-1}H$, denn $aH \,\hat{\cdot}\, a^{-1}H = (a \cdot a^{-1})H = eH$.
- $\hat{\cdot}$ ist assoziativ, denn $(aH \,\hat{\cdot}\, bH) \,\hat{\cdot} cH = (a \cdot b)H \,\hat{\cdot}\, cH = ((a \cdot b) \cdot c)H = (a \cdot (b \cdot c))H = aH \,\hat{\cdot}\, (b \cdot c)H = aH \,\hat{\cdot}\, (bH \,\hat{\cdot} cH)$.

Π ist ein Gruppenhomomorphismus, denn $\Pi(a \cdot b) = (a \cdot b)H = aH \,\hat{\cdot}\, bH = \Pi(a) \,\hat{\cdot}\, \Pi(b)$. Außerdem ist Π surjektiv, denn $G/H \ni aH = \Pi(a) \in \mathrm{Bild}\ \Pi$ und es gilt $\mathrm{Kern}\ \Pi = \{a \in G \,|\, \Pi(a) = eH\} = H$, denn $\Pi(a) = eH \Leftrightarrow aH = (e \cdot a)H = eH \Leftrightarrow a \in H$. $\square$

SATZ 4.21 *Für jeden Gruppenhomomorphismus $\phi : \;\; (G, \cdot) \mapsto (\tilde{G}, \tilde{\cdot})$ ist $\mathrm{Bild}\ \phi$ isomorph zu $G/(\mathrm{Kern}\ \phi)$, das heisst, die Abbildung $j : G/(\mathrm{Kern}\ \phi) \to \mathrm{Bild}\ \phi$, die gegeben ist durch $a\,\mathrm{Kern}\ \phi \mapsto \phi(a)$, ist ein Isomorphismus (d.h. ein bijektiver Homomorphismus).*

$$
\begin{tikzcd}
G \arrow[rr, "\phi"] \arrow[dr, "\Pi"'] & & \text{Bild } \phi \subset \tilde{G} \\
& G/(\text{Kern } \phi) \arrow[ur, "j"'] &
\end{tikzcd}
$$

Beweis:. Wir setzen $H := \text{Kern } \phi = \{g \in G \mid \phi(g) = \tilde{e}\}$.
j ist wohldefiniert, denn für 2 Repräsentanten von $[a]$ gilt:

$$a, b \in aH \quad \Leftrightarrow \quad \exists h \in H \, : \, b = a \cdot ht \quad \Leftrightarrow \quad a^{-1} \cdot b = h \in H \, . \tag{4.9}$$

Daraus folgt

$$j(bH) = \phi(b) = \phi(a \cdot a^{-1} \cdot b) = \phi(a) \,\tilde{\cdot}\, \phi(\underbrace{a^{-1}b}_{\in \text{Kern } \phi}) = \phi(a) \,\tilde{\cdot}\, \tilde{e} = \phi(a) = j(aH) \, .$$

j ist homomorph, denn $j(aH \,\hat{\cdot}\, bH) = j((a \cdot b)H) = \phi(a \cdot b) = \phi(a) \,\tilde{\cdot}\, \phi(b) = j(aH) \,\tilde{\cdot}\, j(bH)$.
j ist injektiv, denn mit (4.9) folgt

$$\begin{aligned} j(aH) = j(bH) \;&\Leftrightarrow\; \phi(a) = \phi(b) \;\Leftrightarrow\; \phi(a^{-1} \cdot b) = \phi(a^{-1}) \,\tilde{\cdot}\, \phi(b) = \phi(a^{-1}) \,\tilde{\cdot}\, \phi(a) = \tilde{e} \\ &\Leftrightarrow\; a^{-1}b \in H \;\Leftrightarrow\; bH = aH \, . \end{aligned}$$

Außerdem ist j surjektiv, denn falls $g \in \text{Bild } \phi$, dann existiert $a \in G$ mit $g = \phi(a)$. Damit folgt aber sofort $g = j(aH)$.

□

Beispiel: Nach Satz 4.13 ist $\text{sign} : \mathcal{S}_n \to \{-1, +1\}$ ein Gruppenhomomorphismus. Dann ist $A_n := \text{Kern sign} = \{\sigma \in \mathcal{S}_n \mid \text{sign } \sigma = 1\} \lhd \mathcal{S}_n$ die sogenannte **alternierende Gruppe**. Für jede Transposition $\tau \in \mathcal{S}_n$ gilt $\mathcal{S}_n = A_n \cup \tau A_n$ und $A_n \cap \tau A_n = \emptyset$, also ist $\mathcal{S}_n$ eine disjunkte Vereinigung von A_n und τA_n.

4.4 Determinanten im $\mathbb{K}^n$

Sei nun $\mathbb{K}$ ein beliebiger Körper. Um die Determinante analog zum $\mathbb{R}^2$ und $\mathbb{R}^3$ einzuführen, definieren wir zunächst:

DEFINITION 4.22 *Sei V ein Vektorraum über $\mathbb{K}$ mit $\dim V = n$.*

1. *Eine Funktion $f : \underbrace{V \times \ldots \times V}_{p \text{ mal}} \longrightarrow \mathbb{K}$ heisst **p-Multilinearform** (oder einfach **p-Form**) auf V :$\Leftrightarrow$ für alle $j = 1, \ldots p$ ist die Abbildung $f(a_1, \ldots, a_{j-1}, \cdot, a_{j+1}, \ldots a_p) : V \to \mathbb{K}$ linear.*

2. *f heisst **alternierend** :⇔ für alle $a \in V$ gilt $f(\ldots, a, \ldots, a, \ldots) = 0$.*

3. *Die Menge der alternierenden p-Multilinearformen auf V bezeichnen wir mit $\operatorname{Alt}^p V$.*

NOTIZ 4.23 $\operatorname{Alt}^p V$ *ist ein Vektorraum über* $\mathbb{K}$.

Die Definition einer alternierenden Multilinearform ist etwas anders als wir es bei den Eigenschaften der Determinante in $\mathbb{R}^2$ und $\mathbb{R}^3$ verwendet haben, es gilt aber

LEMMA 4.24 *Für jedes $f \in \operatorname{Alt}^p \mathbb{K}^n$ gilt: (1)⇔(2)⇒(3), wobei*

(1) f ist alternierend.

(2) $f(a_1, ..., a_p) = 0$ falls $(a_1, ..., a_p)$ linear abhängig sind.

(3) $f(..., a_i, ..., a_j, ...) = -f(..., a_j, ..., a_i, ...)$ für alle $i, j = 1, \ldots p$.

Falls $2 \neq 0$ in $\mathbb{K}$, gilt sogar (2)⇔(3).

Beweis:. (1)⇐ (2) ist klar, da $(\ldots, a, \ldots, a, \ldots)$ linear abhängig ist.
(1)⇒ (2): Falls das Tupel $(a_1, ..., a_p)$ linear abhängig ist, dann ist ein a_i eine Linearkombination der anderen, es gibt also $\lambda_{ij} \in \mathbb{K}, j \neq i$, so dass $a_i = \sum_{i \neq j} \lambda_{ij} a_j$. Aus der Multilinearität von f und (1) folgt $f(a_1, ..., a_p) = \sum_{i \neq j} \lambda_{ij} f(a_1, ..., a_{i-1}, a_j, a_{i+1}, ...) = 0$
(2)⇒ (3): Die Multilinearität von f liefert $f(..., a_i, ..., a_j, ...) + f(..., a_j, ..., a_i, ...) = f(..., a_i + a_j, ..., a_i + a_j, ...) = 0$ und damit (3).
(3)⇒ (1) falls $2 \neq 0$ in $\mathbb{K}$: Aus (2) folgt sofort $f(..., a, ..., a, ...) = -f(..., a, ..., a, ...)$ und damit $2f(..., a, ..., a, ...) = 0$. Aus $2 \neq 0$ folgt damit (1).

□

Um die Determinante in $\mathbb{K}^n$ zu defnieren, verwenden wir die Eigenschaften D1, D2, D3 der Determinante in $\mathbb{R}^2$ und $\mathbb{R}^3$:

DEFINITION 4.25 *Sei* $\det : \underbrace{\mathbb{K}^n \times ... \times \mathbb{K}^n}_{n \text{ mal}} \longrightarrow \mathbb{K}$ *eine Abbildung mit den Eigenschaften*

D1) $\det$ *ist eine n-Multilinearform,*

D2) $\det$ *ist alternierend,*

D3) $\det$ *ist normiert, also* $\det \mathbf{1} = \det(e_1, ..., e_n) = 1$,

dann heisst $\det(a_1, \ldots a_n)$ *Determinante von* $(a_1, \ldots a_n)$ *bzw. Determinante der Matrix* $A \in \mathcal{M}(n \times n, \mathbb{K})$, *deren Spalten durch die Vektoren* $a_1, \ldots a_n$ *gegeben ist.*

Die Determinante eines Tupels oder einer Matrix ist eindeutig bestimmt, denn es gilt

SATZ 4.26 *Es gibt genau eine Abbildung* $\det : \underbrace{\mathbb{K}^n \times ... \times \mathbb{K}^n}_{n\ mal} \longrightarrow \mathbb{K}$ *mit den Eigenschaften D1-D3. Für die Zeilenvektoren* $a_j = (a_{j1}, ..., a_{jn})$ *der Matrix* $A = (a_{ij}) \in \mathcal{M}(n \times n; \mathbb{K})$ *gilt*

$$\det(a_1, ..., a_n) = \det\ A = \sum_{\pi \in \mathcal{S}_n} (\text{sign}\ \pi)\ a_{1\pi(1)}, ..., a_{n\pi(n)}\ . \tag{4.10}$$

Beweis:. Wir zeigen zunächst, dass jede Abbildung $\det$, die Definition 4.25 erfüllt, durch die rechte Seite von (4.10) gegeben ist. Das impliziert dann die Eindeutigkeit dieser Abbildung.
Sei $\det$ eine Determinante nach Definition 4.25, dann gilt

$$\begin{aligned}
\det(a_1, ..., a_n) &= \det\left(\sum_{j_1=1}^{n} a_{1_{j_1}} e_{j_1}, ...\right) \\
&\overset{(D1)}{=} \sum_{j_1,...,j_n=0}^{n} a_{1_{j_1}}, ..., a_{n_{j_n}} \qquad \underbrace{\det(e_{j_1}, ..., e_{j_n})}_{=0, \text{ falls } (j_1,...,j_n) \text{ \underline{keine} Permutation } (\pi(1),...,\pi(n)) \text{ ist.}} \\
&\overset{(D2)}{=} \sum_{\pi \in \mathcal{S}_n} a_{1\pi(1)}, ..., a_{n\pi(n)} \qquad \underbrace{\det(e_{\pi(1)}, ..., e_{\pi(n)})}_{\pi \text{ enthält } k \text{ Transpositionen, also } (-1)^k = \text{sign } \pi} \\
&\overset{(D3)}{=} \sum_{\pi \in \mathcal{S}_n} a_{1\pi(1)}, ..., a_{n\pi(n)}\ \text{sign}\ \pi\ \underbrace{\det \mathbf{1}}_{=1}\ .
\end{aligned}$$

Also ist det eindeutig durch D1-D3 bestimmt und aus diesen Eigenschaften folgt (4.10).

Um die Existenz einer Abbildung $\det$ aus Definition 4.25 zu beweisen, zeigen wir, dass (4.10) eine solche Abbildung definiert.
Sei $f : \mathbb{K}^n \times \ldots \times \mathbb{K}^n \to \mathbb{K}$ durch die rechte Seite von (4.10) definiert. Dann gilt:
(D1) f ist multilinear, denn für alle $j = 1, \ldots n$ gilt

$$\begin{aligned}
f(a_1, ..., \underbrace{a+b}_{j.\ \text{Eingang}}, ..., a_n) &= \sum_{\pi \in \mathcal{S}_n} a_{1\pi(1)} ... (a+b)_{\pi(i)} ... a_{n\pi(n)}\ (\text{sign}\ \pi) \\
&= \sum_{\pi} (\text{sign}\ \pi)\ a_{1\pi(1)} ... a_{\pi(j)} ... a_{n\pi(n)} + \sum_{\pi} ... b_{\pi(j)} ... \\
&= f(...a...) + f(...b...)
\end{aligned}$$

Weiterhin folgt sofort aus der Formel $f(..., \lambda a, ...) = \lambda f(..., a, ...)$ für $\lambda \in \mathbb{K}$.
(D3) f ist normiert, denn es gilt

$$f(e_1, ..., e_n) = \sum_{\pi \in \int_n} (\text{sign}\ \pi)\ \underbrace{e_{1\pi(1)} ... e_{n\pi(n)}}_{=0 \text{ falls } \pi \neq \mathbf{1}} = \text{sign}\ \mathbf{1} = 1\ .$$

(D2) Um zu zeigen, dass f alternierend ist, verwenden wir das letzte Beispiel in Abschnitt 4.3, d.h. die disjunkte Vereinigung $\mathcal{S}_n = A_n \cup \tau A_n = A_n \cup A_n \tau$ für die Transposition τ, die i und j vertauscht. Dann gilt für $b \in \mathbb{K}^n$

$$
\begin{aligned}
f(a_1, ..., \underbrace{b}_{i.\text{ Eingang}}, ..., \underbrace{b}_{j.\text{ Eingang}}, ...a_n) &= \sum_{\pi \in A_n \cup \tau A_n} \text{sign } \pi \; a_{1\pi(1)}...b_{\pi(i)}...b_{\pi(j)}...a_{n\pi(n)} \\
&= \sum_{\pi \in A_n} \underbrace{(\text{sign } \pi)}_{=1} a_{1\pi(1)}...b_{\pi(i)}...b_{\pi(j)}...a_{n\pi(n)} + \sum_{\sigma \in A_n\tau} \underbrace{(\text{sign } \sigma)}_{=-1} a_{1\sigma(1)}...b_{\sigma(i)}...b_{\sigma(j)}...a_{n\sigma(n)} \\
&= \sum_{\pi \in A_n} a_{1\pi(1)}...b_{\pi(i)}...b_{\pi(j)}...a_{n\pi(n)} - \sum_{\sigma \in A_n\tau} a_{1\sigma(1)}...b_{\sigma(i)}...b_{\sigma(j)}...a_{n\sigma(n)} = 0 \;,
\end{aligned}
$$

wobei die letzte Gleichung gilt, da für jedes $\sigma \in A_n\tau$ ein $\pi \in A_n$ existiert, so dass $\sigma = \pi \circ \tau$ gilt und damit $a_{1\sigma(1)}...b_{\sigma(i)}...b_{\sigma(j)}...a_{n\sigma(n)} = a_{1\pi(1)}...b_{\pi(j)}...b_{\pi(i)}...a_{n\pi(n)}$.

□

NOTIZ 4.27 *Falls* $2 \neq 0$ *in* $\mathbb{K}$*, dann benötigt man die Zerlegung* $\mathcal{S}_n = A_n \cup A_n\tau$ *nicht.*

KOROLLAR 4.28 $\text{Alt}^n\,\mathbb{K}^n$ *ist ein* $\mathbb{K}$*-Vektorraum der Dimension 1. Insbesondere gilt*

$$\forall\, f \in \text{Alt}^n\,\mathbb{K}^n \;\exists \lambda \in \mathbb{K} : \quad f = \lambda \det \;.$$

Beweis:. Die Überprüfung der Vektorraumeigenschaften ist eine Übungsaufgabe.
Wir setzen $W := \text{Alt}^n\,\mathbb{K}^n$ und betrachten die Abbildung $\lambda : W \to \mathbb{K}$ mit $\lambda(f) := f(\mathbf{1})$ (hier bezeichnet $\mathbf{1} = (e_1, \ldots e_n)$ die Einheitsmatrix mit den Einheitsvektoren als Spalten). Offensichtlich ist $\lambda \in \text{Hom}(W, \mathbb{K})$ (selber!). Weiterhin gilt $\text{Kern } \lambda = \{0\}$, sonst gäbe es ein $f \in W$ mit $f \neq 0$ und $f(\mathbf{1}) = 0$. Dann wäre aber $(\det + f)(\mathbf{1}) = 1$ und damit $\det + f$ eine weitere Abbildung, verschieden von $\det$, mit den Eigenschaften D1, D2 und D3, im Widerspruch zu Satz 4.26. Also gilt $\dim \text{ Kern } \lambda = 0$ und da $\dim \text{ Bild } \lambda = \dim \mathbb{K} = 1$, folgt mit der Dimensionsformel $\dim W = 0 + 1 = 1$. Der Wert $\lambda(f) \in \mathbb{K}$ der oben definierten Abbildung genügt wegen D3 der Gleichung $f = \lambda(f) \det$.

□

Aus Definition 4.25 bzw. (4.10) folgt

KOROLLAR 4.29 *Für* $A \in \mathcal{M}(n \times n; \mathbb{K})$ *gilt*

$$\det \; A = \det \; A^t \,,$$

wobei A^t *die transponierte Matrix bezeichnet (Definition 3.29).*

Beweis:. Sei $A = (a_{ij})$ also $A^t = (b_{ij})$ mit $b_{ij} = a_{ji}$. Wegen $\text{sign } \pi = \text{sign } \pi^{-1}$ gilt

$$
\begin{aligned}
\det \; A^t &= \textstyle\sum_{\pi \in \mathcal{S}_n} (\text{sign } \pi) \underbrace{b_{1\pi(1)}...b_{n\pi(n)}}_{=a_{\pi(1)1}...a_{\pi(n)n} = a_{1\pi^{-1}(1)}...a_{n\pi^{-1}(n)}} \\
&= \textstyle\sum_{\sigma = \pi^{-1} \in \mathcal{S}_n} (\text{sign } \sigma) \; a_{1\sigma(1)}...a_{n\sigma(n)} = \det A
\end{aligned}
$$

□

SATZ 4.30 *Für* $A, B \in \mathcal{M}(n \times n; \mathbb{K})$ *gilt*

$$\det\ A \cdot B = \det\ A\ \det\ B\ . \tag{4.11}$$

Beweis:. Für $A = (a_{ik})$ und $B = (b_{kj}) = (b_1, \dots b_n)$ ist $A \cdot B = C = (c_{ij})$ mit $c_{ij} = \sum_k a_{ik}\ b_{kj}$. Für die jeweiligen Spaltenvektoren b_j und c_j gilt dann $c_j = Ab_j$ und wir haben

$$A \cdot B = (A \cdot b_1, \dots A \cdot b_n)\ . \tag{4.12}$$

Es gilt mit (4.12)

$$\det(A \cdot B) = \det(Ab_1, ..., Ab_n) = \underbrace{\det(A\cdot, ..., A\cdot)}_{=:f_A}[b_1, ..., b_n] = f_A(b_1, \dots b_n)$$

und $f_A \in \mathrm{Alt}^n\, \mathbb{K}^n$, denn für $b_j = b_k$ folgt $A \cdot b_j = A \cdot b_k$ und es gilt $A \cdot (a + b) = A \cdot a + A \cdot b$. Damit folgt die Behauptung aus den Eigenschaften D1 und D2 der Determinante.
Mit Korollar 4.28 gilt die Gleichung $f_A(b_1, \dots b_n) = \lambda \det(b_1, \dots, b_n)$ für $\lambda = f_A(\mathbf{1}) = \det(A \cdot \mathbf{1}) = \det A$ und damit folgt

$$\det(A \cdot B) = f_A(B) = \det A \det(B)\ .$$

□

SATZ 4.31 *Für* $A \in \mathcal{M}(n \times n; \mathbb{K})$ *gilt*

$$\det\ A \neq 0 \quad \Leftrightarrow \quad A \in \mathrm{GL}(n, \mathbb{K})\ . \tag{4.13}$$

Beweis:. Sei $A \in \mathrm{GL}(n, \mathbb{K})$, dann existiert nach Definition 3.20 ein $A^{-1} \in \mathcal{M}(n \times n; \mathbb{K})$ mit $A \cdot A^{-1} = \mathbf{1}$ und wegen Satz 4.30 und D3 gilt $1 = \det \mathbf{1} = \det(A \cdot A^{-1}) = \det A\ \det A^{-1}$ und damit insbesondere $\det A \neq 0 \neq \det A^{-1}$.

□

NOTIZ 4.32 *Wir können Satz 4.31 natürlich mit der Verneinung und Satz 3.27 äquivalent formulieren:*

$$\det\ A = 0 \quad \Leftrightarrow \quad \mathrm{Rang}\ A < n\ .$$

KOROLLAR 4.33 *Sei* V *ein* $\mathbb{K}$*-Vektorraum,* $T \in \mathrm{End}(V)$ *und seien* $\mathcal{A}, \mathcal{B}$ *Basen von* V*. Dann gilt:*

$$\det\ (\Phi^{\mathcal{A}}_{\mathcal{A}} T) = \det\ (\Phi^{\mathcal{B}}_{\mathcal{B}} T) \tag{4.14}$$

Das heißt, dass die Determinante der Matrixdarstellung eines Endomorphimus ist unabhängig von der Wahl einer Basis.

Beweis:. Mit Satz 4.30 gilt

$$\det\ (\Phi^{\mathcal{B}}_{\mathcal{B}}(T)) = \det \left(\underbrace{\Phi^{\mathcal{A}}_{\mathcal{B}}(id)}_{=:A}\ \Phi^{\mathcal{A}}_{\mathcal{A}}(T)\ \underbrace{\Phi^{\mathcal{B}}_{\mathcal{A}}(id)}_{=A^{-1}} \right) = \underbrace{(\det\ A^{-1})(\det\ A)}_{=\det(A^{-1} \cdot A) = 1} \det\ \Phi^{\mathcal{A}}_{\mathcal{A}}(T)$$

□

Das Korollar erlaubt uns die folgende Definition:

DEFINITION 4.34 *Sei V ein Vektorraum und $T \in \mathrm{End}(V)$ ein Endomorphimus auf V. Dann ist die Determinante von T definiert durch*

$$\det\, T := \det\, (\Phi_{\mathcal{A}}^{\mathcal{A}} T)\ ,$$

wobei , $\mathcal{A}$ eine beliebige Basis von V ist.

4.5 Berechnung und Anwendung von Determinanten

SATZ 4.35 $\det(\cdot)$ *ist invariant unter Zeilen- bzw. Spaltenumformungen vom Typ 3 (siehe Definition 3.38), das heisst die Addition des λ-fachen einer Zeile (Spalte) zu einer anderen Zeile (Spalte) ändert den Wert der Determinante nicht:*

$$\det(...a_i...a_j...) = \det(...a_i..., (a_j + \lambda a_i), ...)\ .$$

Beweis:. Aus D1 folgt $\det(...a_i..., (a_j+\lambda a_i), ...) = \det(\ldots a_i, \ldots a_j, \ldots) + \lambda \det(\ldots a_i, \ldots a_i, \ldots)$ und wegen D2 ist der zweite Summand null.

□

SATZ 4.36 *Für eine obere Dreiecksmatrix* $A = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix}$ *gilt*

$$\det\ A = \lambda_1 \cdot \lambda_2 \cdot ... \cdot \lambda_n.$$

Beweis:. Falls eines der $\lambda_i = 0$, dann folgt $\mathrm{Rang}\ \ A < n$ und mit Satz 4.31 gilt $\det\ \ A = 0$ und damit die Behauptung.
Seien nun $\lambda_1, ..., \lambda_n \neq 0$. Dann folgt aus D1

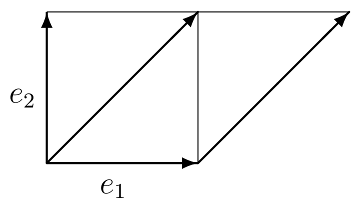
$$\det\ A = \lambda_1\ \det \begin{pmatrix} 1 & & * \\ \lambda_2 & & \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \lambda_1 \ldots \lambda_n\ \det \begin{pmatrix} 1 & & * \\ 1 & & \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \tag{4.15}$$

Die Matrix auf der rechten Seite von (4.15) kann durch elementare Zeilenumformungen in die Einheitsmatrix $\mathbf{1}$ umgeformt werden. Wegen Satz 4.35 folgt damit sofort

$$\det\ A = \lambda_1 \ldots \lambda_n\ \det\ \mathbf{1} = \lambda_1 \ldots \lambda_n\ .$$

□

NOTIZ 4.37 $\det \begin{pmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = \det \mathbf{1} = 1$ *ist die algebraische Formulierung der Scherungsinvarianz von Volumina.*



Satz 4.36 liefert uns das folgende

Berechnungsverfahren für die Determinante einer Matrix:
Zur Berechnung von $\det\ A,\ A \in \mathcal{M}(n \times n; \mathbb{K})$ verwenden wir das Gaußverfahren, um A auf eine obere Dreiecksmatrix zu transformieren.

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

Dann gilt:

$$\det\ A = (-1)^k \lambda_1 ... \lambda_n$$

Dieser Schritt ist für $n \geq 4$ sinnvoll.

Jetzt lösen wir ein Gleichungssystem $Ax = b \in \mathbb{K}^n,\ A \in \mathrm{GL}(n, \mathbb{K})$:
Falls $Ax = b$ eindeutig lösbar ist, gilt $\det\ A \neq 0$. Seien $(a_1, ..., a_n)$ die Spaltenvektoren von A. Wende nun die Linearform $\det(a_1, ..., a_{i-1}, \cdot, a_{i+1}, ..., a_n) :\ K^n \to K$ an:

$$\det(..., a_{i-1}, b, a_{i+1}, ...) = \det(...a_{i-1}, , Ax = \sum_{k=1}^{n} a_k x_k, a_{i+1}, ...) = x_i\ \det(a_1, ..., a_i, ..., a_n)\ .$$

Das beweist die **Cramer‘sche Regel**:

SATZ 4.38 (CRAMER'SCHE REGEL) *Sei* $A \in \mathrm{GL}(n, \mathbb{K})$ *und seien* a_j *die Spaltenvektoren von* A, *dann hat das Gleichungssystem* $Ax = b$ *für* $i = 1, ..., n$ *die eindeutig bestimmte Lösung*

$$x_i = \frac{1}{\det\ A}\ \det(a_1, ..., a_{i-1}, b, a_{i+1}, ..., a_n)\ . \tag{4.16}$$

Die Cramer'sche Regel liefert uns eine Berechungsformel für die **Inverse Matrix**:

Zu $A \in \mathrm{GL}(n, \mathbb{K})$ mit Spaltenvektoren $(a_1, ..., a_n)$ suchen wir die inverse Matrix $A^{-1} = X = (x_1, ..., x_n) = (x_{ij})$. Mit (4.12) gilt $\mathbf{1} = A \cdot X = (Ax_1, \ldots, Ax_n)$ und damit $Ax_j = e_j$ für $1 \leq j \leq n$. Die Cramersche Regel (4.16) liefert

$$x_{ij} = \frac{1}{\det\, A}\, \underbrace{\det\,(a_1, ..., a_{i-1}, \underbrace{e_j}_{\text{Spalte } i}, a_{i+1}, ..., a_n)}_{:=\tilde{A}_{ji}} = \frac{\det\, \tilde{A}_{ji}}{\det\, A} \,. \tag{4.17}$$

Nach $(i-1)$ Spaltenvertauschungen und $(j-1)$ Zeilenvertauschungen (wobei sich jeweils das Vorzeichen der Determinante ändert) ergibt sich

$$\tilde{A}_{ji} = \begin{pmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ * & \cdots & 1 & \cdots & * \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & * \cdots * \\ 0 & \\ \vdots & \boxed{\quad A_{ji} \quad} \\ 0 & \end{pmatrix} \tag{4.18}$$

DEFINITION 4.39 *Sei* $A \in \mathcal{M}(n \times n, \mathbb{K})$*, dann ist die* (j, i)*-te* ***Streichungsmatrix*** $A_{ji} \in \mathcal{M}((n-1) \times (n-1), \mathbb{K})$ *von* A *die Matrix, die aus* A *durch Streichung der* j*-ten Zeile und* i*-ten Spalte entsteht.*

Es gilt mit Satz 4.35 und Satz 4.36:

$$\det\, \tilde{A}_{ji} = (-1)^{j-1+i-1}\, \det\, A_{ji} \,. \tag{4.19}$$

Wir haben damit gezeigt:

SATZ 4.40 (MATRIX-INVERSION) *Sei* $A \in \mathrm{GL}(n, \mathbb{K})$ *und* $A_{ji} \in \mathrm{GL}(n-1, \mathbb{K})$ *die* (j, i)*-te Streichungsmatrix von* A*. Dann sind die Komponenten* x_{ij} *der zu* A *inversen Matrix* $A^{-1} = (x_{ij})$ *gegeben durch*

$$x_{ij} = (-1)^{i+j}\, \frac{\det\, A_{ji}}{\det\, A} \,.$$

Der folgende Satz, der **Laplace'sche Entwicklungssatz**, dient zur induktiven Berechnung von Determinanten.

SATZ 4.41 (LAPLACE'SCHER ENTWICKLUNGSSATZ) *Sei* $A = (a_{ij}) \in \mathcal{M}(n \times n; \mathbb{K})$ *und* A_{ij} *die* (i, j)*-te Streichungsmatrix von* A*, dann gilt*

$$\det\, A = \sum_{i=1}^{n} a_{ij}\, (-1)^{i+j}\, \det\, A_{ij} \,. \tag{4.20}$$

Wir sagen, $\det A$ *wird entwickelt nach der* j*-ten Spalte von* A*.*

Beweis:. Mit D1 und (4.19) gilt

$$\begin{aligned}\det A &= \det(a_1, ..., a_j, ..., a_n) = \det(a_1, ..., \underbrace{\sum a_{ij} e_i}_{\text{Spalte } j}, ..., a_n) \\ &= \sum_{i=1}^{n} a_{ij} \underbrace{\det(a_1, ..., e_i, ..., a_n)}_{\det\ \tilde{A}_{ij}} = \sum_{i=1}^{n} a_{ij} (-1)^{i+j} \det\ A_{ij}\end{aligned}$$

□

NOTIZ 4.42 (ENTWICKLUNG NACH ZEILEN) *Wegen* $\det A = \det A^t$ *(Korollar 4.29) ist analog zur Spaltenentwicklung nach Satz 4.41 auch die Entwicklung nach Zeilen möglich, d.h. die Determinante lässt sich berechnen durch*

$$\det\ A = \sum_{j=1}^{n} a_{ij}\ (-1)^{i+j}\ \det\ A_{ij}\ .$$

4.6 Orientierung und Wegzusammenhang von $\mathrm{GL}(n, \mathbb{R})$

DEFINITION 4.43 $A \in \mathrm{GL}(n, \mathbb{R})$ *heißt* ***stetig deformierbar*** *in* $B \in \mathrm{GL}(n, \mathbb{R}) :\Leftrightarrow$
Es gibt eine stetige Abbildung $\Phi :\ [0, 1] \to \mathrm{GL}(n, \mathbb{R})$ *mit* $\Phi(0) = A$ *und* $\Phi(1) = B$.
Φ *heißt auch* ***Homotopie*** *von* A *nach* B.

NOTIZ 4.44 *Eine stetige Deformation der Basis* $\mathcal{A} = (a_1, ..., a_n)$ *in die Basis* $\mathcal{B} = (b_1, ..., b_n)$ *ist definiert als stetige Deformation der zugehörigen Matrizen* A *und* B *mit den Basisvektoren als Spalten.*

Fundamental ist der folgende

SATZ 4.45 *Für* $A \in \mathrm{GL}(n, \mathbb{R})$ *sind die folgenden Aussagen äquivalent.*

(1) $\det\ A > 0$

(2) A *ist stetig deformierbar in* $\mathbf{1} \in \mathrm{GL}(n, \mathbb{R})$

Beweisskizze:. $(2) \Rightarrow (1)$: Sei Φ eine Homotopie von A nach $\mathbf{1}$, dann ist die Abbildung $\gamma : [0, 1] \to \mathbb{R}$, die gegeben ist durch $t \mapsto \gamma(t) = \det\ \Phi(t)$ stetig und es gilt $\gamma(0) = \det\ A$ und $\gamma(1) = \det\ \mathbf{1} = 1$.
Wäre nun $\det\ A < 0$, dann folgte aus dem Zwischenwertsatz, dass $t_0 \in [0, 1]$ existiert mit $\gamma(t_0) = \det\ \Phi(t_0) = 0$. im Widerspruch zu $\Phi(t_0) \in \mathrm{GL}(n, \mathbb{R}) = \{A \in \mathcal{M}(n \times n; \mathbb{R} \mid \det\ A \neq 0\}$.
$(1) \Rightarrow (2)$: Dieser Beweisteil ist viel schwerer. Zu benutzen ist hier der Gaußalgorithmus mit Einheitsmatrizen: $A \sim ... \sim \mathbf{1}$, implementiert durch Permutation von Zeilen und Addition des Vielfachen einer Zeile zu einer anderen Zeile (beides entspricht der Multiplikation mit Elementarmatrizen in $\mathrm{GL}(n, \mathbb{R})$).

□

DEFINITION 4.46 *Zwei Basen* $\mathcal{A} = (a_1, ..., a_n)$, $\mathcal{B} = (b_1, ..., b_n)$ *eines reellen Vektorraums* V *heißen* ***gleichorientiert*** $:\Longleftrightarrow \quad \exists f \in \mathrm{Aut}(V)$ *mit* $\det f > 0$ *und* $f(a_j) = b_j$.

NOTIZ 4.47 *Die Definition 4.46 liefert eine Äquivalenzrelation auf der Menge aller Basen von* V, *die genau zwei Äquivalenzklassen besitzt. Diese werden als Orientierung des* $\mathbb{R}$*-Vektorraums* V *bezeichnet. Eine orientierter* $\mathbb{R}$*-Vektorraum ist ein Paar* (V, or) *aus einem* $\mathbb{R}$*-Vektorraum und einer der beiden Orientierungen.*

KOROLLAR 4.48 *Seien* $\mathcal{A}$ *und* $\mathcal{B}$ *Basen eines* $\mathbb{R}$*-Vektorraums* V. *Dann gilt:*
$\mathcal{A}$ *und* $\mathcal{B}$ *gleichorientiert* $\Leftrightarrow$ *Es gibt stetige Deformationen von* $\mathcal{A}$ *nach* $\mathcal{B}$, *also*
$f \in C([0,1], Aut(V))$ *mit* $f(0)a_j = a_j$ *und* $f(1)a_j = b_j$

Beweis:. Sei $g \in \mathrm{Aut}(V)$ mit $g(a_j) = b_j$. Dann gilt $\det g > 0 \quad \Leftrightarrow \quad \det \Phi^{\mathcal{A}}_{\mathcal{A}}(g) > 0 \quad \Leftrightarrow$ Die Matrix $A := \Phi^{\mathcal{A}}_{\mathcal{A}}(g)$ ist stetig deformierbar in $\mathbf{1}$ vermöge $A(t) \Leftrightarrow \quad f(t) = (\Phi^{\mathcal{A}}_{\mathcal{A}})^{-1}\, A(t)$ liefert die Deformation von $\mathcal{A}$ in $\mathcal{B}$. □

4.7 Übungen

1. Beweisen Sie Satz 4.3!

2. Beweisen Sie Notiz 4.23.

3. Zum Beweis von Satz 4.45: Schreibe $A = PU^-DU^+$, wobei P eine Permutationsmatrix ist, D diagonal und U^+, U^- Dreiecksmatrizen, genauer

$$U^- \in \left\{ \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ * & & 1 \end{pmatrix} ;\ * \text{ hat Koeffizienten im } \mathbb{R} \right\} =: N^-$$

$$U^+ \in \left\{ \begin{pmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{pmatrix} ;\ * \text{ hat Koeffizienten im } \mathbb{R} \right\} =: N^+$$

$$D \in \{\mathrm{diag}(\lambda_1, ..., \lambda_n); \lambda_j \in \mathbb{R}^*\} =: T$$

T ist der "maximale Torus" (die maximale kommutative Untergruppe von $\mathrm{GL}(n, \mathbb{R})$). Dann ist

$$DU^+ \in \left\{ \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} ;\ \lambda_j \in \mathbb{R}^*, ... \right\} =: B^+$$

B^+ ist die Borel-Untergruppe. Diese Anteile werden einzeln deformiert :.

1) Bei $U^\pm$ schalten wir die Nicht-Diagonalelemente stetig aus:

für $0 \leq t \leq 1$ ist $U^+(t) = \mathbf{1} + t \begin{pmatrix} 0 & & * \\ & \ddots & \\ & & 0 \end{pmatrix}$ und $U^-(t) = \mathbf{1} + t \begin{pmatrix} 0 & & 0 \\ & \ddots & \\ * & & 0 \end{pmatrix}$.

2) Bei D werden die Diagonalelemente einzeln deformiert, also $D(t) = \mathrm{diag}(...\lambda_j(t)...)$, mit $\lambda_j(t) := \frac{\lambda_j}{|\lambda_j|} \underbrace{(t|\lambda_j| + 1 - t)}_{>0 \text{ für } 0 \leq t \leq 1}$.

3) Die Permuationsmatrix P ergibt sich also Produkt von Permutationsmatrizen P_i, die jeweils die i-te und die $i+1$-te Zeile vertauschen, also $P = \prod_i P_i$ mit

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

Diese Vertauschung wird als Deformation mittels einer Drehung ausgeführt, also

$$P_i(t) = \begin{pmatrix} 1 & & & & & 0 \\ & \ddots & & & & \\ & & -\cos t(\pi/2) & \sin t(\pi/2) & & \\ & & \sin t(\pi/2) & \cos t(\pi/2) & & \\ 0 & & & & \ddots & \\ & & & & & 1 \end{pmatrix}$$

Dann gilt

$$P_i(0) = \begin{pmatrix} 1 & & & & & 0 \\ & \ddots & & & & \\ & & -1 & & & \\ & & & 1 & & \\ 0 & & & & \ddots & \\ & & & & & 1 \end{pmatrix} \quad \text{und} \quad P_i(1) = P_i \quad \text{und} \quad \det P_i(t) = \pm 1$$

Insgesamt ergibt sich $A(t) = P(t)U^-(t)D(t)U^+(t)$ mit $A(0) = \begin{pmatrix} \pm 1 & & * \\ & \ddots & \\ 0 & & \pm 1 \end{pmatrix}$ und $A(1) = A$.

Kapitel 5

Lineare Differentialgleichungen

5.1 Exponentialfunktion von Matrizen

DEFINITION 5.1 *Sei V ein reeller oder komplexer Vektorraum und $A \in \mathrm{End}(V) =: \mathcal{L}(V)$ und sei $\|\cdot\|$ eine Norm in V. Dann ist die **uniforme Norm** oder **Operatornorm** von A definiert als*

$$\|A\|_{\mathcal{L}} := \sup_{\substack{x \in V \\ \|x\|=1}} \|Ax\| \,.$$

LEMMA 5.2 *Für jede Norm $\|\cdot\|$ in V ist $\|\cdot\|_{\mathcal{L}} : \ \mathcal{L}(V) \to \mathbb{R}$ eine Norm und es gilt für alle $A, B \in \mathcal{L}(V)$*

$$\|A\|_{\mathcal{L}} = \sup_{y \in V} \frac{\|Ay\|}{\|y\|} \tag{5.1}$$

$$\|A \cdot B\|_{\mathcal{L}} \leq \|A\|_{\mathcal{L}} \, \|B\|_{\mathcal{L}} \,. \tag{5.2}$$

Beweis:. 1) $\|\cdot\|_{\mathcal{L}}$ ist eine Norm (Übung)
2) Beweis von (5.1):

$$\sup_{\substack{x \in V \\ \|x\|=1}} \|Ax\| = \sup_{y \in V} \left\| A \frac{y}{\|y\|} \right\| = \sup_{y \in V} \frac{1}{\|y\|} \|Ay\| \,.$$

3) Beweis von (5.2): Aus (5.1) folgt sofort $\|Ay\| \leq \|A\|_{\mathcal{L}} \|y\|$ für alle $y \in V$ und damit $\|A \cdot Bx\| \leq \|A\|_{\mathcal{L}} \|Bx\| \leq \|A\|_{\mathcal{L}} \|B\|_{\mathcal{L}} \|x\|$ für alle $x \in V$. Division durch $\|x\| \neq 0$ und Supremumsbildung liefert die Behauptung.

□

SATZ 5.3 *Für $V = \mathbb{R}^n$ oder $V = \mathbb{C}^n$ ist der normierte Raum $(\mathcal{L}(V), \|\cdot\|_{\mathcal{L}})$ vollständig.*

Erinnerung: Ein normierter Vektorraum $(V, \|.\|)$ ist vollständig, falls jede Cauchy-Folge konvergiert, genauer: für jede Folge $(a_j)_{j \in \mathbb{N}}$ mit $a_j \in V$ und

$$\forall \epsilon > 0 \ \exists N \in \mathbb{N} \ \forall m, n > N : \ \|a_n - a_m\| < \epsilon$$

existiert ein $a \in V$, so dass gilt

$$\forall \epsilon > 0 \ \exists N \in \mathbb{N} \ \forall n > N \ : \ \|a_n - a\| < \epsilon \ .$$

Für den Beweis des Satzes verwenden wir die Äquivalenz von Normen.

DEFINITION 5.4 *Sei V ein Vektorraum und seien $\|\cdot\|_1$ und $\|\cdot\|_2$ zwei Normen auf V. Dann heissen $\|\cdot\|_1$ und $\|\cdot\|_2$* ***äquivalent*** $:\Leftrightarrow$

$$\exists C > 0 \ \forall x \in V \ : \qquad \frac{1}{C}\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1 \ .$$

Bemerkung: Das definiert eine Äquvalenzrelation auf der Menge aller Normen auf einem Vektorraum. Konvergenz von Folgen und Offenheit von Mengen hängen nur von der Äquivalenzklasse ab.

SATZ 5.5 *Alle Normen auf $\mathbb{R}^n$ sind äquivalent.*

Beweis. Sei $(e_1, \ldots, e_n)$ die Standardbasis von $\mathbb{R}^n$, dann ist durch

$$\|v\|_1 := \Big\|\sum_{j=1}^n v_j e_j\Big\|_1 := \max_{j \in \{1,\ldots n\}} |v_j| \ , \quad v \in \mathbb{R}^n \ ,$$

eine Norm auf $\mathbb{R}^n$ gegeben (die Supremumsnorm), wobei $|\cdot|$ den Betrag in $\mathbb{R}$ bezeichnet.
$\|\cdot\|_2$ sei eine beliebige andere Norm auf $\mathbb{R}^n$. Wir werden zeigen, dass $\|\cdot\|_2$ äquivalent zu $\|\cdot\|_1$ ist, dann folgt aus der Transitivität der Äquivalenz die Behauptung.
Sei $S_1^{n-1} := \{v \in \mathbb{R}^n \mid \|v\|_1 = 1\}$ die $n-1$-Sphäre in $\mathbb{R}^n$ bzgl. $\|\cdot\|_1$. Dann ist S_1^{n-1} das Urbild der abgeschlossenen Menge $\{1\} \subset \mathbb{R}$ unter der stetigen Abbildung $\|\cdot\|_1 : (\mathbb{R}^n, \|\cdot\|_1) \to (\mathbb{R}, |\cdot|)$ (die Abbildung ist stetig, da aus der Dreiecksungleichung folgt $\big|\|v_1\| - \|v_2\|_1\big| \leq \|v_1 - v_2\|_1$). Damit (mit der topologische Definition von Stetigkeit) ist S_1^{n-1} ebenfalls abgeschlossen. Mit dem Satz von Heine-Borel ist S_1^{n-1} damit kompakt in $(\mathbb{R}^n, \|\cdot\|_1)$.
Die Abbildung $\|\cdot\|_2 : \ (\mathbb{R}^n, \|\cdot\|_1) \to (\mathbb{R}, |\cdot|)$ ist stetig, denn (mit der Dreiecksungleichung)

$$\begin{aligned}\Big|\|x\|_2 - \|y\|_2\Big| \leq \|x-y\|_2 = \Big\|\sum_{j=1}^n (x_j - y_j)e_j\Big\|_2 &\leq \sum_{j=1}^n |x_j - y_j| \|e_j\|_2 \\ &\leq \max_j |x_j - y_j| \sum_j \|e_j\|_2 \leq \|x-y\|_1 C \ .\end{aligned}$$

Nun nehmen stetige Funktionen auf Kompakta Minimum und Maximum an, es existieren also $C, C_1, C_2 > 0$ so dass

$$\frac{1}{C} \leq C_1 \leq \|\hat{x}\|_2 \leq C_2 \leq C \ , \qquad \hat{x} \in S_1^{n-1} \ .$$

Für $x = \|x\|_1 \frac{x}{\|x\|_1} = \|x\|_1 \hat{x} \in \mathbb{R}^n$ folgt dann

$$\frac{1}{C}\|x\|_1 \leq \|x\|_1 \|\hat{x}\|_2 = \|x\|_2 \leq C\|x\|_1 \ .$$

□

Beweis von Satz 5.3. Der Raum $\mathcal{L}(\mathbb{C}^n)$ ist isomorph zu $\mathbb{C}^{n^2}$ und damit vollständig z.B. in der Norm $\|A\|_m := \max_{i,j} |A_{ij}|$. Aber in $\mathbb{C}^{n^2}$ sind alle Normen äquivalent (mit Satz 5.5), also folgt die Behauptung.

$\square$

LEMMA 5.6 *In einem vollständigen normierten Vektorraum (= Banachraum) $(V, \|\cdot\|)$ ist jede absolut konvergente Reihe auch konvergent.*

Beweis. Wir müssen zeigen: falls $\sum_{j=1}^{\infty} a_j$ absolut konvergent ist (also falls $\sum_j \|a_j\| < \infty$), dann ist die Folge der Partialsummen $S_m := \sum_{j=1}^{M} a_j$ eine Cauchyfolge in V.
Sei also $\sum_{j=1}^{\infty} \|a_j\|$ konvergent, dann ist die Folge der Partialsummen $\tilde{S}_m := \sum_{j=1}^{m} \|a_j\|$ konvergent und damit insbesondere eine Cauchy-Folge. Weiterhin gilt (sei o.B.d A. $m > n$)

$$\|S_m - S_n\| = \|\sum_{j=n+1}^{m} a_j\| \leq \sum_{j=n+1}^{m} \|a_j\| = |\tilde{S}_m - \tilde{S}_n| \ .$$

Da $\tilde{S}_k$ eine Cauchy-Folge ist, existiert zu jedem $\epsilon > 0$ ein $M \in \mathbb{N}$, so dass für alle $m, n > M$ gilt $|\tilde{S}_m - \tilde{S}_n| < \epsilon$ und damit ist (S_k) eine Cauchy-Folge..

$\square$

SATZ 5.7 *Für $A \in \mathcal{L}(\mathbb{C}^n)$ ist*

$$\exp A = e^A := \sum_{k=0}^{\infty} \frac{1}{k!} A^k$$

absolut konvergent und konvergent und es gilt

1. *falls $AB = BA$, dann ist $e^{A+B} = e^A \cdot e^B$.*
2. *die Abbildung $\mathbb{C} \ni t \mapsto e^{tA} \in \mathcal{L}(\mathbb{C}^n)$ ist differenzierbar und es gilt $\frac{d}{dt} e^{tA} = A \cdot e^{tA}$.*
3. *e^A ist invertierbar mit $(e^A)^{-1} = e^{-A}$.*

Beweisskizze:. Mit Lemma 5.2 ist $\|A^k\|_{\mathcal{L}} \leq \|A\|_{\mathcal{L}}^k$ und damit hat $\sum_k \left\| \frac{A^k}{k!} \right\|$ die absolut konvergente Majorante $\sum_k \frac{\|A\|_{\mathcal{L}}^k}{k!} = e^{\|A\|_{\mathcal{L}}}$. Also ist die Reihe $\sum_k \frac{A^k}{k!}$ absolut konvergent und mit Satz 5.3 und Lemma 5.6 auch konvergent. Also ist e^A wohldefiniert (und $\|e^A\| \leq e^{\|A\|}$).
zu 1) Wegen $AB = BA$ berechnet man $e^A \cdot e^B = \left(\sum_k \frac{A^k}{k!}\right)\left(\sum_j \frac{B^k}{j!}\right)$ über das Cauchy-Produkt absolut konvergenter Reihen wie für komplexe Zahlen. Mit dem Binomialsatz

$$(A + B)^n = \sum_{j=0}^{n} \binom{n}{j} A^j B^{n-j}$$

folgt dann die Behauptung.
zu 2) Mit 1) gilt $\frac{e^{tA} - e^{t_0 A}}{t - t_0} = e^{t_0 A} \cdot \frac{e^{(t-t_0)A} - \mathbf{1}}{t - t_0}$ und die Norm von der rechten Seite konvergiert für $t \to t_0$ gegen A (wie im Beweis für komplexe Zahlen).
zu 3) klar mit 1).

$\square$

SATZ 5.8 *Sei* $A \in \mathcal{L}(\mathbb{C}^n)$ *und* $x_0 \in \mathbb{C}^n$*, dann ist* $x(t) := e^{tA}x_0$ *die eindeutig bestimmte Lösung der Anfangswertproblems (AWP) in* $\mathcal{C}^1(\mathbb{R}, \mathbb{C}^n)$

$$\dot{x}(t) := \frac{d}{dt}x(t) = Ax(t) \,, \qquad x(0) = x_0 \,. \tag{5.3}$$

Beweis:. Die Funktion $e^{tA}x_0$ ist differenzierbar (sogar $\mathcal{C}^\infty(\mathbb{R}, \mathbb{C}^n)$) und offensichtlich Lösung von (5.3) (mit Satz 5.7). Sei nun $y(t) \in \mathcal{C}^1(\mathbb{R}, \mathbb{C}^n)$ eine Lösung von (5.3). Dann gilt (mit der Produktregel für Matrixmultiplikation)

$$\frac{d}{dt}\Big(e^{-tA}y(t)\Big) = -Ae^{-tA}y(t) + e^{-tA} \cdot Ay(t) = 0$$

da A und e^{tA} kommutieren, und damit folgt $e^{-tA}y(t) = const = x_0$, also $y(t) = e^{tA}x_0$.

□

5.2 Differentialoperator zweiter Ordnung

DEFINITION UND SATZ 5.9 *Für* $a, b \in \mathbb{C}$ *heisst*

$$P = \frac{d^2}{dt^2} + a\frac{d}{dt} + b \;:\; \mathcal{C}^2(\mathbb{R}, \mathbb{C}) \to \mathcal{C}^0(\mathbb{R}, \mathbb{C})$$

Differentialoperator zweiter Ordnung mit konstanten Koeffizienten*. Wir schreiben auch* $\frac{d}{dt} = \partial_t$*.* P *ist eine lineare Abbildung und* $\operatorname{Kern} P$ *besteht aus den Lösungen der homogenen Differentialgleichung* $Pu = 0$*.*

NOTIZ 5.10 *Sei* $P = \partial_t^2 + a\partial_t + b$ *ein Differentialoperator zweiter Ordnung. Dann lässt sich* $Pu = 0$ *transformieren auf ein System von Differentialgleichungen erster Ordnung:*
Sei $x = (x_1, x_2) : \mathbb{R} \to \mathbb{C}^2$ *gegeben durch* $x_1 = u$ *und* $x_2 = \dot{u} := \partial_t u$ *und* $u \in \mathcal{C}^2(\mathbb{R}, \mathbb{C})$ *löse* $Pu = \ddot{u} + a\dot{u} + bu = 0$*. Dann folgt*

$$\dot{x} = \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} \dot{u} \\ \ddot{u} \end{pmatrix} = \begin{pmatrix} \dot{u} \\ -a\dot{u} - bu \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix} \begin{pmatrix} u \\ \dot{u} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \,.$$

Es gilt also $\dot{x} = Ax$ *mit* $A = \left(\begin{smallmatrix} 0 & 1 \\ -b & -a \end{smallmatrix}\right)$*.*
Ist andererseits $x \in \mathcal{C}^1(\mathbb{R}, \mathbb{C}^2)$ *Lösung von* $\dot{x} = Ax$*, dann ist* $u = x_1 \in \mathcal{C}^2(\mathbb{R}, \mathbb{C})$ *und löst* $Pu = 0$*.* *(selber nachrechnen!)*

SATZ 5.11 *Sei* P *Differentialoperator zweiter Ordnung wie in Defintion 5.9. Dann ist das Anfangswertproblem (AWP)*

$$Pu = 0 \,, \quad u(0) = u_0 \,, \quad \dot{u}(0) = u_1 \tag{5.4}$$

eindeutig lösbar in $\mathbb{C}^2$*.*

Beweis:. $u \in \mathcal{C}^2$ löst (5.4) $\Leftrightarrow x(t) := \begin{pmatrix} u(t) \\ \dot{u}(t) \end{pmatrix}$ löst $\dot{x}(t) = Ax(t)$ und $x(0) = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} =: x_0 \in \mathbb{C}^2$. Dieses AWP hat nach Satz 5.8 die eindeutige Lösung $x(t) = e^{tA}x_0$. □

KOROLLAR 5.12 $\mathrm{Kern}\, P$ *ist ein Unterverktorraum von* $\mathcal{C}^2(\mathbb{R}, \mathbb{C})$ *der Dimension 2. Eine Basis von* $\mathrm{Kern}\, P$ *heisst Fundamentalsystem von* P.

Beweis:. Sei $j : \mathbb{C}^2 \to \mathrm{Kern}\, P$, gegeben durch $\begin{pmatrix} u_0 \\ u_1 \end{pmatrix} \mapsto u(\cdot)$, wobei u das AWP $Pu = 0$ mit $u(0) = u_0$ und $\dot{u}(0) = u_1$ löst. Dann ist j linear (nachrechnen) und bijektiv (wegen Satz 5.11), also ein Isomorphismus. □

5.2.1 Berechnung des Kerns von P

Wir machen für die Lösung von $Pu = 0$ (für $P = \partial_t^2 + a\partial_t + b$) einen Exponentialansatz: $u(t) = e^{\lambda t}$. Setzen wir diesen Ansatz in die Gleichung ein, so liefert das $(\lambda^2 + a\lambda + b)e^{\lambda t} = 0$. Da die Exponentialfunktion nie null wird, folgt hieraus $\lambda^2 + a\lambda + b = 0$ und damit $\lambda_\pm = -\frac{a}{2} \pm \sqrt{\frac{a^2}{4} - b} =: \frac{1}{2}\big(-a \pm \sqrt{D}\big)$. Für $D \neq 0$ gibt es zwei Lösungen λ_+ und λ_- (wohldefiniert für $a, b \in \mathbb{C}$).

DEFINITION 5.13 $D = a^2 - 4b$ *heisst* ***Diskriminante*** *der quadratischen Gleichung* $\lambda^2 + a\lambda + b = 0$, *der sogenannten* ***charakteristischen Gleichung*** *des Differentialoperators* $P = \partial_t^2 + a\partial_t + b$.

SATZ 5.14 *Sei* $P = \partial_t^2 + a\partial_t + b$ *ein Differentialoperator zweiter Ordnung und* D *die Diskriminante der charakteristischen Gleichung von* P. *Dann gilt:*

1. *für* $D \neq 0$ *bildet* $\left\{ e^{\lambda_+ t}, e^{\lambda_- t} \right\}$ *ein Fundamentalsystem der Differentialgleichung* $Pu = 0$.
2. *für* $D = 0$ *bildet* $\left\{ e^{\lambda t}, te^{\lambda t} \right\}$ *ein Fundamentalsystem der Differentialgleichung* $Pu = 0$. *Dabei ist* $\lambda = \lambda_- = \lambda_+ = -\frac{a}{2}$ *die einzige Lösung der charakteristischen Gleichung von* P.

Beweisskizze:.

1. Für $D \neq 0$ prüft man nach, dass $Pe^{\lambda_\pm t} = 0$ und dass $\left(e^{\lambda_- t}, e^{\lambda_+ t}\right)$ linear unabhängig sind.
2. Heuristik: Hier ist $\lambda_+ = \lambda_-$. Für $\lambda := \lambda_-$ betrachten wir $\lim_{\lambda_+ \to \lambda} \frac{e^{\lambda_+ t} - e^{\lambda t}}{\lambda_+ - \lambda} = \frac{d}{d\lambda} e^{\lambda t} = te^{\lambda t}$.

□

5.2.2 Ausblick

Ist P ein Differentialoperator zweiter Ordnung wie in Satz 5.14, dann hat die inhomogene Differentialgleichung $Pu = f$, falls sie lösbar ist, als Lösungsraum einen affinen Unterraum zu dem Vektorraum $\mathrm{Kern}\, P$.

Ist (u_1, u_2) ein Fundamentalsystem von $Pu = 0$ (vgl. Korollar 5.12), dann ist die zugehörige **Wronski-Determinante** gegeben durch

$$W(u_1, u_2) = \det \begin{pmatrix} u_1 & u_2 \\ \dot{u}_1 & \dot{u}_2 \end{pmatrix} = u_1\dot{u}_2 - u_2\dot{u}_1 \ . \tag{5.5}$$

W löst die Differentialgleichung $\dot{W} = -aW$, ist also gegeben durch $W(t) = e^{-at}W(0)$.
Die Funktion

$$K(t,s) := \frac{1}{W(s)}\big(u_2(t)u_1(s) - u_2(s)u_1(t)\big) \tag{5.6}$$

löst dann das Anfangswertproblem $P(\partial_t)K(t,s) = 0$ mit $K(t,t) = 0$ und $\partial_t K(t,s) = 1$ und die Lösung der inhomogenen Differentialgleichung $Pu = f$ ist gegeben durch

$$y(t) = \int_0^t K(t,s)f(s)\,ds \tag{5.7}$$

mit den Anfangswerten $y(0) = 0$ und $\dot{y}(0) = 0$.

NOTIZ 5.15 *Da für die Wronski-Determinante gilt* $W(t) = e^{-at}W(0)$ *folgt* $W(t_0) = 0$ *für ein* $t_0 \in \mathbb{R} \quad \Leftrightarrow \quad W(t) = 0\ \forall t \in \mathbb{R} \quad \Leftrightarrow \quad \begin{pmatrix} u_1 \\ \dot{u}_1 \end{pmatrix}$ *und* $\begin{pmatrix} u_2 \\ \dot{u}_2 \end{pmatrix}$ *sind linear unabhängig* $\Leftrightarrow \quad (u_1, u_2)$ *sind linear unabhängig in* $\mathrm{Kern}\, P$.

5.3 Übungen

1. Führen Sie Punkt 1) im Beweis von Lemma 5.2 durch.

2. Führen Sie den Beweis von Satz 5.14 durch.

Kapitel 6

Summen und Quotienten von Vektorräumen, Dualität

6.1 Summen und direkte Summen

DEFINITION 6.1 *Seien $V_1, ..., V_n$ Untervektorräume des $\mathbb{K}$-Vektorraumes V. Dann heisst die Menge*

$$\sum_{i=1}^{n} V_i := V_1 + \ldots + V_n := \left\{ v \in V \mid v = \sum_{i=1}^{n} v_i \text{ mit } v_i \in V_i \right\}$$

Summe der Untervektorräume $V_1, ..., V_n$.

NOTIZ 6.2 *1) $\sum_i V_i$ ist offensichtlich ein Untervektorraum von V.*

2) Die Verknüpfung "$+$" zwischen den Untervektorräumen ist kommutativ und assoziativ.

3) $\sum_i V_i$ ist der kleinste Untervektorraum von V, der alle $V_1, ..., V_n$ enthält.

Wir benötigen die folgende wichtige Dimensionsformel.

SATZ 6.3 *Seien V_1, V_2 Untervektorräume des endlichdimensinalen $\mathbb{K}$-Vektorraumes V, dann sind $V_1 \cap V_2$ und $V_1 + V_2$ endlichdimensional und es gilt*

$$\dim(V_1 \cap V_2) + \dim(V_1 + V_2) = \dim V_1 + \dim V_2 \ . \tag{6.1}$$

Beweis. Nach Wahl einer Basis von V sieht man sofort, dass alle anderen auftretenden Vektorräume die endliche Dimension als Untervektorräume von V erben. Zu zeigen bleibt die Dimensionsformel (6.1).
Seien $m := \dim V_1 \cap V_2$, $n := \dim V_1$ und $p := \dim V_2$ und sei $B_{1,2} := (a_1, ..., a_m)$ eine Basis von $V_1 \cap V_2$. Nach dem Basisergänzungssatz (Satz 1.21) lässt sie sich erweitern zu einer Basis $B_1 := (a_1, ..., a_m, b_{m+1}, ..., b_n)$ von V_1 und $B_2 := (a_1, ..., a_m, c_{m+1}, ..., c_p)$ von V_2.

Behauptung: $M := (a_1, ..., a_m, b_{m+1}, ..., b_n, c_{m+1}, ..., c_p)$ ist eine Basis von $V_1 + V_2$.
Dann folgt (6.1), denn $\dim(V_1 + V_2) = \#M = n + (p - m) = \dim V_2 + \dim V_1 - \dim(V_1 \cap V_2)$.

Beweis der Behauptung: Offensichtlich ist $V_1 + V_2 = \operatorname{Span} M$. Zu zeigen bleibt die lineare Unabhängigkeit von M:
Widerspruchsannahme: Es gibt eine nichttriviale Linearkombination, d.h. es gilt nicht $x_i = y_j = z_k = 0$ für alle i, j, k, aber

$$\sum_{i=1}^{m} x_i a_i + \sum_{j=m+1}^{n} y_j b_j + \sum_{k=m+1}^{p} z_k c_k = 0 \tag{6.2}$$

Dann muss aber ein $k \in \{m+1, \dots p\}$ existieren mit $z_k \neq 0$, denn sonst wäre schon $\sum x_i a_i + \sum y_j b_j = 0$ eine nichttriviale Linearkombination und damit B_1 keine Basis von V_1. Der Vektor $w := \sum z_k c_k \neq 0$ liegt aber in V_2 da $c_{m+1}, \dots, c_p \in V_2$ nach Definition, wegen (6.2) liegt w aber auch in V_1 als Linearkombination der Elemente von B_1. Damit folgt aber $w \in V_1 \cap V_2$ und da $B_{1,2}$ eine Basis ist, gibt es eine Linearkombination $w = \sum \mu_i a_i$ mit $\mu_i \neq 0$ für mindestens ein i. Daraus folgt aber sofort

$$\sum_{k=m+1}^{p} z_k c_k - \sum_{i=1}^{m} \mu_i a_i = 0 \quad \text{und es existierten } k, i \text{ mit } z_k \neq 0 \neq \mu_i\ .$$

Das ist ein Widerspruch dazu, dass B_2 eine Basis von V_2 ist und die Behauptung ist bewiesen.

□

Wir wollen jetzt den besonders schönen Fall, dass $V_1 \cap V_2$ nur aus der Null besteht, durch einen neuen Begriff besonders auszeichnen. Wir werden später sehen, dass in diesem Fall die Zerlegung von $v \in V_1 + V_2$ in seine Komponenten aus V_1 bzw. V_2 eindeutig wird. Das benutzen wir in der folgenden Definition nun umgekehrt zur Charakterisierung der direkten Summe.

DEFINITION 6.4 *Der $\mathbb{K}$-Vektorraum V heißt **direkte Summe** seiner Teilräume $V_1, ..., V_n$:⇔ Jedes $v \in V$ hat eine **eindeutige** Darstellung*

$$v = \sum_{i=1}^{n} v_i \ \textit{mit } v_i \in V_i.$$

Als Kurzschreibweise wird notiert $V = \bigoplus_{i=1}^{n} V_i = V_1 \oplus \ldots \oplus V_n$, das Zeichen $\bigoplus$ bzw. $\oplus$ wird dabei als "direkte Summe" gelesen.

Den Zusammenhang dieser Definition mit der Eigenschaft eines minimalen Durchschnitts aller beteiligten Untervektorräume liefert der folgende

SATZ 6.5 *Seien $V_1, ..., V_n$ Untervektorräume des $\mathbb{K}$-Vektorraumes V mit $\dim V < \infty$, dann gilt*

$$\begin{aligned} V = \bigoplus_{i=1}^{n} V_i \quad &\Leftrightarrow \quad \sum_{i=1}^{n} V_i = V \ \wedge\ V_j \cap (\sum_{j \neq i} V_i) = \{0\},\ 1 \leq j \leq n \\ &\Leftrightarrow \quad \sum_{i=1}^{n} V_i = V \ \wedge\ \dim V = \sum_{i=1}^{n} \dim V_i \end{aligned}$$

Der Beweis ist eine nicht völlig triviale Übungsaufgabe für fleißige Studenten.

Wir werden nun die direkte Summenzerlegung eines Vektorraums alternativ durch eine Familie von sogenannten Projektoren charakterisieren.

DEFINITION 6.6 *Sei V ein $\mathbb{K}$-Vektorraum, dann heisst* $\Pi \in \mathrm{End}(V)$ ***Projektor*** *oder* ***Projektion*** $:\Leftrightarrow$ Π *ist* ***idempotent****, das heisst* $\Pi^2 = \Pi \circ \Pi = \Pi$.

Bemerkung: Jede direkte Summenzerlegung $V = \bigoplus_{i=1}^n V_i$ induziert Abbildungen

$$\Pi_j : V \longrightarrow V_j \subset V\,, \quad v = \sum_{j=1}^n v_j \longmapsto \Pi_j(v) = v_j\ .$$

Dabei ist $v_j \in V_j$ die eindeutig bestimmte Komponente in der direkten Summenzerlegung von $v \in V$. Offensichtlich gilt für alle $v \in V$:

1) $\Pi_j^2(v) = \Pi_j(v_j) = v_j = \Pi_j v$, also $V_j = \mathrm{Bild}\,\Pi_j$ und $\Pi_j^2 = \Pi_j$.

2) Für $i \neq j$ gilt $\Pi_i\Pi_j(v) = \Pi_i(v_j) = 0$, also $\Pi_i\Pi_j = 0$ für $i \neq j$.

3) Es gilt $\sum_{j=1}^n \Pi_j = 1$, denn für $v \in V$ mit $v = \sum_k v_k = \sum_k \Pi_k v_k$ für $v_k \in V_k$ gilt:

$$\Big(\sum_j \Pi_j\Big)v = \Big(\sum_j \Pi_j\Big)\Big(\sum_k \Pi_k v_k\Big) = \sum_{j,k} \Pi_j\Pi_k v_k$$
$$= \sum_{j,k} \Pi_j \delta_{jk} v_k = \sum_j \Pi_j v_j = \sum_j v_j = \mathbf{1}v$$

Eine Umkehrung dieser Beobachtungen liefert der folgende

SATZ 6.7 *Seien* $\Pi_1, ..., \Pi_n \in \mathrm{End}(V)$ *eine endliche Familie von Projektoren mit* $\sum_{i=1}^n \Pi_i = \mathbf{1}$ *und* $\Pi_i\Pi_j = 0$ *für* $1 \leq i \neq j \leq n$. *Dann gilt* $V = \bigoplus_{i=1}^n \mathrm{Bild}\,\Pi_i$.

Beweis. Nach Voraussetzung gilt $V \ni v = \mathbf{1}v = \sum_{j=1}^n \Pi_j v = \sum_{j=1}^n v_j$ mit $\Pi_j v =: v_j \in V_j :=$ Bild Π_j. Also ist $V = \sum_{j=1}^n V_j$ und zu zeigen bleibt die Direktheit der Summenzerlegung, also die Eindeutigkeit der Zerlegung $v = \sum v_j$ mit $v_j \in V_j$. Wir nehmen an $v = \sum_{j=1}^n w_j$ mit $w_j \in V_j$ ist eine beliebige Summenzerlegung von $v \in V$, dann gilt (da $\Pi_j|_{V_j} = \mathbf{1}$):

$$v_k = \Pi_k v = \Pi_k \sum_j w_j = \Pi_k \sum_j \Pi_j w_j = \sum_j \Pi_k \Pi_j w_j = \sum_j \delta_{jk} w_j = w_k\ .$$

□

NOTIZ 6.8 $v_j = \Pi_j v$ *ist in der Situation des Satzes die eindeutig bestimmte Komponente von* $v \in V$ *in Richtung des Untervektorraums* $V_j = \mathrm{Bild}\,\Pi_j$ *und zwar für eine Projektion auf* V_j *"längs* $\bigoplus_{i\neq j} V_i$*".*

NOTIZ 6.9 *Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und V_1 Untervektorraum von V. Dann gibt es einen nicht eindeutig bestimmten sogenannten* ***komplementären Unterraum*** *V_2 von V mit $V = V_1 \oplus V_2$.*
Man erhält ein V_2, indem man eine Basis $\{a_1, ..., a_m\}$ von V_1 mit dem Basisergänzungssatz zu einer Basis $\{a_1, ..., a_m, b_1, ..., b_p\}$ von V ergänzt: $V_2 := \mathrm{Span}(b_1, ..., b_p)$ ist dann ein direktes Komplement von V_1 in V.

Bis jetzt haben wir den Fall betrachtet, dass die einzelnen Vektorräume $V_1, ..., V_n$ in einem gemeinsamen Oberraum V enthalten waren. Im Allgemeinen muss dieser Oberraum mit dem mittlerweile ja hoffentlich vertrauten Mengen erst herbei gezaubert werden.

DEFINITION 6.10 *Seien $V_1, ..., V_n$ $\mathbb{K}$-Vektorräume. Wir definieren die* ***externe direkte Summe*** *$V := \bigoplus_{j=1}^n V_j$ mit der $\mathbb{K}$-Vektorraum-Struktur $(V, +, \cdot)$ wie folgt:*

1) *Als Menge ist V das kartesische Produkt $V_1 \times ... \times V_n$, d.h.*
   $V = \{v = (v_1, ..., v_n) \,|\, v_j \in V_j,\ j = 1, ..., n\}$.

2) *Addition in V und Multiplikation mit Skalaren $\lambda \in \mathbb{K}$ sind komponentenweise erklärt:*

$$\begin{aligned} v + \tilde{v} &= (v_1, ..., v_n) + (\tilde{v}_1, ..., \tilde{v}_n) := (v_1 + \tilde{v}_1, ..., v_n + \tilde{v}_n) \\ \lambda v &= \lambda(v_1, ..., v_n) = (\lambda v_1, ..., \lambda v_n) \end{aligned}$$

SATZ 6.11 *Sind $V_1, \ldots V_n$ $\mathbb{K}$-Vektorräume, dann ist die externe direkte Summe $V := \bigoplus_{j=1}^n V_j$ ebenfalls ein $\mathbb{K}$-Vektorraum.*

NOTIZ 6.12 *Ist V die externe direkte Summe von $V_1, \ldots V_n$, dann ist die sogenannte* ***natürliche Injektion*** *$\phi_j :\ V_j \to V\ ;\ v_j \mapsto (0, ..., 0, v_j, 0, ...)$ eine injektive Abbildung und liefert eine Identifikation von V_j mit einem Teilraum von V. Offensichtlich ist dann $V = \bigoplus_{j=1}^n \phi_j(V_j)$ eine interne direkte Summenzerlegung im alten Sinn. In diesem Sinn ist nichts neues passiert.*
Solche Identifizierungen hatten wir natürlich schon immer vorgenommen. Zum Beispiel bei

$$\mathbb{K}^n = \bigoplus_{j=1}^{n} \mathbb{K} \overset{als Menge}{=} \underbrace{\mathbb{K} \times ... \times \mathbb{K}}_{n\,mal} .$$

NOTIZ 6.13 *Die Notation einer direkten Summe übertragen wir auch auf lineare Abbildungen, die eine direkte Summenzerlegung des Vektorraums respektieren:*
Seien $V = \bigoplus_{j=1}^n V_j$ und $W = \bigoplus_{j=1}^n W_j$. Weiter sei $f \in \mathrm{Hom}(V, W)$ mit $f(V_j) \subset W_j$, $(j = 1, ..., n)$. Wir setzen dann $f_j := f|_{V_j}$ und vereinbaren die Notation $f = \bigoplus_{j=1}^n f_j$.
Diese Schreibweise bedeutet, dass f die direkte Summenzerlegung von V und W respektiert.
Bezüglich der Wahl von Basen $\mathcal{A}_j$ von V_j, $\mathcal{B}_j$ von W_j und Basen $\mathcal{A} = (\mathcal{A}_j)$, $\mathcal{B} = (\mathcal{B}_j)$ von V bzw. W hat f eine Darstellung als Blockmatrix:

$$\Phi_{\mathcal{B}}^{\mathcal{A}}(f) = \begin{pmatrix} A_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & A_n \end{pmatrix} , \quad \textit{mit} \quad A_j := \Phi_{\mathcal{B}_j}^{\mathcal{A}_j}(f_j)$$

Zur Konstruktion von Quotientenvektorräumen und ihrer geometrischen Interpretation als Äquivalenzklassen von translatierten Untervektorräumen eines Oberraums verweise ich auf andere Literatur. Stattdessen erfolgt hier ein

6.2 Exkurs zur Dualität

Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum. Bekanntlich ist der Dualraum $V^* = \mathrm{Hom}(V, \mathbb{K})$ isomorph zu V (zu jeder Basis $\mathcal{A} = (a_1, \ldots a_n)$ von V war ja $\mathcal{A}^* = (a^1, \ldots a^n)$ mit $a^j(a_k) = \delta_{jk}$ wiederum eine Basis von V^*, die sogenannte duale Basis). Der durch $\mathcal{A}$ induzierte Isomorphismus $V \to V^*$ ist aber abhängig von der Auswahl der speziellen und durch nichts ausgezeichneten Basis $\mathcal{A}$, also nicht kanonisch.
Das ändert sich, wenn man den Dualraum des Dualraums betrachtet, den sogenannten **Bidualraum** $V^{**} = \mathrm{Hom}(V^*, \mathbb{K})$.

DEFINITION UND SATZ 6.14 *1. Die Abbildung*

$$\varepsilon_V : V \to V^{**} = \mathrm{Hom}(V^*, \mathbb{K}) \,, \quad \varepsilon_V(v)(\ell) := \ell(v)$$

*heisst kanonischer Isomorphismus (der sogenannte **Auswertungsisomorphismus**).*
ε_V ist eine bijektive lineare Abbildung.

2. *Die Abbildung $(\cdot, \cdot) : V^* \times V \to \mathbb{K}$, die definiert ist durch $(\ell, v) := \ell(v)$, heisst **duale Paarung** zwischen V^* und V. Wir schreiben $\big(\varepsilon_V(v), \ell\big) =: (v, \ell)$*
$(\cdot, \cdot)$ ist bilinear, das heißt für $\mu, \lambda \in \mathbb{K}$, $\ell, f \in V^*$ und $v, w \in V$ gilt $(\mu\ell + \lambda f, v) = \mu(\ell, v) + \lambda(f, v)$ und $(\ell, \mu v + \lambda w) = \mu(\ell, v) + \lambda(f, w)$.

In den Koordinaten von V bzw. V^* bezüglich einer Basis $\mathcal{A} = (a_1, ..., a_n)$ bzw. der dualen Basis $\mathcal{A}^* = (a^1, ..., a^n)$ ist die duale Paarung nicht anderes als die Matrixmultiplikation von Zeilen, die Linearformen in V^* repräsentieren (Kovektoren), mit Spalten, die die Elemente in V repräsentieren (Vektoren). Wegen der Bilinearität und

$$(a^j, a_i) = \delta_{ij} = (\varepsilon_V(a_i), a^j) = (a_i, a^j),$$

gilt für Linearformen $\ell = \sum_{j=1}^n y_i a^j$ und Vektoren $v = \sum_{i=1}^n x_i a_i$

$$\begin{aligned}(\ell, v) &= \sum_{i,j=1}^n y_i x_i (a^j, a_i) = \sum_i y_i x_i = (y_1, ..., y_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= y^t x = x^t y = (x_1, ..., x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = (\varepsilon_V(v), \ell) = (v, \ell)\,.\end{aligned}$$

DEFINITION UND SATZ 6.15 *Seien V und W endlichdimensionale $\mathbb{K}$-Vektorräume und $f \in \operatorname{Hom}(V,W)$. Dann gibt es einen eindeutig bestimmten Homomorphismus $f^* \in \operatorname{Hom}(W^*,V^*)$ mit*

$$\big(f^*(w^*),v\big) = \big(w^*,f(v)\big)\,, \qquad \forall w^* \in W^*,\ v \in V\ . \tag{6.3}$$

f^* heißt die zu f ***duale Abbildung****. Entsprechend dem kommutativen Diagramm*

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[dr, "f^*w^*"'] & W \arrow[d, "w^* \in \operatorname{Hom}(W,K)"] \\ & \mathbb{K} \end{tikzcd}$$

heißt $f^*(w^*) := w^* \circ f$ manchmal auch die Zurückliftung von w^* unter f.

Beweisskizze. Eindeutigkeit:
Sind $f^*, \tilde{f}^* \in \operatorname{Hom}(W^*,V^*)$ zwei Abbildungen, die bzgl. f der Gleichung (6.3) genügen, so folgt sofort $\Big((f^* - \tilde{f}^*)(w^*),v\Big) = 0$ für alle $w^* \in W^*$ und $v \in V$ und damit $(f^* - \tilde{f}^*)(w^*) = 0$ in V^*, also $f^* = \tilde{f}^*$.
Existenz:
Bei festem $w^* \in W^*$ ist die Zurückliftung $f^*(w^*) :\ V \to \mathbb{K},\, v \mapsto (w^*,f(v))$, linear und damit $f^*(w^*)$ wohldefiniert als ein Element in V^*. Somit haben wir eine Abbildung $f^* :\ W^* \to V^*$ erhalten, die Gleichung (6.3) genügt. Zu zeigen bleibt, dass f^* linear ist. Das folgt sofort durch Ausschreiben der Definitionen unter Verwendung der Bilinearität und (6.3).

□

Wir zeigen jetzt, dass unter Verwendung von $\mathcal{A}$ und $\mathcal{A}^*$ der Übergang von f zu f^* nichts anderes ist, als das altbekannte Transponieren von Matrizen.

SATZ 6.16 *Seien V und W endlichdimensionale $\mathbb{K}$-Vektorräume mit den Basen $\mathcal{A} = (a_1,...,a_n)$ und $\mathcal{B} = (b_1,...,b_m)$ und dualen Basen $\mathcal{A}^* = (a^1,...,a^n)$ sowie $\mathcal{B}^* = (b^1,...,b^m)$ von V^* und W^*. Dann gilt für die Matrixdarstellungen von $f \in \operatorname{Hom}(V,W)$ und $f^* \in \operatorname{Hom}(W^*,V^*)$*

$$\Phi^{\mathcal{B}^*}_{\mathcal{A}^*}(f^*) = \big(\Phi^{\mathcal{B}}_{\mathcal{A}}(f)\big)^t\ .$$

Beweis. Setze $A = (A_{ij}) = \Phi^{\mathcal{A}}_{\mathcal{B}}(f)$ und $B = (B_{ij}) = \Phi^{\mathcal{B}^*}_{\mathcal{A}^*}(f^*)$, dann ist $f(a_j) = \sum_i A_{ij}b_i$ und $f^*(b^k) = \sum_i B_{ik}a^i$, also:

$$\big(b^k,fa_j\big) = \sum_i A_{ij}b^k(b_i) = A_{kj} \quad \text{und mit (6.3)}$$
$$\big(b^k,fa_j\big) = \big(f^*(b^k),a_j\big) = \sum_i B_{ik}a^i(a_j) = B_{jk}\ .$$

□

Bemerkung: Ganz analog zu den Eigenschaften der Matrixtransposition gilt für $f,g \in \operatorname{Hom}(V,W)$:

1. $(f+g)^* = f^* + g^*$ und $(\lambda f)^* = \lambda f^*$ (Linearität).

2. $(fg)^* = g^* f^*$ (Vertauschung der Reihenfolge).

3. $\mathbf{1}^* = \mathbf{1}$ und $0^* = 0$.

4. Mit der kanonischen Identifikation $V^{**} \cong V,\ W^{**} \cong W$ vermöge ε_V und ε_W wird $f^{**} : V^{**} \to W^{**}$ identifiziert mit $f:\ V \to W$. Entsprechend gilt $(A^t)^t = A$ für die charakteristische Matrix.

Die ganzen Dualitätsfragen werden ein wenig übersichtlicher, wenn man anstatt allgemeiner $\mathbb{K}$-Vektorräume reelle oder komplexe Vektorräume $(V, \langle\cdot,\cdot\rangle)$ mit Skalarprodukt betrachtet. Der Einfachheit halber betrachten wir endlichdimensionale Vektorräume. Dann gibt es nach Gram-Schmidt eine Orthonormalbasis $\mathcal{E} := (e_1, ..., e_n)$ von V und die Linearformen

$$e^j := \langle e_j,\ \cdot\,\rangle \in \mathrm{Hom}(V, \mathbb{K}),\ \mathbb{K} = \mathbb{R}, \mathbb{C}$$

bilden gerade die zu $\mathcal{E}$ duale Basis $\mathcal{E}^*$ von V^*. Damit sehen wir

NOTIZ 6.17 *Die Abbildung*

$$\phi_V : V \longrightarrow V^*, \quad v \longmapsto \phi_V(v) := \langle v, \cdot\rangle \in V^*$$

bildet Basen in Basen ab, ist also ein (Semi-) Isomorphismus. Das heißt genauer, dass ϕ für $\mathbb{K} = \mathbb{R}$ linear und für $\mathbb{K} = \mathbb{C}$ semilinear ist (d.h. $\phi(\lambda v) = \bar{\lambda}\phi(v)$ für $\lambda \in \mathbb{C}, v \in V$).

Bemerkung: Diese Dualität überträgt sich auch auf den Fall ∞-dimensionaler Hilberträume. Diese tauchen in der Physik vor allem in der Quantenmechanik auf und in diesem Zusammenhang ist eine auf Dirac zurückzuführende Schreibweise extrem populär geworden. In ihrer beabsichtigten Primitivität ist sie so suggestiv, dass man schon gar nicht mehr erkennt, was passiert ist:
Das Skalarprodukt $\langle\cdot,\cdot\rangle :\ V \times V \to \mathbb{K} = \mathbb{R}, \mathbb{C}$ schreibt man grundsätzlich als $\langle\cdot|\cdot\rangle$ und stilisiert **bracket** (=Skalarprodukt) weiter zu BRA-KET. Vektoren im ersten Eingang des Skalarprodukts, was in der Physik durch letztere Schreibweise immer der antilineare Eingang ist, heißen entsprechend BRA-Vektoren und werden als $\langle x|$ notiert. Bei uns sind das Elemente des Dualraums bzw. Kovektoren oder Linearformen, die man mit $\langle x, \cdot\rangle$ bezeichnet. Der zweite Eingang liefert KET-Vektoren, notiert als $|y\rangle$. Bei uns heißen diese schlicht Vektoren $y \in V$. Kurz:

$$\underbrace{\langle x|}_{BRA} \text{ wirkt auf } \underbrace{|y\rangle}_{KET} \text{ via } \langle x|y\rangle$$

$V \cong V^{**}$ erscheint dann als Umkehrung der Leserichtung von links-rechts auf rechts-links: KET $|y\rangle$ wirkt auf BRA $\langle x|$ via BRA-KET ebenfalls als $|y\rangle$ und $\langle x| \mapsto \langle x|y\rangle \in \mathbb{K} = \mathbb{R}, \mathbb{C}$

Seien nun wieder V, W endlich dimensionale Vektorräume mit Skalarprodukt $\langle\cdot,\cdot\rangle_V$ bzw. $\langle\cdot,\cdot\rangle_W$. Identifiziert man V, W mit ihren Dualräumen via Skalarprodukt, so erhält man für $f \in \mathrm{Hom}(V, W)$

$$\begin{tikzcd}
V^* & W^* \arrow[l, "f^*"'] \\
V \arrow[u, "\phi_V"] & W \arrow[l, "f^\dagger"'] \arrow[u, "\phi_W"']
\end{tikzcd}$$

den sogenannten **adjungierten Operator** $f^\dagger : W \to V$.

DEFINITION UND SATZ 6.18 *Zu $f \in \mathrm{Hom}(V,W)$ heißt gemäß dem obigen Diagramm $f^\dagger := \phi_V^{-1} \circ f^* \circ \phi_W : W \longrightarrow V$ die zu f* ***adjungierte Abbildung****.*
$f^\dagger$ ist eindeutig bestimmt durch die Eigenschaft

$$\langle w, fv\rangle_W = \langle f^\dagger w, v\rangle_V\,, \quad v \in V, w \in W \tag{6.4}$$

und die Abbildung $\dagger : \mathrm{Hom}(V,W) \to \mathrm{Hom}(W,V)$, die gegeben ist durch $f \mapsto \dagger(f) := f^\dagger$, ist ein Semi-Isomorphismus.

Beweis:. Offensichtlich ist $f^\dagger \in \mathrm{Hom}(W,V)$ wohldefiniert. Wegen der Antilinearität von ϕ_V^{-1} ist $\dagger$ ebenfalls antilinear und wegen $\dagger^2 = \mathbf{1}$ also ein Semi-Isomorphismus. Weiter gilt (wobei $(\cdot,\cdot) : V^* \times V \to \mathbb{K}$ die duale Paarung (siehe Def. 6.14) bezeichnet)

$$\begin{aligned}
\langle w, f(v)\rangle_W &= \big(\phi_W(w), f(v)\big) \overset{(6.3)}{=} \big(f^*(\phi_W(w)), v\big) = \big[f^*(\phi_W(w))\big](v) \\
&= \langle \phi_V^{-1}(f^*(\phi_W(w))), v\rangle = \langle f^\dagger w, v\rangle
\end{aligned}$$

Also gilt (6.4) und zweifelnde Gemüter verifizieren als Übung im Ausschreiben von Isomorphismen noch selber, dass (6.4) den adjungierten Operator auch tatsächlich eindeutig festlegt.

□

Notation
Im Folgenden werden wir in Vektorräumen mit Skalarprodukt systematisch V mit V^* identifizieren und die hilfreichen $\phi_V,\ \phi_W$ schlicht weglassen. Damit wird dem Leser aufgebürdet, je nach Kontext zu unterscheiden zwischen

- der dualen Abbildung $f^* \in Hom(W^*, V^*)$, auch **Banachraum-Adjungierte**,
- und der adjungierten Abbildung $f^\dagger \in Hom(W,V)$, auch **Hilbertraum-Adjungierte**.

Ab jetzt schreiben wir nämlich auch f^* für den adjungierten Operator $f^\dagger \in \mathrm{Hom}(W,V)$.

SATZ 6.19 *Seien V und W endlich-dimensionale Vektorräume mit Skalarprodukten $\langle\cdot,\cdot\rangle_V$ und $\langle\cdot,\cdot\rangle_W$ und sei $f \in \mathrm{Hom}(V,W)$, dann gilt:*

$$(\mathrm{Bild}\, f)^\perp = \mathrm{Kern}\, f^* \quad \textit{und} \quad (\mathrm{Bild}\, f^*)^\perp = \mathrm{Kern}\, f$$

Beweis. Es gilt $x \in \operatorname{Kern} f^* \Leftrightarrow f^*(x) = 0 \Leftrightarrow \langle f^*(x), y\rangle = 0$ für alle $y \in V$. Nach Definition ist das äquivalent zu $\langle x, f(y)\rangle = 0$ für alle $y \in V$ und damit zu $x \perp \operatorname{Bild} f$.
Die zweite Gleichung folgt aus der ersten mit $f^* =: g \in \operatorname{Hom}(W, V)$ und $f^{**} = f$.

□

Bemerkung: Ist $f \in \operatorname{Hom}(V, W)$, dann haben die darstellenden Matrizen $A = \Phi^{\mathcal{A}}_{\mathcal{B}}(f)$ und $B = \Phi^{\mathcal{B}}_{\mathcal{A}}(f^*)$ von f und f^* bezüglich Orthonormalbasen $\mathcal{A} = (a_1, ..., a_n)$ von V und $\mathcal{B} = (b_1, ..., b_n)$ von W die Komponenten:

$$A_{ij} = \langle b_i, f(a_j)\rangle = \langle f^*(b_i), a_j\rangle = \overline{\langle a_j, f^*(b_i)\rangle} = \overline{B_{ji}}$$

Also gilt $\overline{A^t} = B$.
Warnung: Das gilt nur für Orthonormalbasen!

DEFINITION 6.20 *In einem Vektorraum mit Skalarprodukt* $(V, \langle\cdot,\cdot\rangle)$ *heißt* $f \in \operatorname{End}(V)$ ***selbstadjungiert*** *oder* ***symmetrisch*** $:\Leftrightarrow f^* = f$.
Ein selbstadjungierter Projektor $\Pi \in \operatorname{End}(V)$ *heißt auch* ***orthogonal****.*

NOTIZ 6.21 *Zu jedem Projektor* $\Pi \in \operatorname{End}(V)$ *erhält man wegen* $(\mathbf{1} - \Pi)^2 = \mathbf{1} - \Pi^2 - 2\Pi = (\mathbf{1} - \Pi)$ *mit* $(\mathbf{1} - \Pi)$ *die Projektion auf ein direktes Komplement von* $\operatorname{Bild}\Pi$*. Für einen orthogonalen Projektor gilt damit sogar (in Rechtfertigung des Namens):*
Die direkte Summenzerlegung $V = \operatorname{Bild}\Pi \oplus \operatorname{Bild}(\mathbf{1} - \Pi)$ *ist orthogonal, das heisst*

$$\langle \Pi x, (\mathbf{1} - \Pi)y\rangle = \langle \Pi x, y\rangle - \langle \Pi x, \Pi y\rangle = \langle \Pi x, y\rangle - \langle \Pi^*\Pi x, y\rangle = 0\,, \quad x, y \in V\,,$$

es gilt also tatsächlich $\operatorname{Bild}\Pi \perp \operatorname{Bild}(1 - \Pi)$.

Bemerkung und Beispiel: In einem endlichdimensionalen Vektorraum mit Skalarprodukt $(V, \langle\cdot,\cdot\rangle)$ erhält man die orthogonale Projektion Π_U auf einem Untervektorraum U von V einfach durch Wahl einer Orthonormalbasis $\mathcal{E} = (e_1, ..., e_k)$ von U. Dann ist (in der Bra-Ket-Schreibweise)

$$\Pi_U = \sum_{j=1}^{k} |e_j\rangle\langle e_j|\,, \quad \text{also} \quad \Pi_U x = \sum_j e_j \langle e_j, x\rangle\,,\ x \in V\,.$$

Diese Formel ist eigentlich schon bekannt, da sie gerade die Projektionen von x auf die Koordinatenebene aufsummiert. Formal gilt in BRA-KET-Notation $\big(|a\rangle\langle b|\big)^* = |b\rangle\langle a|$ (lies $\langle x|a\rangle\langle b|y\rangle$ von rechts nach links), also:

$$\begin{aligned}
\Pi^* &= \left(\sum_{j=1}^{k} |e_j\rangle\langle e_j|\right)^* = \sum_{j=1}^{k} |e_j\rangle\langle e_j| = \Pi \\
\Pi^2 &= \left(\sum_{j=1}^{k} |e_j\rangle\langle e_j|\right)\left(\sum_{\ell=1}^{k} |e_\ell\rangle\langle e_\ell|\right) = \sum_{j,\ell=1}^{k} |e_j\rangle\delta_{j\ell}\langle e_\ell| = \sum_j |e_j\rangle\langle e_j| = \Pi
\end{aligned}$$

In ∞-dimensionalen Hilberträumen gelten diese Formeln für abgeschlossene Untervektorräume U mit der Vorsichtsmaßnahme, dass unendliche Summen immer als konvergent bezüglich der Hilbertraumnorm $\|x\| = \sqrt{\langle x, x\rangle}$ interpretiert werden und der Begriff Orthonormalbasis entsprechend weiter gefasst wird.

6.3 Übungen

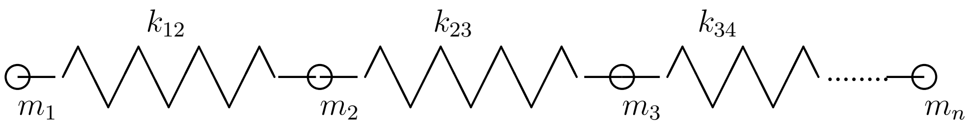
1. Beweisen Sie Notiz 6.2, 1) und 3).
2. Beweisen Sie Satz 6.5.
3. Beweisen Sie Satz 6.11.
4. Beweisen Sie Satz 6.14.
5. Zeigen Sie für V, W Vektorräume und $f \in \mathrm{Hom}(V, W)$, dass gilt $\mathrm{Rang}\, f = \mathrm{Rang}\, f^*$.

Kapitel 7

Das Eigenwertproblem

7.1 Motivation: Lösung linearer Differentialgleichungen

Wir betrachten eine Kette von Oszillatoren, das heißt durch Federn mit den Federkonstanten $k_{12},\ k_{23}, .., k_{(n-1)n}$ gekoppelte Massepunkte mit den Massen $m_1,\ m_2, ..., m_n$.



Sei $x_j(t) \in \mathbb{R}$ die Auslenkung des Massenpunktes $m_j,\ j = 1, ..., n$ aus seiner Ruhelage zum Zeitpunkt $t \in \mathbb{R}$. Die Gesetze von Newton und Hooke liefern (mit der Notation $\frac{d^2}{dt^2}x = \ddot{x}$) das folgende System von Differentialgleichungen:

$$
\begin{aligned}
m_1\ddot{x}_1 &= -k_{12}(x_1 - x_2) \\
m_2\ddot{x}_2 &= -k_{23}(x_2 - x_3) + k_{12}(x_1 - x_2) \\
\vdots &= \vdots \\
m_n\ddot{x}_n &= -k_{(n-1)n}(x_{n-1} - x_n)
\end{aligned}
$$

Das bedeutet in Vektor-Notation:

$$
\mathbb{R}^{\ltimes} \ni \ddot{x} = \begin{pmatrix} \ddot{x}_1 \\ \vdots \\ \ddot{x}_n \end{pmatrix} = \begin{pmatrix} -\frac{k_{12}}{m_1} & +\frac{k_{12}}{m_1} & 0 & \cdots & 0 \\ +\frac{k_{12}}{m_2} & -\left(\frac{k_{12}+k_{23}}{m_2}\right) & +\frac{k_{23}}{m_2} & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} =: Ax
$$

NOTIZ 7.1 *Die Matrix A hat die Struktur* $A = \begin{pmatrix} * & * & 0 & 0 \\ * & \ddots & \ddots & * \\ 0 & \ddots & * & * \end{pmatrix}$, *sie ist also eine sogenannte Bandmatrix (nur auf der Haupt- und den beiden Nebendiagonalen sind Einträge $\neq 0$).*
Für $m_1 = m_2 = ... = m_n,\ k_{ij} = k = const$ ist $A = A^*$.

Zur Lösung der Differentialgleichung $\ddot{x} = Ax$ mit $x \in \mathcal{C}^2(\mathbb{R}, \mathbb{R}^n)$ verwenden wir den Exponentialansatz:

$$x(t) = e^{i\omega t} x_0 \quad \text{mit } x_0 \in \mathbb{R}^n \ .$$

Als reelle Lösung dienen dann Realteil und Imaginärteil von $x(t)$. Damit dieser Ansatz die Differentialgleichung löst, ergibt sich als notwendige Bedingung:

$$\ddot{x}(t) = -\omega^2 \underbrace{e^{i\omega t} x_0}_{x(t)} = A \underbrace{x(t)}_{e^{i\omega t} x_0}$$

Fazit: $x(t) = e^{i\omega t} y$ löst $\ddot{x} = Ax \ \Leftrightarrow\ Ay = \lambda y$ für $\lambda = -\omega^2 \in \mathbb{C}$.
Gilt $Ay = \lambda y$, dann ist $y \in \mathbb{R}^n$ ist ein sogenannter Eigenvektor von A zum Eigenwert $\lambda \in \mathbb{C}$.

Bemerkung: Für unser Differentialgleichungssystem $\ddot{x} = Ax$ liefern die Eigenvektoren $y \in \mathbb{R}^n$ von A gerade bestimmte ausgezeichnete **Normalschwingungen** des Systems, die sogenannten **Schwingungsmoden**. Bei drei Massepunkten ist eine Schwingungsmode die Schwingung der beiden äußeren Massepunkte gegeneinander, der innere Massepunkt bleibt in Ruhelage, bei einer anderen Schwingungsmode schwingt der innere Massepunkt gegen die gleich schwingenden äußeren Massepunkte.

7.2 Grundlegende Eigenschaften

DEFINITION 7.2 *Sei V ein $\mathbb{K}$-Vektorraum. $\lambda \in \mathbb{K}$ heißt **Eigenwert** von $f \in \mathrm{End}(V)$* $:\Leftrightarrow$

$$\exists v \in V, v \neq 0 \quad \textit{mit} \quad f(v) = \lambda v \ .$$

*Jedes solche $v \neq 0$ heißt **Eigenvektor** von f zum Eigenwert λ.*

Warnung: $0 \in \mathbb{K}$ ist als Eigenwert möglich, $v = 0$ als Eigenvektor ist jedoch nicht zugelassen.

NOTIZ 7.3 *Die Eigenvektoren $v_1, ..., v_r \in V$ eines Endomorphismus $f \in \mathrm{End}(V)$ liefern die einfachsten f-**invarianten Teilräume** des $\mathbb{K}$-Vektorraums, für die also gilt $f(U) \subset U$:*

1. *$W := \mathrm{Span}_{\mathbb{K}}(v_1)$ ist f-invariant, denn mit $f(v_1) = \lambda v_1$ folgt aus der Linearität $f(\mu v_1) = \mu f(v_1) = \mu\lambda_1 v_1$ und damit $f(w) \in W$ für alle $w \in W$.*

2. *Analog ist $W = \mathrm{Span}_{\mathbb{K}}(v_1, ..., v_r)$ f-invariant:*

$$f(\mu_1 v_1 + ... + \mu_r v_r) = \mu_1\lambda_1 v_1 + ... + \mu_r\lambda_r v_r \in W$$

Beispiel: Sei $f \in SO(3,\mathbb{R})$ eine Drehung des $\mathbb{R}^3$, $f \neq \mathbf{1}$. Unsere geometrische Anschauung sagt, dass f eine Drehachse $e \in \mathbb{R}^3$ besitzt, welche ein Eigenvektor (zum Eigenwert 1) ist: $fe = e$, denn e ist fix unter der Drehung f. Weitere Eigenvektoren besitzt $f \neq \mathbf{1}$ im Allgemeinen nicht mehr. Als weiterer f-invarianter Unterraum tritt das Orthogonalkomplement $(\operatorname{Span}_{\mathbb{K}} e)^{\perp} \subset \mathbb{R}^3$ der Drehachse auf.

Die Existenz genügend vieler Eigenvektoren gestattet die Zerlegung von V in eine direkte Summe f-invarianter Unterräume und liefert darüberhinaus eine **Diagonalisierung** von $f \in \operatorname{End}(V)$. Die schönste Situation beschreibt

SATZ 7.4 *Sei V ein $\mathbb{K}$-Vektorraum und $f \in \operatorname{End}(V)$. V habe eine Basis $\mathcal{A} = (a_1, ..., a_n)$ aus Eigenvektoren von f zu den Eigenwerten $\lambda_1, ..., \lambda_n \in \mathbb{K}$, die nicht notwendig paarweise verschieden sein müssen (es gilt also $fa_j = \lambda_j a_j,\ j = 1, ..., n$). Dann gilt:*

$$\Phi_{\mathcal{A}}^{\mathcal{A}}(f) = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix} = \operatorname{diag}(\lambda_1, ..., \lambda_n) \tag{7.1}$$

Beweis. Die Spalten sind die Koordinaten der Bilder der Basisvektoren und $fv_j = 0 + ... + 0 + \lambda_j v_j + 0 + ... + 0$.

□

DEFINITION 7.5 *$f \in \operatorname{End}(V)$ heißt **diagonalisierbar**:$\Leftrightarrow$ Es gibt eine Basis $\mathcal{A}$ von V mit*

$$\Phi_{\mathcal{A}}^{\mathcal{A}}(f) = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix}.$$

SATZ 7.6 *$f \in \operatorname{End}(V)$ ist **diagonalisierbar** $\Leftrightarrow$ V zerfällt in eine direkte Summe von eindimensionalen f-invarianten Unterräumen $V_1, ..., V_n$, den sogenannten **Eigenräumen** $V_j = \operatorname{Span} v_j$ zu den Eigenwerten $\lambda_j \in \mathbb{K}$.*

Die Berechnung von Eigenvektoren und Eigenwerten beruht auf den folgenden einfachen Äquivalenzen:

NOTIZ 7.7 *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $f \in \operatorname{End}(V)$. Dann gilt:*
$v \neq 0$ ist ein Eigenvektor von f zum Eigenwert $\lambda \in \mathbb{K}$ (also $fv = \lambda v$)
$\Leftrightarrow (f - \lambda\mathbf{1})v = 0$ für $v \neq 0$
$\Leftrightarrow (f - \lambda\mathbf{1})$ ist nicht injektiv und $v \in \operatorname{Kern}(f - \lambda\mathbf{1})$
$\Leftrightarrow \det(f - \lambda\mathbf{1}) = 0$ und $v \in \operatorname{Kern}(f - \lambda\mathbf{1})$

Die hier aufgeführten Objekte sind so wichtig, dass sie besondere Namen erhalten.

DEFINITION 7.8 *Sei V ein $\mathbb{K}$-Vektorraum und $f \in \operatorname{End}(V)$. Ist $\lambda \in \mathbb{K}$ ein Eigenwert von f, so heißt der Untervektorraum $\operatorname{Kern}(f - \lambda\mathbf{1}) =: \operatorname{Eig}(\lambda) =: E_\lambda$ der **Eigenraum** von f zum Eigenwert λ. $\dim E_\lambda$ heißt **geometrische Vielfachheit** des Eigenwerts λ.*

$$p_f(\lambda) := \det(f - \lambda\mathbf{1}) \tag{7.2}$$

*heißt **charakteristisches Polynom** von f.*

SATZ 7.9 *Sei V ein $\mathbb{K}$-Vektorraum mit $\dim_{\mathbb{K}} V = n$ und $f \in \mathrm{End}(V)$ mit charakteristischem Polynom p_f, dann gibt es $a_0, ..., a_n \in \mathbb{K}$ mit*

$$p_f(\lambda) = (-1)^n \lambda^n + a_{n-1}\lambda^{n-1} + ... + a_1\lambda + a_0. \tag{7.3}$$

Insbesondere gilt $\deg p_f = n = \dim V$ und $a_0 = \det f$ und $a_{n-1} = (-1)^{n-1} \operatorname{Tr} f$.

Beweisidee. Sei $\mathcal{A}$ eine Basis von V, $A = \Phi^{\mathcal{A}}_{\mathcal{A}}(f) = (a_{ij})$ und seien $b_{ij} := a_{ij} - \lambda\delta_{ij}$ die Komponenten von $(A - \lambda\mathbf{1})$. Dann gilt:

$$\begin{aligned}
\det(A-\lambda) &= \sum_{\Pi \in \mathcal{S}_n} (\text{sign } \Pi) b_{1\Pi(1)} \cdots b_{n\Pi(n)} \\
&= (a_{11} - \lambda) \cdots (a_{nn} - \lambda) + \underbrace{\sum_{\Pi \in \mathcal{S}_n \setminus \{\mathbf{1}\}} .....}_{} \\
&\quad \text{jeder Term der Summe enthält } \leq (n-2) \text{ Diagonalelemente} \\
&= (-1)^n \lambda^n + (-1)^{n-1}[a_{11} + ... + a_{nn}]\lambda^{n-1} + \underbrace{Q_1(\lambda) + Q_2(\lambda)}_{\text{grad } Q_j \leq n-2}
\end{aligned}$$

Das Ausmultiplizieren liefert die Existenz der $a_0, ...a_n$. Man liest ab:

$$\begin{aligned}
a_{n-1} &= (-1)^{n-1}[a_{11} + ... + a_{nn}] =: (-1)^{n-1} \operatorname{Tr} \ A \\
a_0 &= \det(f - 0) = \det A
\end{aligned}$$

□

Zur Lösung des **Eigenwertproblems** $fv = \lambda v$ bestimmt man also die **Nullstellen des charakteristischen Polynoms** $p_f(\lambda)$ (vgl. Notiz 7.7).
So erhält man die Eigenwerte $\lambda_1, ..., \lambda_r$. Zu λ_j berechnet man den Eigenraum $E_{\lambda_j} = \mathrm{Kern}(f - \lambda_j \mathbf{1})$ via Gaußalgorithmus, nachdem f durch Basiswahl als eine Matrix $A = \Phi^{\mathcal{A}}_{\mathcal{A}}(f)$ dargestellt wurde. Der Determinanten-Multiplikationssatz 4.30 stellt sicher, dass das charakteristische Polynom $p_f(\lambda)$ basisunabhängig ist.

Warnung: Wir haben hier stillschweigend das Polynom $p_f(\lambda)$ mit der zugehörigen sogenannten Polynomfunktion $p_f(\cdot) : \mathbb{K} \to \mathbb{K}$ identifiziert. Für $\mathbb{K} = \mathbb{R}, \mathbb{C}$ sind diese beiden Begriffe tatsächlich identisch, aber für allgemeine endliche Körper $\mathbb{K}$ müssen wir den Begriff des Polynoms in einem Exkurs präzisieren. Zur Bestimmung der Eigenwerte als Nullstellen des charakteristischen Polynoms reicht der naive Begriff der Polynomfunktion aus.

Als Schritt auf dem Weg zur Diagonalisierung eines Endomorphismus f in einer Basis aus Eigenvektoren zeigen wir zunächst, dass Eigenvektoren zu paarweise verschiedenen Eigenwerten linear unabhängig sind.

LEMMA 7.10 *Sei V ein $\mathbb{K}$-Vektorraum und sind $v_1, ..., v_r \in V$ Eigenvektoren von $f \in \mathrm{End}(V)$ zu Eigenwerten $\lambda_1, ..., \lambda_r \in \mathbb{K}$. Falls für alle $i \neq j$ gilt $\lambda_i \neq \lambda_j$, so sind $(v_1, ..., v_r)$ linear unabhängig.*

Beweis. Induktion nach der Anzahl $r \in \mathbb{N}$:
Induktionsverankerung $r = 1$: Ein Eigenvektor (v_1) ist linear unabhängig, da $v_1 \neq 0$ nach Voraussetzung gilt.
Induktionsschluss $(r) \to (r+1)$:
Seien $(v_1, ..., v_{r+1})$ Eigenvektoren zu Eigenwerten $(\lambda_1, ..., \lambda_{r+1})$ mit $\lambda_i \neq \lambda_j,\ (i \neq j)$. Wir betrachten die Linearkombinationen

$$\begin{aligned} \alpha_1 v_1 + ... + \alpha_{r+1} v_{r+1} &= 0 \\ \Rightarrow \alpha_1 \lambda_1 v_1 + ... + \alpha_{r+1} \lambda_{r+1} v_{r+1} &= 0 \end{aligned} \tag{7.4}$$

Obige Implikation folgt durch das Anwenden von f auf beiden Seiten. Multiplikation der Linearkombination mit λ_{r+1} liefert

$$\lambda_{r+1} \alpha_1 v_1 + ... + \lambda_{r+1} \alpha_{r+1} v_{r+1} = 0.$$

Also resultiert nach Subtraktion beider Gleichungen

$$(\lambda_1 - \lambda_{r+1}) \alpha_1 v_1 + ... + (\lambda_r - \lambda_{r+1}) \alpha_r v_r = 0$$

Nach Voraussetzung ist $(v_1, ..., v_r)$ linear unabhängig. Also gilt:

$$(\lambda_1 - \lambda_{r+1}) \alpha_1 = ... = (\lambda_r - \lambda_{r+1}) \alpha_r = 0$$

$\Rightarrow \alpha_1 = \alpha_2 = ... = \alpha_r = 0$, da $\lambda_i \neq \lambda_j$ für $i \neq j$. Einsetzen in (7.4) liefert $\alpha_{r+1} = 0$. □

KOROLLAR 7.11 *Ein Endomorphismus $f \in \mathrm{End}(V)$ eines n-dimensionalen $\mathbb{K}$-Vektorraums V hat höchstens n verschiedene Eigenwerte. In diesem Fall ist f diagonalisierbar in einer Basis aus den zugehörigen Eigenvektoren.*

Bemerkung: Mit unseren neuen Begriffen können wir auch sagen:
$f \in \mathrm{End}(V)$ ist diagonalisierbar $\Leftrightarrow$ Die Summe der geometrischen Vielfachheiten $\dim E_\lambda$, mit λ Eigenwert von f, ist genau $n = \dim V$.

Warnung: Nicht jedes $f \in \mathrm{End}(V)$ besitzt eine Basis aus Eigenvektoren. Zum Beispiel $f = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ hat die Eigenwerte $\lambda_1 = \lambda_2 = 0$ und einen einzigen Eigenvektor $e = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = e_1 \in \mathbb{R}^2$.

7.3 Eigenräume symmetrischer Endomorphismen: Der Spektralsatz

Sei im Folgenden $(V \langle \cdot, \cdot \rangle)$ ein endlichdimensionaler reeller oder komplexer Vektorraum mit Skalarprodukt. Wir erinnern uns an folgende Sachverhalte:
Zu $f \in \mathrm{End}(V)$ ist der adjungierte Operator f^* eindeutig bestimmt durch

$$\langle fx, y \rangle = \langle x, f^* y \rangle \,, \qquad \forall x, y \in V \,. \tag{7.5}$$

Wir sagen: f ist selbstadjungiert oder symmetrisch $\Leftrightarrow f = f^*$.

LEMMA 7.12 *Sei* $f = f^* \in \mathrm{End}(V)$*, dann sind alle Nullstellen des charakteristischen Polynoms* $p_f(\lambda)$ *reell.*

Beweis. 1. Fall: V unitär
Sei $\lambda \in \mathbb{C}$ mit $p_f(\lambda) = \det(f - \lambda\mathbf{1}) = 0$, also λ Eigenwert von f. Dann existiert $x \in V, x \neq 0$ mit $fx = \lambda x$. Dann gilt aber

$$\lambda\|x\|^2 = \langle x, fx\rangle = \langle f^*x, x\rangle = \overline{\lambda}\|x\|^2$$

und damit $\lambda = \overline{\lambda}$, also $\lambda \in \mathbb{R}$.

2. Fall: V euklidisch
Um aus obigen Schlussketten etwas Interessantes zu erhalten, müssen wir V komplexifizieren, das heißt grob gesprochen, die Multiplikation mit reellen Skalaren zu einer Multiplikation mit komplexen Skalaren zu erweitern. Besonders einfach ist das für eine Matrixdarstellung von f: Sei $\mathcal{A}$ eine Orthonormalbasis von V, dann ist die Matrixdarstellung $\Phi^{\mathcal{A}}_{\mathcal{A}}(f) = A_f$ symmetrisch. Oder genauer: $A_f \in \mathcal{M}(n \times n, \mathbb{R})$ und $A_f^t = A_f$, mit $n = \dim_{\mathbb{R}} V$. Damit ist A_f als Element von $\mathrm{End}(\mathbb{C}^n)$ selbstadjungiert. Somit gilt nach Fall 1:
$p_{A_f}(\lambda)$ hat nur reelle Nullstellen. Da aber $p_f(\lambda) = p_{A_f}(\lambda)$ folgt die Behauptung. □

KOROLLAR 7.13 *Jeder symmetrische Endomorphismus eines euklidischen Vektorraums besitzt mindestens einen Eigenwert* $\lambda \in \mathbb{R}$.

Beweis. Nach dem Fundamentalsatz der Algebra gibt es mindestens eine komplexe Nullstelle $\lambda \in \mathbb{C}$ des charakteristischen Polynoms. Dieses λ ist aber reell nach Lemma 7.12 und damit ein Eigenwert. □

LEMMA 7.14 *Sei* $f = f^* \in \mathrm{End}(V)$ *und* $E \subset V$ *ein* f*-invarianter Untervektorraum, also* $f(E) \subset E$*, dann ist auch* $E^\perp$ f*-invariant.*

Beweis. Sei $y \in E^\perp$, dann gilt $0 = \langle y, fx\rangle \overset{f=f^*}{=} \langle fy, x\rangle$ für alle $x \in E$ und damit $fy \in E^\perp$. □

SATZ 7.15 (SPEKTRALSATZ (1. FORM)) *Sei* $f = f^* \in \mathrm{End}(V)$ *und seien* $\lambda_1, ..., \lambda_r \in \mathbb{R}$ *die paarweise verschiedenen Eigenwerte von* f *und* $\mathrm{Kern}(f - \lambda_j\mathbf{1}) =: E_j$ *die zugehörigen Eigenräume. Dann gilt*

$$E_j \perp E_i \qquad \forall i \neq j,\ i, j \in \{1, ..., r\} \tag{7.6}$$

$$V = \bigoplus_{j=1}^{r} E_j\ . \tag{7.7}$$

KOROLLAR 7.16 *Jeder selbstadjungierte Endomorphismus* f *ist diagonalisierbar, da* V *eine Basis aus Eigenvektoren von* f *besitzt.*

Beweis von Satz 7.15. Gleichung (7.6):
Sei $x \in E_j,\ y \in E_i,\ i \neq j$, dann ist $fx = \lambda_j x$ und $fy = \lambda_i y$. Daraus folgt

$$\lambda_j\langle x, y\rangle = \langle fx, y\rangle = \langle x, fy\rangle = \lambda_i\langle x, y\rangle$$

und damit $(\lambda_j - \lambda_i)\langle x, y\rangle = 0$. Da aber $\lambda_i \neq \lambda_j$ folgt $\langle x, y\rangle = 0$.
Gleichung (7.7):
f besitzt nach Korollar 7.13 mindestens einen Eigenwert $\lambda_1 \in \mathbb{R}$ mit zugehörigem Eigenraum E_1. Nach Lemma 7.14 ist $f|_{E_1^\perp} \in \mathrm{End}(E_1^\perp)$ und dies Einschränkung ist auch selbstadjungiert. Erneute Anwendung von Korollar 7.13 liefert mindestens einen Eigenwert $\lambda_2 \in \mathbb{R},\ \lambda_2 \neq \lambda_1$ von $f|_{E_1^\perp}$. Die Abspaltung des zugehörigen Eigenraums E_2 liefert wiederum $f|_{(E_1 \oplus E_2)^\perp} \in \mathrm{End}\big((E_1 \oplus E_2)^\perp\big)$ selbstadjungiert mit mindestens einem Eigenwert $\lambda_3 \in \mathbb{R}$ usw. Da $\dim\ V < \infty$, bricht das Verfahren ab und es folgt $V = \bigoplus_{j=1}^r E_j$.

$\square$

SATZ 7.17 (SPEKTRALSATZ (2. FORM)) *Sei $f = f^* \in \mathrm{End}(V)$ und $\lambda_1, ..., \lambda_r \in \mathbb{R}$ die paarweise verschiedenen Eigenwerte von f. Seien Π_j die orthogonalen Projektoren auf die zugehörigen Eigenräume $E_{\lambda_j} = \mathrm{Kern}(f - \lambda_j \mathbf{1}),\ j = 1, ..., r$. Dann gilt*

$$\Pi_i \Pi_j = \Pi_j \Pi_i \;=\; \delta_{ij}\Pi_i \ (i, j \in \{1, \dots, r\}) \qquad \textit{und} \qquad \sum_{j=1}^r \Pi_j = \mathbf{1} \tag{7.8}$$

$$f \;=\; \sum_{j=1}^r \lambda_j \Pi_j \tag{7.9}$$

Beweis. Gleichung (7.8):
$\Pi_i \Pi_j = 0\ (i \neq j)$ folgt aus $E_i \perp E_j$ (Satz 7.15). Weiterhin ist $\sum_{j=1}^r \Pi_j$ der orthogonale Projektor auf $E_1 \oplus E_2 \oplus ... \oplus E_r = V$.

Gleichung (7.9):
Jedes $x \in V$ besitzt die mit $x_j = \Pi_j x \in E_j$ die (eindeutige) Zerlegung $x = (\sum_{j=1}^r \Pi_j)x = \sum_{j=1}^r x_j$. Da $f(x_j) = \lambda_j x_j$ gilt für alle $x \in V$:

$$f(x) = \sum_j f(x_j) = \sum_j \lambda_j x_j = \Big(\sum_j \lambda_j \Pi_j\Big) x\ .$$

$\square$

Mit diesen Resultaten lässt sich nun einfach der sogenannte Funktionalkalkül für selbstadjungierte Operatoren eines endlichdimensionalen euklidischen oder unitären Vektorraums V skizzieren.

7.3.1 Funktionalkalkül für $f = f^* \in \mathrm{End}(V)$

Einsetzen von f in die Monome $p_k(t) = t^k$ führt mit (7.8) und (7.9) auf die selbstadjungierten Operatoren

$$\begin{aligned}
p_0(f) = \mathbf{1} &= \sum \Pi_j \\
p_1(f) = f &= \sum \lambda_j \Pi_j \\
p_2(f) = f^2 &= \Big(\sum_{j=1}^r \lambda_j \Pi_j\Big)\Big(\sum_{k=1}^r \lambda_k \Pi_k\Big) = \sum_{j,k=1}^r \lambda_j \lambda_k \Pi_j \Pi_k = \sum_{j=1}^r \lambda_j^2 \Pi_j
\end{aligned}$$

und durch Induktion nach n erhalten wir $\forall n \in \mathbb{N}$:

$$f^n = \underbrace{f \circ ... \circ f}_{nMal} = \sum \lambda_j^n \Pi_j \; -$$

Damit ergeben sich für beliebige Polynome $q(t) = \sum_{k=0}^{m} a_k t^k \in \mathbb{C}[t]$ durch Einsetzen von f die selbstadjungierten Operatoren $q(f) \in \mathrm{End}(V)$ mit

$$q(f) = \sum_{k=0}^{m} a_k f^k = \sum_{k=0}^{m} a_k \sum_{j=1}^{r} \lambda_j^k \Pi_j = \sum_{j=1}^{r} q(\lambda_j) \Pi_j \; .$$

Dies lässt sich verallgemeinern auf absolut konvergente Potenzreihen $s(t) = \sum_{k=0}^{\infty} a_k t^k$ mit $a_k \in \mathbb{C}$:

$$s(f) = \sum_{k=0}^{\infty} a_k \sum_{j=1}^{r} \lambda_j^k \Pi_j = \sum_{j=1}^{r} s(\lambda_j) \Pi_j \; . \tag{7.10}$$

Der selbstadjungierte Operator $s(f) \in \mathrm{End}(V)$ ist als Limes der Partialsummen

$$s_N(f) = \sum_{k=0}^{N} a_k \sum_{j=1}^{r} \lambda_j^k \Pi_j$$

bezüglich der Operatornorm $\| \cdot \|$ definiert. $s(\lambda_j)$ ist der Limes der Partialsummen $s_N(\lambda_j) = \sum_{k=0}^{N} a_k \lambda_j^k$ im komplexen Betrag $| \cdot |$.

Diese Resultate legen nun nahe, für allgemeine Funktionen das "Einsetzen eines selbstadjungierten Endomorphismus" durch völlig analoge Formeln zu definieren.

DEFINITION 7.18 *Für $f \in \mathrm{End}(V)$ heißt die Menge der Eigenwerte auch das* ***Spektrum*** *$\sigma(f)$ von f. Für jede Funktion $q : \ \sigma(f) \to \mathbb{C}$ auf dem Spektrum von f definiert man im Fall $f = f^* = \sum_j \lambda_j \Pi_j$:*

$$q(f) := \sum_j q(\lambda_j) \Pi_j \in \mathrm{End}(V) \tag{7.11}$$

NOTIZ 7.19 *Für absolut konvergente Potenzreihen, das heißt analytische Funktionen q auf $\sigma(f)$, stimmt dies mit der alternativen Definition* (7.10) *durch Einsetzen in die absolut konvergente Reihe überein. Für stetige Funktionen $q \in \mathcal{C}\,(\sigma(f); \mathbb{C})$ werden wir spätestens in der Funktionalanalysis des vierten Semesters sehen, dass* (7.11) *mit der Approximation von q durch Polynome in der sup-Norm $\| \cdot \|_\infty$ verträglich ist.*

Alle Resultate der vorangegangenen Abschnitte benutzen wesentlich $\dim \ V < \infty$. *Im unendlich-dimensionalen Fall ist es notwendig, den Begriff des Spektrums eines selbstadjungierten Operators $f \in \mathrm{End}(V)$ auf mehr als Eigenwerte zu verallgemeinern. Es kann zusätzlich ein sogenanntes* ***kontinuierliches Spektrum*** *auftreten. Eine verallgemeinerte Spektralzerlegung eines selbstadjungierten Operators f in einem ∞-dimensionalen Hilbertraum und einen analogen Funktionalkalkül werden wir in der Funktionalanalysis kennenlernen. Sie ist wesentlich für das Verständnis der Quantentheorie.*

7.4 Der Spektralsatz für normale Endomorphismen

DEFINITION 7.20 *Sei* $(V, \langle\cdot,\cdot\rangle)$ *ein unitärer bzw. euklidischer Vektorraum.* $f \in \mathrm{End}(V)$ *heißt*

1. ***unitär*** *(im Fall* $\mathbb{K} = \mathbb{C}$*)* $:\Leftrightarrow f^* f = \mathbf{1}$ *(oder äquivalent* $\langle fx, fy\rangle = \langle x, y\rangle$ *für alle* $x, y \in V$*).*
2. ***orthogonal*** *(im Fall* $\mathbb{K} = \mathbb{R}$*)* $:\Leftrightarrow f^t f = \mathbf{1}$ *(oder äquivalent* $\langle fx, fy\rangle = \langle x, y\rangle$ *für alle* $x, y \in V$*).*
3. ***normal*** $:\Leftrightarrow [f, f^*] := ff^* - f^* f = 0$.

Beispiele: Der Endomorphismus $f - f^*$ ist normal, ebenso sind die selbstadjungierten und die unitären bzw. orthogonalen Abbildungen immer normal.

Warnung: In einem reellen Vektorraum V gilt $[f, f^*] = 0$ für orthogonale Abbildungen f. Diese brauchen aber keinen reellen Eigenwert zu besitzen (im Gegensatz zu den symmetrischen Abbildungen). Die im Allgemeinen komplexen Nullstellen des charakteristischen Polynoms $p_f(\lambda)$ sind streng genommen nur Eigenwerte für die von f erzeugte lineare Abbildung in der Komplexifizierung des reellen Vektorraums V. Daher machen wir die folgende Betrachtung gleich in einem komplexen Vektorraum mit Skalarprodukt.

DEFINITION 7.21 *Sei* V *ein unitärer Vektorraum.* $f \in \mathrm{End}(V)$ *besitzt eine* ***Spektralzerlegung***$:\Leftrightarrow$ *Es gibt orthogonale Projektoren* Π_j *und* $\lambda_j \in \mathbb{C},\ j = 1, ..., r, ,$ *mit:*

$$\begin{aligned} \sum_{j=1}^{r} \Pi_j &= \mathbf{1}\,, \qquad \Pi_i \Pi_j = \Pi_j \Pi_i = \delta_{ij} \Pi_j\,, \quad 1 \le i, j \le r, \\ f &= \sum_{j=1}^{r} \lambda_j \Pi_j\,. \end{aligned}$$

SATZ 7.22 *Sei V ein unitärer Vektorraum und* $f \in \mathrm{End}(V)$*, dann besitzt* f *eine Spektralzerlegung genau dann, wenn* f *normal ist.*

Beweis. "$\Rightarrow$"

$$\begin{aligned} ff^* &= \left(\sum \lambda_j \Pi_i\right)\left(\sum \lambda_i \Pi_i\right)^* = \left(\sum_j \lambda_j \Pi_j\right)\left(\sum_i \overline{\lambda_i} \Pi_i\right) \\ &= \sum_{i,j} \lambda_j \overline{\lambda_i} \delta_{ij} \Pi_i = \sum_j |\lambda_j|^2 \Pi_j = ... = f^* f \end{aligned} \tag{7.12}$$

"$\Leftarrow$"
Es reicht, den Spektralsatz in der 1. Form für normale (statt selbstadjungierte) Abbildungen zu zeigen. Das geht weitestgehend analog zum selbstadjungierten Fall (Satz 7.15) mit folgenden Lemmata.

LEMMA 7.23 *Sei $f \in \mathrm{End}(V)$ normal und seien $x \in V, x \neq 0$ und $\lambda \in \mathbb{C}$ mit $fx = \lambda x$, dann gilt: $f^*x = \overline{\lambda}x$.*

Beweis des Lemmas.

$$\|fx\|^2 = \langle fx, fx\rangle = \langle f^*fx, x\rangle \overset{f \text{ normal}}{=} \langle ff^*x, x\rangle = \langle f^*x, f^*x\rangle = \|f^*x\|^2$$

Nun ist $(f-\lambda)$ normal, da f normal ist und $(f-\lambda)^* = f^* - \overline{\lambda}$. Also gilt $\forall y \in V$: $\|(f-\lambda)y\| = \|(f^* - \overline{\lambda})y\|$ und damit $\|(f^* - \overline{\lambda})x\| = 0$ und also $(f^* - \overline{\lambda})x = 0$ für den Eigenvektor x.

□

LEMMA 7.24 *Seien $\lambda_1, ..., \lambda_r \in \mathbb{C}$ die paarweise verschiedenen Eigenwerte von f und $E_1, ..., E_r$ die zugehörigen Eigenräume. Dann gilt $E_i \perp E_j$ für alle $i \neq j, 1 \leq i, j \leq n$.*

Beweis des Lemmas. Mit $fx = \lambda_i x,\ fy = \lambda_j y$ folgt aus Lemma 7.23

$$\lambda_i\langle x, y\rangle = \langle \overline{\lambda_i}x, y\rangle = \langle f^*x, y\rangle = \langle x, fy\rangle = \lambda_j\langle x, y\rangle\ .$$

Also folgt aus $\lambda_j \neq \lambda_i$ sofort $x \perp y$. □

LEMMA 7.25 *Ein normaler Endomorphismus $f \in \mathrm{End}(V)$ besitzt mindestens einen Eigenwert $\lambda_1 \in \mathbb{C}$ mit zugehörigem Eigenraum E_1. Die Unterräume E_1 und $E_1^\perp$ sind f-invariant.*

Beweis des Lemmas. Die Existenz des Eigenwertes folgt aus dem Fundamentalsatz der Algebra, die Invarianz von E_1 folgt sofort aus der Eigenwertgleichung $fx = \lambda_1 x$ für $x \in E_1$. Sei $x \in E_1$, dann gilt für alle $y \in E_1^\perp$:

$$\langle fy, x\rangle = \langle y, f^*x\rangle = \langle y, \overline{\lambda_1}x\rangle = 0\,,$$

also $fy \in E_1^\perp$. □

Aus Lemma 7.25 folgt induktiv, wie bei selbstadjungierten Abbildungen (Beweis Satz 7.15), die direkte Summenzerlegung $V = \bigoplus_{j=1}^r E_j$. Zusammen mit Lemma 7.24 liefert das eine Spektralzerlegung für jeden normalen Endomorphismus f. □

Bemerkung: Kombiniert man den Funktionalkalkül für selbstadjungierte Endomorphismen mit der Spektralzerlegung eines normalen $f \in \mathrm{End}(V)$, so findet man, dass es zu jedem normalen $f \in \mathrm{End}(V)$ einen selbstadjungierten $T \in \mathrm{End}(V)$ und eine Funktion $q : \sigma(T) \to \mathbb{C}$ gibt mit $q(T) = f$. Normale Endomorphismen sind also genau die Funktionen von selbstadjungierten Endomorphismen. Hierdurch kann man alternativ die Menge der normalen Endomorphismen beschreiben.

SATZ 7.26 *Sei V ein unitärer Vektorraum.*

a) Ein normaler Endomorphismus $U \in \mathrm{End}(V)$ ist genau dann unitär, wenn $\sigma(U) \subset S^1 := \{\lambda \in \mathbb{C}; |\lambda| = 1\}$ ist.

b) Ein normales $A \in \mathrm{End}(V)$ ist genau dann selbstadjungiert, wenn $\sigma(A) \subset \mathbb{R}$ gilt.

c) Zu jedem unitären $U \in \mathrm{End}(V)$ gibt es ein selbstadjungiertes $A \in \mathrm{End}(V)$ mit $U = \exp(iA)$.

Beweis. ad a) Sei $U = \sum_j \lambda_j \Pi_j$ die Spektralzerlegung des normalen $U \in \mathrm{End}(V)$. Dann gilt

$$\begin{aligned} \mathbf{1} = UU^* &= \left(\sum_j \lambda_j \Pi_j\right)\left(\sum_k \overline{\lambda_k}\Pi_k\right) = \sum_{i,j} \lambda_j \overline{\lambda_k}\delta_{jk}\Pi_k \\ &= \sum |\lambda_j|^2 \Pi_j \Leftrightarrow |\lambda_j| = \mathbf{1}\,, \quad \forall j \end{aligned}$$

Die λ_j sind aber gerade die Elemente von $\sigma(U)$.

ad b) Sei $A = \sum \lambda_j \Pi_j$ die Spektralzerlegung von A, dann gilt

$$A = A^* \Leftrightarrow \sum_j (\lambda_j - \overline{\lambda_j})\Pi_j = 0 \Leftrightarrow \lambda_j \in \mathbb{R}\,,\ \forall j\,.$$

ad c) Sei $U = \sum \lambda_j \Pi_j$ die Spektralzerlegung eines unitären $U \in \mathrm{End}(V)$. Nach a) gilt dann $\lambda_j \in S^1$, also $\lambda_j = \exp(i\phi_j)$ für $\phi_j \in \mathbb{R},\ j = 1, ..., r$.
Unter Beibehaltung der selbstadjungierten Spektralprojektoren Π_j folgt $A := \sum_j \phi_j \Pi_j \in \mathrm{End}(V)$ ist selbstadjungiert, da $A^* = \sum_j \overline{\phi}_j \Pi_j^* = A$, und mit dem Funktionalkalkül folgt

$$e^{iA} = \sum_j e^{i\phi_j}\Pi_j = \sum \lambda_j \Pi_j = U\,.$$

□

Kapitel 8

Die Jordan- Normalform

8.1 Motivation und Ziel

Die Matrix $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathrm{End}(\mathbb{C}^2)$ hat das charakteristische Polynom

$$p_A(\lambda) = \det(A - \lambda\mathbf{1}) = (1-\lambda)^2 \quad \Rightarrow \quad \sigma(A) = \{1\}\,,$$

die Matrix hat also nur den Eigenwert 1. Der zugehörige Eigenvektor x ergibt sich mit

$$(A - \mathbf{1})x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \quad \text{zu} \quad x = e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Damit ist die geometrische Vielfachheit des Eigenwertes 1 gerade $\dim E_1 = \dim \mathrm{Kern}(A-1) = 1$ und damit kleiner als seine algebraische Vielfachheit (also die Vielfachheit der Nullstelle im charakteristischen Polynom), die 2 ist.

Beachte: Die Matrix $N = A - \mathbf{1} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ ist **nilpotent**, d.h. $N^k = 0$ für ein $k < \infty$. Der **Nilpotenzgrad** ist der kleinste Exponent $n \in \mathbb{N}$ mit $N^n = 0$. (Hier ist $N^2 = 0$, der Nilpotenzgrad von N ist also 2).

Wir haben eine Summenzerlegung:

$$A = A_s + A_n = \mathbf{1} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

A_s ist dabei diagonal (i.Allg. diagonalisierbar=semi-simple) und A_n ist nilpotent.

Ziel: Finde für $f \in \mathrm{End}(V)$ eine Zerlegung $f = f_s + f_n$ in einen diagonalisierbaren Anteil f_s und einen nilpotenten Anteil f_n. Wir vermuten die zur Diagonalisierung fehlenden Vektoren im Untervektorraum $(f - \lambda\mathbf{1})^k, k = 2, 3, ...$ für $\lambda \in \sigma(f)$.

In unserem obigen Beispiel gilt $Ne_1 = 0$ und $Ne_2 = e_1$, also $N^2 e_2 = 0$, damit wäre der fehlende Vektor also gerade der 2. Standardvektor e_2. Dabei fassen wir $\mathbb{C}^2 = \mathrm{Span}(e_1, e_2)$ als **verallgemeinerten Eigenraum** E_1 von A zum Eigenwert 1 auf:

$$\mathbb{C}^2 = E_1 := \bigcup_{k\in\mathbb{N}} \mathrm{Kern}(f - \mathbf{1})^k \ .$$

DEFINITION 8.1 *1) Eine Matrix*

$$J_r(\lambda) := \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix} \in \mathcal{M}(r \times r; \mathbb{K}) \tag{8.1}$$

heißt ***r-Jordan-Block*** *mit dem Eigenwert* $\lambda \in \mathbb{K}$.

2) Eine ***Jordan-Matrix*** *besteht aus Jordanblöcken längs der Diagonalen.*

$$J = \begin{pmatrix} \boxed{J_{r_1}(\lambda_1)} & 0 & \cdots & 0 \\ 0 & \boxed{J_{r_2}(\lambda_2)} & & 0 \\ \vdots & & \ddots & 0 \\ 0 & 0 & 0 & \boxed{J_{r_l}(\lambda_l)} \end{pmatrix} \tag{8.2}$$

Dabei sind $\lambda_1, ..., \lambda_l$ *nicht notwendig paarweise verschieden.*

3) Eine Basis $\mathcal{A}$ *des* $\mathbb{K}$*-Vektorraums* V *heißt* ***Jordan-Basis*** *für* $f \in \mathrm{End}(V) :\Leftrightarrow \Phi^{\mathcal{A}}_{\mathcal{A}}(f)$ *ist eine Jordan-Matrix.*

NOTIZ 8.2 *Für* $f = A \in \mathcal{M}(n \times n, \mathbb{K})$ *sind die folgenden Aussagen äquivalent:*

- *Es gibt eine Jordan-Basis* $\mathcal{A}$ *von* $V = \mathbb{K}^n$ *für* A.
- *Es gibt ein* $X \in \mathrm{GL}(n, \mathbb{K})$*, so dass* XAX^{-1} *eine Jordan-Matrix ist.*

Präziser für unser Ziel ist der folgende

SATZ 8.3 *Sei* V *ein endlichdimensionaler* $\mathbb{K}$*-Vektorraum und* $\mathbb{K}$ *algebraisch abgeschlossen, dann gibt es für jedes* $f \in \mathrm{End}(V)$ *eine Jordan-Basis* $\mathcal{A}$ *von* V.

Bemerkung: Die algebraische Abgeschlossenheit von $\mathbb{K}$ kann durch schwächere Bedingungen an $\mathbb{K}$ ersetzt werden. Für die speziellen Endomorphismen f, für die bei beliebigem $\mathbb{K}$ das charakteristische Polynom $p_f(\lambda)$ vollständig in Linearfaktoren zerfällt (d.h. $p_f(t) = (t-\lambda_1)^{r_1} \cdots (t-\lambda_k)^{r_k}$

für $\lambda_j \in \mathbb{K}$, $r_j \in \mathbb{N}$), benötigt man überhaupt keine Einschränkung an $\mathbb{K}$. In diesem Fall kann man den im Folgenden gegebenen Beweis verallgemeinern.

Beispiel:
Sei $V = \mathrm{Span}_{\mathbb{C}}\{p(x)e^{\lambda x} \,|\, p(\cdot) \in \mathbb{C}[x]$ ist ein Polynom mit $\mathrm{grad} p \leq (n-1)\}$ für ein festes $\lambda \in \mathbb{C}$, dann ist $\dim\ V = n$ und

$$\partial_x \left(p(x)e^{\lambda x}\right) = e^{\lambda x} \underbrace{(\lambda p(x) + p'(x))}_{\text{Polynom mit grad}\leq\text{grad } p} ,$$

insbesondere gilt also $\partial_x \in \mathrm{End}(V)$. Sei nun $a_j(x) := \frac{x^j}{j!}e^{\lambda x}$, $j = 1, ..., n-1$, dann bildet $\mathcal{A} := (a_o, ..., a_{n-1})$ eine Basis von V und es gilt

$$\partial_x a_j(x) = e^{\lambda x}\left\{\lambda \frac{x^j}{j!} + \frac{x^{j-1}}{(j-1)!}\right\} = \lambda a_j(x) + a_{j-1}(x) \, .$$

Daraus folgt, dass die Matrixdarstellung von ∂_x in der Basis $\mathcal{A}$ gegeben ist durch

$$\Phi_{\mathcal{A}}^{\mathcal{A}}(\partial_x) = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix} ,$$

also ist $\mathcal{A}$ eine Jordan-Basis für ∂_x.

Wir werden diese spezielle Basis aus sogenannten Quasi-Polynomen beim Lösen von Differentialgleichungen wiederfinden.

Der Beweis der Jordanzerlegung (Satz 8.3) ist lang und zerfällt in die folgenden Schritte:

1) Bestimme eine Normalform für nilpotente Endomorphismen.

2) Zerlege V in f-invariante Untervektorräume assoziiert zu $\lambda \in \sigma(f)$, die sogenannten algebraischen Eigenräume.

3) Kombiniere 1) und 2) um die verfeinerte Zerlegung in Jordan-Blöcke zu erhalten.

8.2 Die Klassifikation nilpotenter Endomorphismen

DEFINITION 8.4 *Sei V ein $\mathbb{K}$-Vektorraum. Ein r-dimensionaler Untervektorraum $W \subset V$ heißt* ***zyklisch*** *für $f \in \mathrm{End}(V) :\Leftrightarrow$ Es gibt einen Vektor $e \in W$ so dass $\mathcal{A} = \{e, fe, ..., f^{r-1}e\}$ eine Basis von W ist und $f^r e = 0$ gilt.*

Eine solche Basis $\mathcal{A}$ heißt auch ***zyklisch*** *für f. e heißt* ***erzeugender Vektor*** *bzw.* ***zyklischer Vektor*** *des f-invarianten Untervektorraums W (bzw. der Basis $\mathcal{A}$).*

Ein zyklischer Untervektorraum von f ist automatisch f-invariant!

NOTIZ 8.5 *Sei* $V = W_1 \oplus ... \oplus W_l$, *mit* W_j *zyklisch für* $f \in \mathrm{End}(V)$. *In einer zyklischen Basis* $\mathcal{A}_j = \{a_1, \ldots, a_{r_j}\}$ *von* W_j *gilt mit* $a_1 = f^{r-1}e_j,\ a_2 = f^{r-2}e, ...$

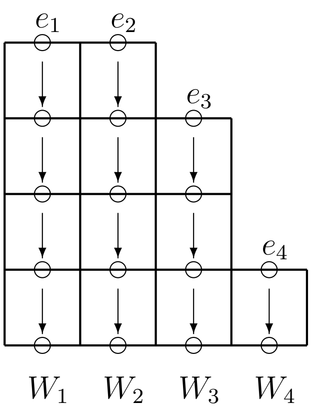
$$\left.\begin{array}{c} fa_1 = 0 \\ fa_2 = a_1 \\ \vdots \end{array}\right\} \Rightarrow \Phi^{\mathcal{A}_j}_{\mathcal{A}_j}(f|_{W_j}) = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = J_{r_j}(0)$$

mit $r_j = \dim W_j$
Also: In der Basis $\mathcal{A} = \{\mathcal{A}_1, ..., \mathcal{A}_l\}$ *von* V *zerfällt* f *in Jordan-Kästchen entsprechend der Zerlegung in zyklische Summanden:*

$$\Phi^{\mathcal{A}}_{\mathcal{A}}(f) = \begin{pmatrix} J_{r_1}(0) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & J_{r_l}(0) \end{pmatrix}$$

Diese Matrix hat Jordan-Normalform. Umgekehrt liefert jede Jordan-Basis $\mathcal{A}$ *von* V *für ein nilpotentes* f *eine Zerlegung* $V = \bigoplus_j W_j$ *in zyklische Summanden* W_j. *Dabei entspricht ein zyklischer Summand gerade einem Jordan-Kästchen* J_{r_j} *zum Eigenwert* $\lambda = 0$.

Besonders übersichtlich wird die Aktion von f auf der Basis $\mathcal{A} = \{\mathcal{A}_1, ..., \mathcal{A}_l\}$ aus den zyklischen Basen $\mathcal{A}_j$ (mit erzeugenden Vektoren e_j) der zyklischen Summanden $W_j \subset V$ im folgenden **Young-Tableau**



$$\begin{array}{c} \mathrm{Kern}\, f^5 \\ \cup \\ \mathrm{Kern}\, f^4 \\ \cup \\ \mathrm{Kern}\, f^3 \\ \cup \\ \mathrm{Kern}\, f^2 \\ \cup \\ \mathrm{Kern}\, f \end{array}$$

Hier bezeichnen die blobs
Elemente der Basis $\mathcal{A}$,
und $\downarrow$ steht für die Aktion von f.

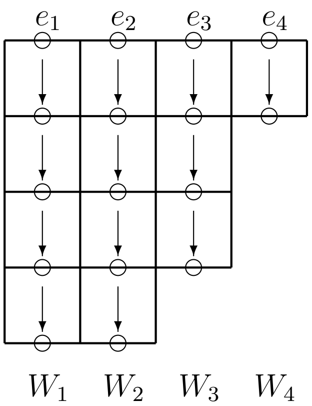
Jeder vertikale Turm steht für einen zyklischen Summanden W_j (genauer für eine zyklische Basis $\mathcal{A}_j$) in der Zerlegung $V = \bigoplus_j W_j$ mit dem zyklischen Vektor e_j on top.

NOTIZ 8.6 *Jedes Young-Tableau mit n blops beschreibt eine sogenannte Partition von* $n \in \mathbb{N}^*$, *das heißt eine Summenzerlegung* $n = r_1 + \ldots + r_\ell\,;\ r_1, ...r_\ell \in \mathbb{N}^*$. *Dabei ist* r_j *die Anzahl der blops im Turm* W_j.

Offensichtlich liegen die blops der Basiszeile in $\mathrm{Kern}\ f$, denn nach der Definition 8.4 der zyklischen Basis gehen sie unter erneuter Anwendung von f in die Null. Sie sind sicher linear unabhängig und man sieht leicht:

NOTIZ 8.7 *Die blops der Basiszeile spannen* $\mathrm{Kern}\ f$ *auf, das heißt sie bilden eine Basis von* $\mathrm{Kern}\ f$.

Um eine Jordan-Basis für ein nilpotentes f zu finden, verschieben wir die Kästchen im Young-Tableau folgendermaßen:



$$\begin{array}{c} V \\ \cup \\ \mathrm{Bild}\, f \\ \cup \\ \mathrm{Bild}\, f^2 \\ \cup \\ \mathrm{Bild}\, f^3 \\ \cup \\ \mathrm{Bild}\, f^4 \end{array}$$

Natürlich wissen wir noch nicht, dass f überhaupt eine Jordan-Basis besitzt. Wir werden jetzt aber schrittweise das nebenstehende Diagramm konstruieren. Das liefert gerade die Existenz einer Jordan-Basis.

SATZ 8.8 *Sei* V *ein* $\mathbb{K}$*-Vektorraum mit* $\dim\ V = n$ *und* $f \in \mathrm{End}(V)$ *nilpotent mit Nilpotenzgrad* $r+1$ *für* $r \in \mathbb{N}$. *Dann gilt:*

1) Es gibt eine Zerlegung $V = W_1 \oplus ... \oplus W_\ell$ *in für* f *zyklische Summanden* W_j. *Bezüglich zyklischer Basen* $\mathcal{A}_j$ *von* W_j *ist* f *durch ein Jordan-Kästchen* $J_{r_j}(0)$, $r_j = \dim\ W_j$, *repräsentiert.*

2) Die Anzahl ℓ *dieser Unterverktorräume* W_j *und ihre Dimension* r_j *mit* $r = r_1 \geq r_2 \geq ... \geq r_\ell$ *sind durch* f *eindeutig bestimmt. Diese Daten bestimmen umkehrbar eindeutig eine Partition* $n = r_1 + ... + r_\ell$.

3) Zwei nilpotente Endomorphismen $f, \tilde{f} \in \mathrm{End}(V)$ *sind konjugiert, das heißt es existiert ein* $S \in \mathrm{GL}(n, \mathbb{K})$ *mit* $f = S\tilde{f}S^{-1}$, *genau dann wenn die zu* f *und* $\tilde{f}$ *gehörenden Partitionen gleich sind.*

Beweisskizze

ad 1) Da $r+1$ der Nilpotenzgrad von f ist, gilt $f^{r+1} = 0$, aber $f^r \neq 0$. Falls $r = 0$ (der Nilpotenzgrad also 1), dann ist $f = 0$ und damit $\mathrm{Kern}\, f = V = W_1 \oplus \ldots \oplus W_n$, die Unterräume W_j bestehen jeweils nur aus einem Element von V.

Sei nun $r \geq 1$, dann wählen wir eine Basis $\mathcal{B}_r := \{x_1^1, x_2^1, ..., x_{p_1}^1\}$ von $\mathrm{Bild}\ f^r$. Zu jedem $x_j^1 \in \mathcal{B}_r$ gibt es ein Urbild $x_j^r \in V$ unter f^r (den zugehörigen zyklischen Vektor), es gilt also $f^r(x_j^r) = x_j^1$, $j = 1, ..., p_1$. Wir setzen

$$x_j^2 := f^{r-1}(x_j^r), \quad j = 1, ..., p_1$$

Insbesondere gilt dann $f(x_j^2) = x_j^1$. Damit haben wir die zwei untersten Stufen der längsten zyklischen Spalten konstruiert. Um daraus eine Basis $\mathcal{B}_{r-1}$ von $\mathrm{Bild}\ f^{r-1}$ zu erhalten, müssen wir aber noch zunächst verifizieren, dass die Vektoren $\{x_j^1, x_j^2\}_{1\leq j\leq p_1}$ tatsächlich linear unabhängig sind:

$$\begin{aligned}\sum \lambda_j x_j^1 + \sum \mu_j x_j^2 &= 0 \overset{\text{Anwenden von } f}{\Rightarrow} \sum \mu_j x_j^1 = 0 \overset{x_j^1 \text{ Basis von Bild } f}{\Rightarrow} \mu_j = 0,\ \forall 1 \leq j \leq p_1 \\ \Rightarrow \quad \sum \lambda_j x_j^1 &= 0 \quad \Rightarrow \quad \lambda_j = 0 = \mu_j,\ \forall j = 1, ..., p_1\end{aligned}$$

Erweitere nun $\{x_j^k\}_{1\leq j\leq p_1;\ k=1,2}$ zu einer Basis $\mathcal{B}_{r-1}$ von $\mathrm{Bild}\ f^{r-1}$ durch Hinzunahme von $x_{p_1+1}^2$, . . ., $x_{p_2}^2 \in V$ mit $f(x_j^2) = 0$, $j = p_1 + 1, \ldots p_2$. Das liefert jetzt auch noch die fehlenden Elemente der Zeile zwei von unten im obigen Young-Tableau. Diese werden nun wiederum von Elementen $x_j^r,\ j = p_1 + 1, \ldots p_2$ on top erzeugt (d.h. sind Bilder von diesen Elementen unter f^{r-1}). Die Iteration dieses Verfahrens liefert schließlich eine Basis

$$\mathcal{B} = \{x_1^1, \ldots, x_{p_1}^1, x_1^2, \ldots x_{p_2}^2, x_1^3, \ldots\ldots, x_{p_r}^r\}$$

von V und es gilt für alle $k = 1, ..., r$ und $j = 1, ..., p_k$ mit $p_1 \leq p_2 \leq ... \leq p_r$:

$$fx_j^k = \begin{cases} x_j^{k-1} & 1 \leq j \leq p_{k-1} \\ 0 & p_{k-1} < j \leq p_k \end{cases}$$

Das ist aber gerade das Young-Tableau. Ordnet man die x_j^k spaltenweise hintereinander an, kann man aus dem Diagramm die zyklischen Summanden W_j ablesen. Offensichtlich bilden die $\{x_j^k\}$ gerade eine Jordan-Basis dieser W_j.

ad 2) Bei der unter 1) ausgeführten Konstruktion des Young-Tableaus von unten sind jeweils die Längen der Zeilen $p_1 = \dim\ \mathrm{Bild}\ f^r$, $p_2 = \dim\ \mathrm{Bild}\ f^{r-1}$ etc. eindeutig durch f bestimmt. Damit ist die Anzahl $p_r = \ell$ der Spalten (also die Anzahl der zyklischen Summanden) eindeutig bestimmt und es ergeben sich ebenso in eindeutiger Weise die Längen der Spalten und damit die Höhe des Young-Tableaus, da das Young-Tableau eine Partition von $n = \dim\ V$ beschreibt.

ad 3) "⇒": Seien $f, \tilde{f}$ konjugiert, dann existiert $S \in \mathrm{GL}(n, \mathbb{K})$ mit $\tilde{f} = SfS^{-1}$ und daraus folgt direkt $\dim\ \mathrm{Kern}\ f^j = \dim\ \mathrm{Kern}\ \tilde{f}^j$, also Part($\tilde{f}$) =Part($f$).
"⇐": Falls Part($\tilde{f}$) =Part(f), dann gilt in geeigneten Basen $\mathcal{A}, \tilde{\mathcal{A}}$ von V: $\Phi_{\mathcal{A}}^{\mathcal{A}}(f) = \Phi_{\tilde{\mathcal{A}}}^{\tilde{\mathcal{A}}}(\tilde{f})$, also ist f konjugiert zu $\tilde{f}$.

Aus Satz 8.8, 3, folgt sofort, dass es genau $p(n) = \#$ Part$\{n\}$ Konjugationsklassen nilpotenter Matrizen gibt.

8.3 Zerlegung eines Endomorphismus in algebraische Eigenräume

8.3.1 Der Satz von Caley- Hamilton

Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum und $f \in \mathrm{End}(V)$. In Polynome aus $\mathbb{K}[t]$ kann man lineare Abbildungen f einsetzen und erhält so einen Ring-Homomorphismus, oder auch **Algebra-**

Homomorphismus,

$$\Psi_f : \mathbb{K}[t] \longrightarrow \mathrm{End}(V) .$$

Da $\dim_{\mathbb{K}} \mathbb{K}[t] = \infty$ als Vektorraum über $\mathbb{K}$ und $\dim\ \mathrm{End}(V) = n^2$, hat Ψ_f sicher einen nichtleeren Kern. Es gilt sogar

SATZ 8.9 (SATZ VON CALEY-HAMILTON) *Sei $f \in \mathrm{End}(V)$ und $p_f \in \mathbb{K}[t]$ das charakteristische Polynom von f, dann gilt:*

$$p_f(f) = 0 \in \mathrm{End}(V) \tag{8.3}$$

Beweisskizze. Über die Formel mit den Permutationen ist $\det : \ \mathcal{M}(n \times n, R) \to R$ wohldefiniert für Matrizen mit Einträgen in einem kommutativen Ring R mit 1. In unserem Fall wählen wir $R = \mathrm{Bild}\ \Psi_f =: \mathbb{K}[f]$. In einer Basis $\mathcal{B} = (b_1, ..., b_n)$ von V sei $\Phi^{\mathcal{B}}_{\mathcal{B}}(f) =: A = (a_{ij})$ mit

$$fb_j = \sum_i a_{ij} b_i \text{ für alle } j \quad \text{oder alternativ} \quad \sum_i \big(f\delta_{ij} - a_{ij}\big) b_i = 0 \tag{8.4}$$

Hierbei definiert

$$(f\mathbf{1} - A) = \big(f\delta_{ij} - a_{ij}\big)_{i,j} \in \mathcal{M}(n \times n, R)$$

eine Matrix mit Eingängen in R. Wir erinnern uns an die Definition der adjunkten Matrix in der Theorie der Determinanten:
Für $M \in \mathcal{M}(n{\times}n, R)$ ist $\mathrm{Adj}_{kj}(M)$ dadurch gegeben, dass m_{kj} durch 1 und alle übrigen Elemente der k-ten Zeile und der j-ten Spalte durch Nullen ersetzt werden. Für $M \in \mathcal{M}(n \times n, \mathbb{K})$ ist bekannt, dass gilt

$$M \cdot \big(\det \mathrm{Adj}_{kl}(M)\big)^T = \det M\, \mathbf{1} . \tag{8.5}$$

Für $M := t\mathbf{1} - A$ sind beide Seiten von (8.5) Polynome aus $\mathbb{K}[t]$. Einsetzen von f in (8.5) liefert also diese Gleichung auch für $M := f\mathbf{1} - A \in \mathcal{M}(n \times n, R)$ (das gilt sogar allgemein für jeden kommutativen Ring R mit 1, wenn man die Theorie der Determinanten gleich entsprechend allgemein aufzieht). Gleichung (8.5), aufgeschrieben in Komponenten, ergibt also für $M := f\mathbf{1} - A$:

$$\sum_j \big(f\delta_{ij} - a_{ij}\big) \det\big[\mathrm{Adj}_{kj}(f\mathbf{1} - A)\big] = \delta_{ik} \underbrace{\det(f\mathbf{1} - A)}_{=\pm p_f(f)} \tag{8.6}$$

Anwenden der beiden Seiten von (8.6) auf die Elemente b_k der Basis $\mathcal{B}$ liefert nach Summation über $i = 1, ..., n$ wegen (8.4)

$$0 = \sum_{i,j} \det[\mathrm{Adj}_{kj}(f\mathbf{1} - A)](f\delta_{ij} - a_{ij})b_i = p_f(f)b_k , \qquad k = 1, \ldots n$$

und damit $p_f(f) = 0 \in \mathrm{End}(V)$. □

8.3.2 Der Euklidische Algorithmus

Aus dem Satz 8.9 von Caley-Hamilton kann man die Zerlegung von V in die verallgemeinerten Eigenräume eines $f \in \mathrm{End}(V)$ erschließen. Dazu benötigen wir aber noch einige einfache Teilbarkeitsaussagen über Polynome in $\mathbb{K}[t]$. Wir betrachten zunächst ein Verfahren zur Division, den sogenannten Euklidischen Algorithmus.

SATZ 8.10 (EUKLIDISCHER ALGORITHMUS) *Seien* $p, q \in \mathbb{K}[t]$ *mit* $\mathrm{grad}\ \ p = n \geq 0,\ \mathrm{grad}\ \ q = m$ *und* $q \neq 0$. *Dann gibt es eindeutig bestimmte Polynome* $M, r \in \mathbb{K}[t]$ *mit*

$$p = Mq + r \ \ \textit{und} \ \ \mathrm{grad}\ \ r < \mathrm{grad}\ \ q = m \tag{8.7}$$

Beweis. Zunächst beweisen wir die Existenz der gesuchte Polynome M, r durch Induktion nach $n = \mathrm{grad}\ \ p$.
Induktionsverankerung: Für $n = 0$ verifiziert man die Behauptung direkt (mit $M = 0$ und $r = p$ mit $\mathrm{grad}\, r = 0$).
Induktionsschluss: Die Behauptung gelte für Polynome $\tilde{p}, \tilde{q}$ mit $\mathrm{grad}\, \tilde{p} < n$. Sei nun $p \in \mathbb{K}[t]$ mit $\mathrm{grad}\, p = n$.

Falls $n < m$, dann folgt die Behauptung sofort mit $M = 0$ und $r = p$.
Falls $n \geq m$, dann dividieren wir die führende Ordnung:
Sei $p(x) = ax^n +$ Terme niedrigerer Ordnung und $q(x) = bx^m +$ Terme niedrigerer Ordnung mit $b \neq 0$. Für $\tilde{q}_1(x) = \frac{a}{b}x^{n-m}$ gilt dann $\tilde{q}_1 q(x) = ax^n +$ Terme niedrigerer Ordnung und damit

$$\mathrm{grad}(p - \tilde{q}_1 q) = \mathrm{grad}(\text{Terme niedrigerer Ordnung}) < n.$$

Mit $M' = \tilde{q}_1$ folgt dann aus der Induktionsannahme für $\tilde{p} = (p - M'q) \in \mathbb{K}[t]$:
Es gibt $M'' \in \mathbb{K}[t]$ und $r \in \mathbb{K}[t]$ mit $\mathrm{grad}\, r < \mathrm{grad}\, q$, sodass gilt:

$$p - M'q = M''q + r \Rightarrow p = (M' + M'')q + r\,.$$

Die Eindeutigkeit skizzieren wir nur:
Aus der Existenz der Zerlegungen $p = Mq + r = \tilde{M}q + \tilde{r}$ folgt

$$(M - \tilde{M})q + (r - \tilde{r}) = 0.$$

Das Betrachten des ersten nicht verschiedenen Gliedes in $M - \tilde{M} = b_j t^j r,\ b_j \neq 0$, liefert einen Widerspruch zu $\mathrm{grad}\ (r - \tilde{r}) < \mathrm{grad}\ \ q$. Also ist $M = \tilde{M}$ und damit $r = \tilde{r}$. $\square$

Bemerkung: Der Beweis von Satz 8.10 ist konstruktiv und enthält implizit einen Algorithmus für die Durchführung der Division mit Rest.

KOROLLAR 8.11 *Sukzessive Anwendung des Euklidischen Algorithmus (Satz 8.10) liefert für* p_1,

$p_2 \in \mathbb{K}[t]$ *ein abbrechendes Schema mit* $\text{grad } p_n < \text{grad } p_{n-1}$.

$$\begin{aligned} p_1 &= q_1 p_2 + p_3 \\ p_2 &= q_2 p_3 + p_4 \\ &\vdots \\ p_{k-2} &= q_{k-2} p_{k-1} + p_k \\ p_{k-1} &= q_{k-1} p_k + 0 \end{aligned}$$

Beweis. $\forall i$ ist entweder $p_i = 0$ oder aber $\text{grad } p_i < \text{grad } p_{i-1}$, also schließlich $p_i = 0$ für i genügend groß. □

KOROLLAR 8.12 *Das letzte nicht verschwindende* $p_k \in \mathbb{K}[t]$ *im Euklidischen Algorithmus für* $p_1, p_2 \in \mathbb{K}[t]$ *ist der größte gemeinsame Teiler (*ggT*) von* p_1 *und* p_2 *(also* $p_k = \text{ggT}(p_1, p_2)$*).*

Beweis. Betrachtet man den Euklidischen Algorithmus rückwärts, so teilt p_k alle vorangehenden Polynome $p_{k-1},\ p_{k-2}$, usw. bis p_1.
Betrachtet man ihn vorwärts, so folgt, dass alle Teiler p von p_1 und p_2 auch Teiler von $p_3, ..., p_k$ sind und damit muss p_k der größte gemeinsame Teiler von p_1 und p_2 sein. □

KOROLLAR 8.13 *Für alle* $p_1, p_2 \in \mathbb{K}[t]$ *existieren* $Q_1, Q_2 \in \mathbb{K}[t]$*, so dass:*

$$\text{ggT}(p_1, p_2) = Q_1 p_1 + Q_2 p_2 \,. \tag{8.8}$$

Beweis. Die vorletzte Zeile des Euklidischen Algorithmus liefert

$$p_k = Q'_{k-1} p_{k-1} + Q'_{k-2} p_{k-2}$$

Zusammen mit der vorhergehenden Zeile also:

$$\begin{aligned} p_k &= Q''_{k-2} p_{k-2} + Q''_{k-3} p_{k-3} \\ &\vdots \\ &= Q^{(j)}_{k-j} p_{k-j} + Q^{(j)}_{k-j-1} p_{k-j-1} \\ &= Q_2 p_2 + Q_1 p_1 \end{aligned}$$

Mit Korollar 8.12 folgt die Behauptung. □

KOROLLAR 8.14 *Sind* $p_1, p_2 \in \mathbb{K}[t]$ *teilerfremd, so gibt es* $Q_1, Q_2 \in \mathbb{K}[t]$ *mit*

$$1 = Q_1 p_1 + Q_2 p_2 \tag{8.9}$$

Beweis. $p_k = 1$ ist ggT(p_1, p_2). Also klar mit Korollar 8.13. □

Bemerkung Korollar 8.14 kann man verallgemeinern auf k Polynome $p_1, \ldots, p_k \in \mathbb{K}[t]$ mit $\text{ggT}(p_1, ..., p_k) = 1$.

DEFINITION UND SATZ 8.15 *Für $f \in \mathrm{End}(V)$ gibt es genau ein normiertes Polynom $\mu_f \in \mathbb{K}[t]$, sodass jedes Polynom p mit $p(f) = 0$ ein Vielfaches von μ_f ist. μ_f heißt* ***Minimalpolynom*** *von f. Es ist das normierte Polynom kleinsten Grades, welches f annulliert.*

Beweis. Existenz:
Sei $\mathcal{N}_f := \{p \in \mathbb{K}[t] \mid \text{p normiert und } p(f) = 0\}$ (p heisst normiert, falls $p(t) = t^k$+Terme niedrigerer Ordnung). Nach Satz 8.9 ist das charakteristische Polynom $p_f = \pm \det(f - t\mathbf{1}) \in \mathcal{N}_f$, also $\mathcal{N}_f \neq \{0\}$. Es gibt also ein $\mu \in \mathcal{N}_f$ mit $\mathrm{grad}\ \ \mu$ minimal und $\mu \neq 0$. Daraus folgt, dass $\mathrm{grad}\ \ \mu \leq \mathrm{grad}\ \ p_f = n = \dim\, V$. Dann liefert aber der Euklidische Algorithmus Satz 8.10, dass für alle $p \in \mathcal{N}_f$ Polynome M und R mit $p = M\mu + R$ und $\mathrm{grad}\ \ R < \mathrm{grad}\ \ \mu$ existieren. Da aber nach Voraussetzung gilt $p(f) = 0$ folgt daraus $0 = M(f)\mu(f) + R(f) = R(f)$ und damit $R = 0 \in \mathbb{K}[t]$, da wir angenommen hatten, dass $\mathrm{grad}\ \ \mu$ minimal in $\mathcal{N}_f$ ist.
Eindeutigkeit folgt sofort aus der Beziehung $\tilde{\mu} = Q\mu$ für ein $Q \in \mathbb{K}[t]$ und der Normiertheit von μ und $\tilde{\mu}$. □

8.3.3 Zerlegung in verallgemeinerte Eigenräume

Szenarium zur Erinnerung
Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und $f \in \mathrm{End}(V)$, dann gilt.

- Das charakteristische Polynom p_f von f zerfällt vollständig in Linearfaktoren mit $\lambda_j \in \sigma(f) \subset \mathbb{K}$ und r_j als algebraischer Vielfachheit von λ_j, also

$$p_f(t) = \prod_{j=1}^{k} (t - \lambda_j)^{r_j} \ .$$

- Da das Minimalpolynom μ_f nach Konstruktion p_f teilt, gilt mit $m_j \leq r_j$:

$$\mu_f(t) = \prod_j (t - \lambda_j)^{m_j} \ . \tag{8.10}$$

DEFINITION 8.16 *Ein Vektor $v \in V, v \neq 0$, heißt* ***verallgemeinerter Eigenvektor*** *oder* ***Hauptvektor*** *von f zu $\lambda \in \mathbb{K}$:⇔*

$$\exists r \in \mathbb{N} : \quad (f - \lambda)^r v = 0 \tag{8.11}$$

Die Hauptvektoren zu λ liegen in der Filtrierung $\mathrm{Kern}(f - \lambda) \subset \mathrm{Kern}(f - \lambda)^2 \subset ...$

DEFINITION UND SATZ 8.17 *Sei $f \in \mathrm{End}(V)$ und $\lambda \in \mathbb{K}$, dann heißt $E_\lambda := \bigcup_{r \in \mathbb{N}^*} \mathrm{Kern}\,(f - \lambda)^r$* ***verallgemeinerter Eigenraum*** *von f zu λ. Ist $m \in \mathbb{N}$ die Vielfachheit von λ als Nullstelle des Minimalpolynoms μ_f, dann gilt*

$$E_\lambda = \mathrm{Kern}\,(f - \lambda\mathbf{1})^m \tag{8.12}$$

Beweis. Da $\mathrm{Kern}(f-\lambda\mathbf{1}) \subset \mathrm{Kern}(f-\lambda\mathbf{1})^2 \subset \dots$ gilt $\bigcup_{k=1}^m \mathrm{Kern}(f-\lambda\mathbf{1})^k = \mathrm{Kern}(f-\lambda\mathbf{1})^m$. Für die Behauptung (8.12) genügt es also, zu zeigen dass $\mathrm{Kern}(f-\lambda\mathbf{1})^m = \mathrm{Kern}(f-\lambda\mathbf{1})^{m+1} = \dots$. Widerspruchsannahme: Es gebe $v \in V \setminus \{0\}$ mit $(f-\lambda)^n v = 0$, aber $w := (f-\lambda)^m v \neq 0$ für ein $n > m$. Mit der Faktorisierung (8.10) des Minimalpolynoms μ_f von f schreiben wir $\mu_f(t) = q(t)(t-\lambda)^m$. Da $t-\lambda$ kein Teiler von q ist, sind auch $(t-\lambda)^{n-m}$ und $q(t)$ teilerfremd. Mit Korollar 8.14 existieren damit Polynome $Q_1, Q_2 \in \mathbb{K}[t]$, so dass $1 = Q_1(t)q(t) + Q_2(t)(t-\lambda)^{n-m}$. Einsetzen von f statt t und anwenden auf w liefert

$$w = \big[Q_1(f)q(f) + Q_2(f)\underbrace{(f-\lambda)^{n-m}\big]w}_{=0} = Q_1(f)\underbrace{q(f)(f-\lambda)^m}_{=\mu_f(f)} v = 0$$

im Widerspruch zur Annahme. □

NOTIZ 8.18 *1) Es gilt $E_\lambda \neq \{0\} \;\Leftrightarrow\; \lambda$ ist Eigenwert von f (also $\lambda \in \sigma(f)$).*
2) Für $k < m$ ist $\mathrm{Kern}\,(f-\lambda)^k$ eine echte Teilmenge von E_λ.

SATZ 8.19 *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $f \in \mathrm{End}(V)$, dann zerfällt V in eine direkte Summe von verallgemeinerten Eigenräumen E_λ von f, d.h.*

$$V = \bigoplus_{\lambda \in \mathbb{K}} E_\lambda = \bigoplus_{\lambda_j \in \sigma(f)} E_{\lambda_j}\,, \tag{8.13}$$

wobei $\sigma(f)$ die Menge der Eigenwerte λ_i, $i = 1, \dots k$ von f ist.
Für $\lambda_j \in \sigma(f)$ mit Vielfachheit $m_j \in \mathbb{N}$ als Nullstelle des Minimalpolynoms μ_f sei

$$P_j(t) := \prod_{i \neq j} (t-\lambda_j)^{m_i} = \frac{\mu_f(t)}{(t-\lambda_j)^{m_j}}\,, \qquad j = 1, \dots k\,, \tag{8.14}$$

dann gibt es $Q_1, \dots Q_k \in \mathbb{K}[t]$ mit

$$1 = Q_1P_1 + \dots + Q_kP_k\,. \tag{8.15}$$

Sei $q_j := Q_jP_j\,,\ j = 1, \dots k$, dann ist $q_j(f) = Q_j(f)P_j(f)$ die Projektion auf $E_j := E_{\lambda_j}$ längs der direkten Summe der anderen Eigenräume E_i, $i \neq j$ (also längs $\bigoplus_{i \neq j} E_i$).

Bemerkung: Statt aus dem Minimalpolynom kann man die Projektion $q_j(f)$ auch aus dem charakteristischen Polynom $p_f(t)$ erzeugen.

Beweis von Satz 8.19. 1) Die Polynome $P_1, \dots, P_k \in \mathbb{K}[t]$ sind nach Konstruktion relativ prim (d.h. $\mathrm{ggT}(P_1, \dots P_k) = 1$), daher existieren nach Korollar 8.14 $Q_1, \dots, Q_k \in \mathbb{K}[t]$ mit (8.15) und das Einsetzen von f in in die Polynome in (8.15) liefert

$$\mathbf{1} = \sum_{j=1}^k q_j(f)\,. \tag{8.16}$$

Damit folgt (8.13), wenn wir zeigen können, dass $q_j(f)$ ein Projektor auf E_j längs $\bigoplus_{i \neq j} E_i$ ist.
2) Wir zeigen zunächst, dass $q_j(f)$ Projektoren sind, d.h. dass gilt

$$q_i(f)q_j(f) = \delta_{ij}q_i(f)\,, \qquad i, j = 1, \dots k\,. \tag{8.17}$$

Aus (8.14) folgt, dass das Minimalpolynom μ_f für $i \neq j$ das Polynom $P_i P_j$ teilt, also auch das Polynom $q_i q_j$. Es gibt also $Q \in \mathbb{K}[t]$, so dass $q_i q_j = Q\mu_f$ und damit folgt für alle $i \neq j$ aus der Definition des Minimalpolynoms sofort

$$q_i(f) q_j(f) = Q(f)\mu_f(f) = 0 \,, \qquad i \neq j \,,\ i,j = 1, \ldots k \,. \tag{8.18}$$

Aus (8.16) und (8.18) folgt

$$q_i(f) = \sum_{j=1}^{k} q_i(f) q_j(f) = q_i(f)^2 + \sum_{\substack{j=1 \\ j\neq i}}^{k} q_i(f) q_j(f) = q_i^2(f) \,, \qquad i = 1, \ldots k \,, \tag{8.19}$$

Die Gleichungen (8.18) und (8.19) zusammen liefern (8.17), die Abbildungen $q_j(f) \in \mathrm{End}(v)$, $j = 1, \ldots k$ sind also Projektoren auf Bild $q_j(f)$ längs $\bigoplus_{i\neq j}$ Bild $q_i(f)$.
3) Die Gleichungen (8.16) und (8.17) zusammen ergeben mit Satz 6.7 die Zerlegung

$$V = \bigoplus_{j=1}^{k} \mathrm{Bild}\ q_j(f) \,.$$

Zu zeigen bleibt also nur noch

$$\mathrm{Bild}\ q_j(f) = E_j \,.$$

"$\subset$": Wegen (8.14) und $q_j = Q_j P_j$ gilt

$$\big(f - \lambda_j\big)^{m_j} q_j(f) = Q_j(f)\mu_f(f) = 0 \,,$$

da $\mu_f(f) = 0$ nach Definition des Minimalpolynoms. Damit ist $\mathrm{Bild}\ q_j \subset \mathrm{Kern}\ \big(f - \lambda_j\big)^{m_j}$ und mit Satz 8.17 folgt $\mathrm{Bild}\ q_j \subset E_j$.
"$\supset$": Wegen (8.16) gilt für alle $v \in V$

$$q_j(f) v = \Big(\mathbf{1} - \sum_{i\neq j} q_i(f)\Big) v = v - \sum_{i\neq j} q_i(f) v \,, \qquad j = 1, \ldots k \,. \tag{8.20}$$

Wegen (8.14) teilt aber $(t - \lambda_j)^{m_j}$ für $i \neq j$ das Polynom $P_i(t)$ und damit auch auch $q_i(t) = Q_i(t) P_i(t)$. Damit existiert ein Polynom $Q \in \mathbb{K}[t]$ mit $q_i(t) = Q(t)\big(t - \lambda_j\big)^{m_j}$, $i \neq j$, und da $E_j = \mathrm{Kern}\ \big(f - \lambda_j\big)$ wegen Satz 8.17 folgt für alle $v_j \in E_j$

$$q_i(f) v_j = Q(f)\big(f - \lambda_j\big)^{m_j} v_j = 0 \,, \quad i \neq j\, i,j = 1, \ldots k \,.$$

Einsetzen in (8.20) liefert $q_j(f) v_j = v_j$ für $v_j \in E_j$ und damit $E_j \subset \mathrm{Bild}\ q_j(f)$. $\square$

Das Ergebnis lässt sich wie folgt analog zum Spektralsatz für selbstadjungierte Endomorphismen umformulieren und beschreibt die Zerlegung von $f \in \mathrm{End}(V)$ in diagonalen und nilpotenten Anteil.

KOROLLAR 8.20 *Sei* $f \in \mathrm{End}(V)$ *wie oben und* $\Pi_j := q_j(f)$ *der Projektor auf den verallgemeinerten Eigenraum* E_j *zum Eigenwert* $\lambda_j \in \mathbb{K}$*, wobei* $\lambda_1, ..., \lambda_k$ *die paarweise verschiedenen Eigenwerte von* f *sind. Dann gilt*

$$\sum \Pi_j = \mathbf{1} \tag{8.21}$$

$$N_j := \big(f - \lambda_j\big)\Pi_j \text{ ist nilpotent und } [\Pi_j, N_i] = 0 \text{ für } i, j = 1 \dots k \tag{8.22}$$

$$f = \sum_{j=1}^{k} \Big(\lambda_j \Pi_j + N_j\Big) =: f_s + f_n \text{ mit } f_n \text{ nilpotent und } [f_s, f_n] = 0\ . \tag{8.23}$$

Beweis. Gleichung (8.21) ist wegen $q_j(f) = \Pi_j$ einfach (8.16).
Gleichung (8.22): Da $\Pi_j = q_j(f)$ mit f und mit sich selbst vertauscht folgt die Kommutatorrelation für $i = j$ sofort aus der Definition von N_j. Für $i \neq j$ gilt $\Pi_i \Pi_j = 0$ und damit, da f mit Π_j vertauscht, auch $\Pi_j N_i = 0 = N_i \Pi_j$. Da f mit N_i vertauscht, gilt für alle $\ell \in \mathbb{N}_0$ aus der Definition von Π_j und (8.14)

$$N_j^{m_j+\ell} = \big(f - \lambda_j\big)^{m_j+\ell}\Pi_j = \big(f - \lambda_j\big)^{\ell} Q_j(f)\mu_f(f) = 0\,,$$

also ist N_j nilpotent (mit Nilpotenzgrad m_j).
Gleichung (8.23): Mit (8.22) und (8.21) gilt

$$\sum_{j=1}^{k} N_j = \sum_{j=1}^{k} (f - \lambda_j)\Pi_j = f \sum_{j=1}^{k} \Pi_j - \sum_{j=1}^{k} \lambda_j \Pi_j = f - \sum_{j=1}^{k} \lambda_j \Pi_j\ .$$

Weiterhin ist f_n nilpotent, da Summen nilpotenter Operatoren nilpotent sind (selber!). Die Vertauschungsrelation folgt sofort aus (8.22), da Summen kommutierender Operatoren kommutieren (selber!). □

8.4 Die Jordan-Normalform von Endomorphismen

KOROLLAR 8.21 *Sei* V *ein* $\mathbb{K}$*-Vektorraum und* $f \in \mathrm{End}(V)$*. Das charakteristische Polynom* $p_f(t)$ *zerfalle vollständig in Linearfaktoren. Dann hat* V *eine Jordan-Basis* $\mathcal{A}$ *für* f*.*
Genauer gilt: Seien $\lambda_1, ..., \lambda_k \in \mathbb{K}$ *die paarweise verschiedenen Eigenwerte von* f *und* $E_1, \dots E_k$ *die zugehörigen verallgemeinerten Eigenräume.*

1) Für $j = 1, \dots k$ *sei* Part_j *die dem nilpotenten Endomorphismus* $\big(f - \lambda_j\big)|_{E_j} \in \mathrm{End}(E_j)$ *zugeordnete Partition* $\dim E_j = r_{j,1} + ... + r_{j,\ell}$ *und* $\mathcal{A}_j$ *eine Jordan-Basis von* E_j *für* $(f - \lambda_j)|_{E_j}$.

Dann gilt für alle $j = 1, \dots k$

$$\Phi^{\mathcal{A}_j}_{\mathcal{A}_j}(f|_{E_j}) = J_{\mathrm{Part}_j}(\lambda_j) := \begin{pmatrix} J_{r_{j,1}}(\lambda_j) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & J_{r_{j,\ell}}(\lambda_j) \end{pmatrix} \tag{8.24}$$

$$\textit{mit } J_{r_{j,s}}(\lambda_j) = \begin{pmatrix} \lambda_j & 1 & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & \lambda_j \end{pmatrix} \in \mathcal{M}(r_{j,s} \times r_{j,s}, \mathbb{K}), \ s = 1, \dots \ell \,. \tag{8.25}$$

2) Sei $\mathcal{A} = (\mathcal{A}_1, ..., \mathcal{A}_k)$ *die induzierte Jordan-Basis von* V *für* f*. Dann gilt:*

$$\Phi^{\mathcal{A}}_{\mathcal{A}}(f) = \begin{pmatrix} J_{\mathrm{Part}_1}(\lambda_1) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & J_{\mathrm{Part}_k}(\lambda_k) \end{pmatrix} \tag{8.26}$$

3) Die Jordan-Kästchen $J_{r_{j,s}}(\lambda_j)$ *sind bis auf Permutation eindeutig bestimmt.*

Beweis. ad 1) Gemäß der Klassifikation nilpotenter Endomorphismen (Satz 8.8) hat $N_j := \big(f - \lambda_j\big)|_{E_j}$ die Jordan-Normalform

$$\Phi^{\mathcal{A}_j}_{\mathcal{A}_j}(N_j) = \begin{pmatrix} J_{r_{j,1}}(0) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & J_{r_{j,\ell}}(0) \end{pmatrix}$$

in einer Jordan-Basis $\mathcal{A}_j$ von E_j. Addition von $\lambda_j \mathbf{1}_{E_j}$ liefert die behauptete Matrixdarstellung von $f|_{E_j}$.

ad 2) Die Matrixdarstellung von f folgt sofort aus 1).

ad 3) Die Eindeutigkeit folgt aus Satz 8.8, 3, (Eindeutigkeitsresultat für nilpotente Endomorphismen). □

NOTIZ 8.22 *Aus der expliziten Gestalt der Jordan'schen Normalform lässt sich das Minimalpolynom ablesen. Sei* $m_j := \max\{r_{j,1}, ..., r_{j,\ell}\}$ *die Länge eines maximalen Jordanblocks in* $J_{\mathrm{Part}_j}(\lambda_j)$*, dann gilt:*

$$\mu_f = \prod_{j=1}^{k} \big(\lambda_j - t\big)^{m_j}.$$

8.5 Lösung linearer Differentialgleichungen erster Ordnung

Erinnerung: Das Differentialgleichungssystem $\dot{x}(t) = Ax(t)$ mit Anfangswert $x(0) = c \in \mathbb{R}^n$ oder $\mathbb{C}^n$ hat die Lösung $x(t) = e^{At}c$.

Für kommutierende Matrizen $A, B \in \mathcal{M}(n \times n, \mathbb{C})$ gilt $e^{A+B} = e^A e^B$.

Im Folgenden wollen wir uns der Frage zuwenden, wie wir mit Hilfe der Jordan-Basis die Funktion

$$e^{At} = \sum_{j=0}^{\infty} \frac{1}{j!} (At)^j$$

berechnen können.

In einer Jordan-Basis $\mathcal{A}$ für $A \in \mathrm{End}(\mathbb{R}^n)$ gilt (mit S als Basiswechselmatrix)

$$\Phi_{\mathcal{A}}^{\mathcal{A}}(A) = SAS^{-1} = J_A = \begin{pmatrix} J_{\mathrm{Part}_1}(\lambda_1) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & J_{\mathrm{Part}_k}(\lambda_k) \end{pmatrix}$$

und damit

$$A = S^{-1} J_A S\,, \quad \text{also} \quad e^{At} = S^{-1} e^{J_A t} S\,.$$

NOTIZ 8.23 *Es reicht, die Exponentialfunktion für einen einzigen Jordanblock $J_{r_{j,s}}(\lambda_j)$, $\lambda_j \in \sigma(A)$, zu berechnen, denn es gilt*

$$e^{J_A t} = \begin{pmatrix} e^{J_{r_{1,1}}(\lambda_1)t} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & e^{J_{r_{k,\ell}}(\lambda_k)t} \end{pmatrix}.$$

Da für einen Jordanblock gilt $J_r(\lambda) = \lambda \mathbf{1} + N$ mit

$$N = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \textit{und } [N, \lambda \mathbf{1}] = 0$$

folgt

$$e^{J_r(\lambda)t} = e^{\lambda t} e^{Nt} \quad \textit{und} \quad e^{Nt} = \sum_{j=0}^{r-1} \frac{1}{j!} (Nt)^j = \begin{pmatrix} 1 & t & \cdots & \frac{t^{r-1}}{(r-1)!} \\ & \ddots & \ddots & \vdots \\ 0 & & & t \\ & & & 1 \end{pmatrix},$$

treten also gerade die Quasipolynome $\frac{t^j}{j!} e^{\lambda t}$ auf.

Fazit: Sei $A \in \mathcal{M}(n \times n, \mathbb{R})$ und $x(t)$ eine Lösung von $\dot{x} = Ax$, dann sind die Koordinaten $x_j(t)$ von $x(t) \in \mathbb{R}^n$ Linearkombinationen der Funktionen

$$t^j e^{ta} \cos(bt),\ t^\ell e^{ta} \sin(bt)\,, \quad \text{mit } \lambda = a + ib \in \sigma(A)\,,\ b \geq 0\,,\ j, \ell \in \{0, ..., n-1\}\,.$$

(Genauer gilt $j, \ell \leq$ maximale Länge eines Jordanblocks J_r zum Eigenwert λ.)

Das ergibt sich aus den komplexen Quasipolynomen $t^j e^{\lambda t}$ für $\lambda \in \mathbb{C}$, indem man eine Basis aus komplexen Eigenvektoren $\{x_1 + iy_1, ..., x_r + iy_r\}$ eines zyklischen Summanden in E_λ in Real- und Imaginärteil zerlegt. Man kann natürlich auch die Quasipolynome $t^j e^{\lambda t}$ stehen lassen.

Beispiel:
1) Für $y \in C^2([0, \infty), \mathbb{R})$ gelte $\ddot{y} = 0$. Das lässt sich mit $x_1 := y$ und $x_2 := \dot{y}$ umschreiben auf das System

$$\dot{x} = \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x =: Ax\,.$$

Die Matrix A ist schon in Jordan-Form mit Eigenwert $\lambda = 0$ (also nilpotent) und die Lösung ist

$$x(t) = e^{At}x(0) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_0 \\ \dot{y}_0 \end{pmatrix} \text{ mit } x(0) = \begin{pmatrix} y_0 \\ \dot{y}_0 \end{pmatrix}$$

und damit folgt für y als Lösung $y(t) = x_1(t) = y_0 + t\dot{y}_0$.

2) Beim aperiodischen Grenzfall einer Schwingung treten nilpotente Teile in der Zerlegung $A = A_s + A_n$ auf.

3) Entsprechende Terme treten bei gekoppelten Feder-Massenschwingern auf.

8.6 Übungen

1. Beweisen Sie Notiz 8.8.7!

2. Beweisen Sie Notiz 8.8.22!