Lineare Algebra I und II

Skript zur Vorlesung von Prof. Dr. Carl-Friedrich Bödigheimer

Wintersemester 2014/2015, Sommersemester 2015
Stand: 17. Juli 2015

Hinweis:
Im Sommersemester (ab etwa Kapitel 5) wird dieses Skript erstellt von Jürgen Kanzler.
Fehler, Korrekturen oder sonstige Anmerkungen zum Skript können gern an `s6jukanz@uni-bonn.de` geschickt werden.

Inhaltsverzeichnis

Überblick

Achtung:

Diese Version des Skriptes ist nicht vollständig korrigiert! Verbesserungsvorschläge, Rechtschreibfehler und inhaltliche Fehler dürfen gerne an linaskript@tauradian.de geschickt werden. Wir freuen uns über Rückmeldung!

Vorwort

Lineare Algebra:

- Theorie der linearen Gleichungssysteme
- Theorie der Vektorräume und linearen Abbildungen
- Theorie der Matrizen und Vektoren
- Unbekannte / Variablen x kommen nur zur ersten Potenz vor
- Keine gemischten Terme $x_i \cdot x_j$

0 Warum und zu welchem Ende studieren wir Lineare Algebra?

Was ist eigentlich *Lineare Algebra*?

= Theorie der linearen Gleichungssysteme

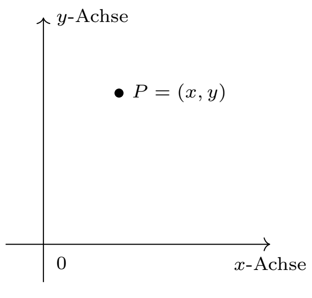
= Theorie der Vektorräume und linearen Abbildungen

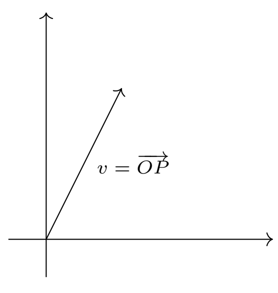
= Theorie der Vektoren und Matrizen

Früher hatten Lehrbücher zu Linearen Algebra öfter als heute noch den Zusatz „… *und Analytische Geometrie*“. Wie ist der Zusammenhang?

Geometrie	Analytische Geometrie
Punkte in der Ebene	„Koordinatengeometrie“
P	$P = (x, y)$
	zwei Koordinaten
Geraden G	Geradengleichung $ax + by = c$
Kreis, Ellipse, …	$ax^2 + by^2 = c$

Geometrie	Analytische Geometrie	Vektorrechnung
P	$P \;= (x, y)$	$v \;= \begin{pmatrix} x \\ y \end{pmatrix}$
$\bullet$	$O \;= (0, 0)$	$= \overrightarrow{OP}$
	Ursprung	Ortvektor $\begin{cases} \text{Richtung} \\ \text{Länge} \end{cases}$





Vektoren kann man addieren, Punkte aber nicht!

Descartes und die Einführung der Koordinaten

Wir nennen heute ein (rechtwinkliges) Koordinatensystem „kartesisch“, weil Koordinaten im Jahre 1637 von *Rene Descartes* in seinem Buch „La Geometrie“ in die Geometrie eingeführt worden sind.

Er hat damit ein - seit der Antike - ungelöstes *Problem des Pappus* (4. Jhdt) gelöst und war mit Recht sehr stolz darauf. Bevor wir das Problem erklären, hier eine schöne Würdigung von Descartes Leistung durch *J. d'Alembert*, ein Mathematiker und Physiker, der zusammen mit *D. Diderot* einer der Herausgeber der berühmten Enzyklopädie:
Er nennt 1751 in der Einleitung der Enyklopädie:

[die Einführung der Koordinaten] eine der umfassendsten und glücklichsten Einfälle, die der menschlichen Geist je gehabt hat und der stets den Schlüssel zu den eingehendsten Untersuchungen nicht nur in der höheren Mathematik, sondern in allen mathematisch-physikalischen Wissenschaft darstellen wird.

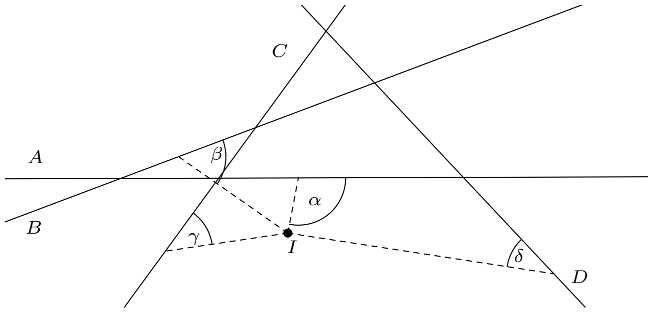
Pappus-Problem

gegeben: n Geraden $A, B, C, D, \dots$ in der Ebene
n Winkel $\alpha, \beta, \gamma, \delta, \dots$
λ Zahl

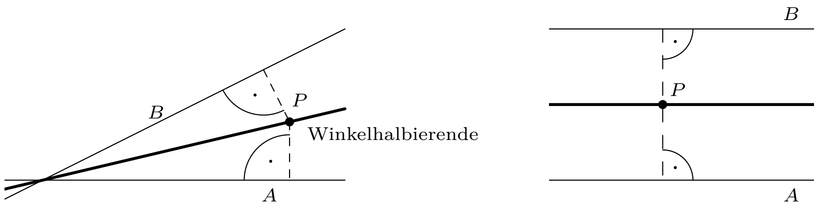
Was ist der Ort aller Punkte P der Ebenen, deren Abstände $a, b, c, d, \dots$ von $A, B, C, D, \dots$ gemessen unter den Winkeln $\alpha, \beta, \gamma, \delta, \dots$ die Gleichung

$$ab = \lambda cd$$

erfüllen?

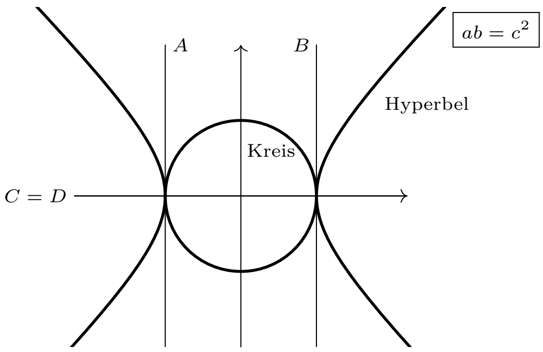


Beispiel 0.1. Zwei Geraden A, B, $\alpha = \beta = 90°$, $\lambda = 1$.



Beispiel 0.2. 4 Geraden:

$$\begin{aligned} A &: x = -1, & \qquad & \alpha = \beta = \gamma = \delta = 90°, \\ B &: x = +1, & & \\ C = D &: y = 0, & & \lambda = 1. \end{aligned}$$



$\boxed{-1 \leq x \leq +1}$

$$\begin{aligned} &(x - (-1))(1 - x) = y^2 \\ &= (1 + x)(1 - x) \\ &= 1 - x^2 \end{aligned}$$

$$\boxed{x^2 + y^2 = 1} \qquad \text{Kreis}$$

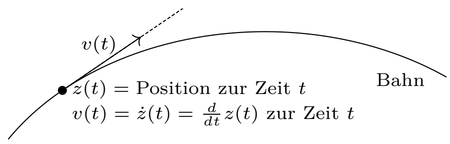
$$\boxed{x \leq -1,\ x \geq +1} \qquad\qquad (x - (-1)(x-1) = y^2$$

$$= x^2 - 1$$

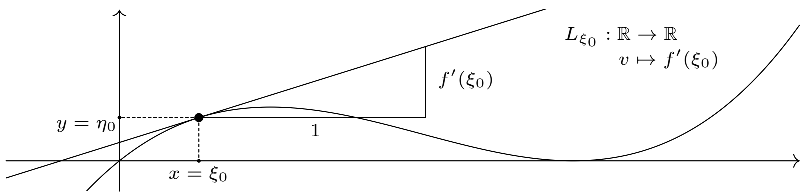
$$\boxed{x^2 - y^2 = 1} \qquad \text{Hyperbel}$$

Vektoren und Physik

Der Begriff ds Vektors stammt aus der Physik: er beschreibt eine Größe mit *Richtung* und *Betrag*, wie ihn die Physik als Kraft oder Geschwindigkeit kennt.
Da sich *Kräfte überlagern* können oer *Geschwindigkeiten addieren* - eine misst die Geschwindigkeit eines bewegten Systems, eine zweite die Geschwindigkeit relativ zum ersten - ist es völlig einleuchtend, dass man Vektoren addieren, subtrahieren, strecken oder stauchen kann, - während man dies mit Punkten nicht kann:



$v(t)$ ist die Richtung, die eine Massepunkt nähme, hörten alle Kräfte, auf ihn zu wirken, plötzlich auf.
Darum betrachtet man in der Mathematik lineare Approximationen an Kurven (= Graphen von Funktionen), als Tangente mit der Ableitung als Steigung - und *nicht*, weil die lineare Approximation „einfach" ist.



- Bei einer Funktion $f : \mathbb{R} \to \mathbb{R}, f(x) = y$ *einer* Variablen ist eine Tangente an den Graphen eine Gerade und i.W.durch ihre Steigung (neben (x_0, y_0)) bestimmt.
- Bei einer Funktion $f : \mathbb{R}^2 \to \mathbb{R}, f(x_1, x_2) = y$ von zwei Variablen haben wir einen Tangentialebene an den Graphen; diese ist nicht mehr durch eine Zahl bestimmt, sondern durch zwei: nämlich die beiden partiellen Ableitungen
$$\frac{\partial f}{\partial x_1}(\xi_0, \eta_0), \qquad \frac{\partial f}{\partial x_2}(\xi_0, \eta_0)$$
$$L_{\xi_0, \eta_0} : \mathbb{R}^2 \longrightarrow \mathbb{R} \qquad \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \longmapsto \frac{\partial f}{\partial x_1}(\xi_0, \eta_0) v_1 + \frac{\partial f}{\partial x_2}(\xi_0, \eta_0) v_2$$
- Bei einer vektorwertigen Funktion
$$F : \mathbb{R}^n \longrightarrow \mathbb{R}^m \quad F(x_1, \dots, x_n) = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix}$$
ist die Ableitung was? - Eine lineare Abbildung
$$L_{\xi_0, \eta_0, \dots} : \mathbb{R}^n \longrightarrow \mathbb{R}^m \qquad \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \longmapsto \begin{pmatrix} a_{11} v_1 + a_{12} v_2 + \ldots + a_{1n} v_n \\ a_{21} v_1 + a_{22} v_2 + \ldots + a_{2n} v_n \\ \vdots \\ a_{n1} v_1 + a_{n2} v_2 + \ldots + a_{nn} v_n \end{pmatrix}.$$

Dies sind m *lineare* (weil kein v_i in höherer Potenz auftritt und auch kein gemischter Term $v_i v_j$) Gleichungen in n Variablen $v_1, \dots, v_n$.

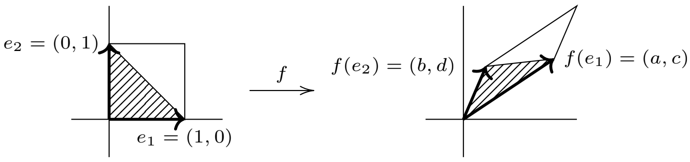
Also lineare Gleichungssysteme!

Beispiel 0.3. Drehungen, Streckungen, ...

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \quad (x,y) \longmapsto f(x,y) = (x', y')$$

mit

$$\begin{aligned} x' &= f_1(x,y) = ax + by, \\ y' &= f_2(x,y) = cx + dy. \end{aligned}$$

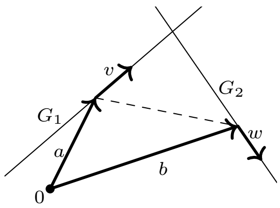


$$\left.\begin{aligned} a &= \cos\alpha & b &= -\sin\alpha \\ c &= \sin\alpha & d &= \cos\alpha \end{aligned}\right\} \text{Drehung um Winkel } \alpha$$

$$\left.\begin{aligned} a &= -1 & b &= 0 \\ c &= 0 & d &= 1 \end{aligned}\right\} \text{Spiegelung an der } y\text{-Achse}$$

Beispiel 0.4. Schnitt zweier Geraden im $\mathbb{R}^2$

$$\begin{aligned} G_1 &:= G_{a,v} : \boxed{a + sv} \quad (v \neq 0) \\ G_2 &:= G_{b,w} : \boxed{b + tw} \quad (w \neq 0) \end{aligned}$$



Schnittpunkt? $\exists t, s \in \mathbb{R}$:

$$a + sv = b + tw$$

$$\left.\begin{array}{rlll} & \{sv - tw & = b - a & = c \\ \mathrm{I} & \begin{cases} v_1 s - w_1 t \\ v_2 s - w_2 t \end{cases} & \begin{array}{l} = b_1 - a_1 \\ = b_2 - a_2 \end{array} & \begin{array}{l} = c_1 \\ = c_2 \end{array} \\ \mathrm{II} & & & \end{array}\right| \begin{array}{l} \text{Lösen?} \\ v_1 \neq 0 \text{ oder } v_2 \neq 0 \\ \text{und} \\ w_1 \neq 0 \text{ oder } w_2 \neq 0 \end{array}$$

Das Gleichungssystem hat keine Lösung, genau eine Lösung oder unendlich viele Lösungen, je nachdem ob G_1, G_2 parallel sind, sich schneiden oder gleich sind.

$$\begin{aligned} w_2\mathrm{I} : &\qquad v_1 w_2 s - w_1 w_2 t = w_2 c_1 \\ w_1\mathrm{II} : &\qquad v_2 w_1 s - w_1 w_2 t = w_1 c_2 \end{aligned}$$

$$w_2\mathrm{I} - w_1\mathrm{II} : \qquad \underbrace{v_1 w_2 - v_2 w_1}_{\substack{:=\Delta \\ \text{Determinante}}} s = \underbrace{w_2 c_1 - w_1 c_2}_{\langle -w^\perp, c\rangle}$$

$$\begin{aligned} v_2\mathrm{I} : &\qquad v_1 v_2 s - v_2 w_1 t = v_2 c_1 \\ v_1\mathrm{II} : &\qquad v_1 v_2 s - v_1 w_2 t = v_1 c_2 \end{aligned}$$

$$v_2\mathrm{I} - v_1\mathrm{II} : \qquad \Delta t = \underbrace{v_2 c_1 - v_1 c_2}_{-v^\perp, c}$$

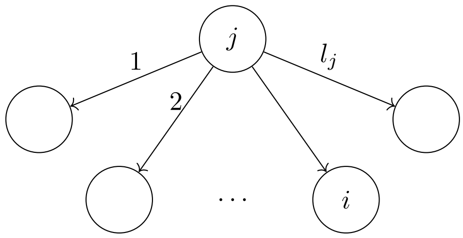
- Falls $\Delta \neq 0$:

$$s = \frac{w_2 c_1 - w_2 c_2}{\Delta} = \frac{\begin{vmatrix} c_1 & w_1 \\ c_2 & w_2 \end{vmatrix}}{\begin{vmatrix} v_1 & w_1 \\ v_2 & w_2 \end{vmatrix}}$$

$$t = \frac{v_2 c_1 - v_1 c_2}{\Delta} = -\frac{\begin{vmatrix} v_1 & c_1 \\ v_2 & c_2 \end{vmatrix}}{\begin{vmatrix} v_1 & w_1 \\ v_2 & w_2 \end{vmatrix}}$$

- Falls $\Delta = 0$ und $\langle w^\perp, c\rangle = \langle v^\perp, c\rangle = 0$: $G_1 = G_2$.
- Falls $\Delta = 0$ und $\langle w^\perp, c\rangle \neq 0$ oder $\langle v^\perp, c\rangle \neq 0$: $G_1 \cap G_2 = \emptyset$.

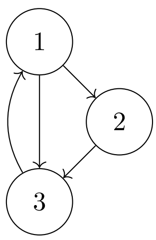
Beispiel 0.5 (Google-Matrix, Page-Rank-Algorithmus)**.** Internetseiten j mit Gewichten $x_j \geq 0$,



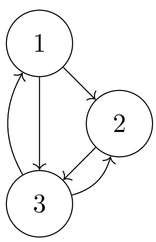
i erhält von j den Bruchteil $\frac{1}{l_j}$ von dessen Gewicht x_j. (l_j sei die Anzahl der Links die von j ausgehen.)

$$x_i = \sum_{j \to i} \frac{x_j}{l_j}$$

<u>Lösungen</u>



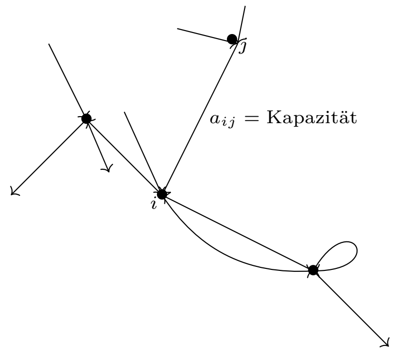
$$x_1 = x_3 \qquad x_2 = \frac{x_1}{2} \qquad x_3 = \frac{x_1}{2} + x_2 \qquad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} u \\ \frac{u}{2} \\ u \end{pmatrix}$$



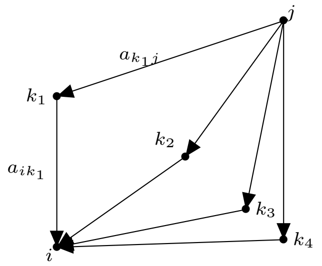
$$x_1 = x_3 \qquad x_2 = \frac{x_1}{2} + \frac{x_3}{2} \qquad x_3 = \frac{x_1}{2} + x_2 \qquad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}u \\ \frac{3}{4}u \\ u \end{pmatrix}$$

- Gibt es eine Lösung?
- Gibt es eine eindeutige Lösung?
- ... mit einer *klaren Ordnung*, d.h. $x_i \neq x_j$ für $i \neq j$?

Beispiel 0.6 (Transportfragen)**.** Graph mit gerichteten Kanten



a_{ij} = Transportmenge an einem Tag



$$a_{ij}^{(2)} := \sum_k a_{ik} a_{kj}$$

$$A = (a_{ij}) \quad \rightsquigarrow A^2 \qquad (a_{ij}^{(2)})$$

$$A = (a_{ij}), B = (b_{ij}) \quad \rightsquigarrow \quad C = (c_{ij}), \qquad c_{ij} := \sum_k a_{ik} b_{kj} \quad \boxed{\text{Produkt von Matrizen}}$$

Wir werden Matrizen A, B Abbildungen f, g zuordnen, so dass das Matrixprodukt C der Komposition $f \circ g$ von Abbildungen entspricht und umgekehrt.

Beispiel 0.7 (Interpolation durch Polynome)**.** gegeben: Zahlen $x_1, \ldots, x_n$ und $y_1, \ldots, y_n$
gesucht: ein Polynom $y = f(x) = a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ mit $f(x_i) = y_i$ für $i = 1, \ldots, n$.

$$\begin{array}{ccc}
(1) & \boxed{x = x_1} & a_{n-1}x_1^{n-1} + \ldots + a_1 x_1 + a_0 = y_1 \\
(2) & \boxed{x = x_2} & a_{n-1}x_2^{n-1} + \ldots + a_1 x_2 + a_0 = y_2 \\
 & \vdots & \vdots \\
(n) & \boxed{x = x_n} & a_{n-1}x_n^{n-1} + \ldots + a_1 x_n + a_0 = y_n
\end{array}$$

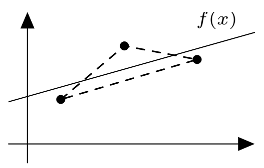
Hier sind: $x_1, \ldots, x_n, y_1, \ldots, y_n$ bekannt und $a_0, \ldots, a_{n-1}$ gesucht.

<u>$n = 2$</u> Gerade durch zwei Punkte

<u>Schema</u> [Wronski]

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Beispiel 0.8 (Methode der kleinsten Quadrate)**.**
gegeben: Messpunkte $x_1, \ldots, x_n, y_1, \ldots, y_n$
gesucht: Gerade „bester“ Interpolation $f(x) = mx + b$ (m, b sind Unbekannte) mit $\sum_{i=1}^n (f(x_i) - y_i)^2$ minimal.



$$\Phi(m,b) = \sum_i (mx_i + b)^2 \overset{!}{=} \text{minimum}$$

partielle Ableitungen nach m und nach b:

$$\begin{aligned}
\frac{\partial\Phi}{\partial m} &= \sum_i 2(mx_i + b)x_i \\
&= \sum_i (2x_i^2 m + bx_i) \\
&= 2\underbrace{\sum_i x_i^2}_{\alpha_{11}} + 2\underbrace{\sum_i x_i}_{\alpha_{12}} b
\end{aligned}$$

$$\begin{aligned}
\frac{\partial\Phi}{\partial b} &= \sum_i 2(mx_i + b)\cdot 1 \\
&= 2\underbrace{\sum_i x_i}_{\alpha_{21}} m + 2\underbrace{\sum_i 1}_{=n=\alpha_{22}} b
\end{aligned}$$

$$\boxed{\begin{aligned} \alpha_{11}m + \alpha_{12}b &= 0 \\ \alpha_{21}m + \alpha_{22}b &= 0 \end{aligned}}$$

Beispiel 0.9 (Chemische Reaktionsgleichgewichte)**.**

$\boxed{S_1} \xrightarrow{a_1} \boxed{S_2}$		chemische Stoffe
y_1	y_2	Konzentrationen

$$\boxed{\begin{aligned} \dot{y}_1 &= -a_1 y_1 \\ \dot{y}_2 &= a_1 y_1 \end{aligned}} \quad \begin{aligned} &S_1 \text{ nimmt mit der Rate } k_1 \text{ ab} \\ &S_2 \text{ ist proportional zu } S_1 \end{aligned}$$

$$\boxed{S_1} \underset{a_2}{\overset{a_1}{\rightleftarrows}} \boxed{S_2}$$

$$\begin{aligned}
\dot{y}_1 &= -a_1 y_1 + a_2 y_2 \\
\dot{y}_2 &= a_1 y_1 - a_2 y_2
\end{aligned}$$

Beispiel

$$\boxed{S_1} \xrightarrow{a_1} \boxed{S_2} \underset{a_3}{\overset{a_2}{\rightleftarrows}} \boxed{S_3}$$

$$\begin{aligned}
\dot{y}_1 &= -a_1 y_1 \\
\dot{y}_2 &= a_1 y_1 - a_2 y_2 + a_2 y_3 \\
\dot{y}_3 &= \phantom{-a_1 y_1 -{}} a_2 y_2 - a_3 y_3
\end{aligned}$$

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \end{pmatrix} = \begin{pmatrix} -a_1 & 0 & 0 \\ a_1 & -a_2 & a_3 \\ 0 & a_2 & -a_3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

(1) nur eine Funktion und eine Gleichung

$$\dot{y} = ay \qquad\qquad \underline{\text{Ansatz:}}\ y = e^{at}, \quad \dot{y} = ae^{at}$$

(2) mehrere Funktionen und ein Gleichungssystem

$$\dot{y} = Ay \quad (A \text{ Matrix}) \qquad \underline{\text{Ansatz:}} y = e^{At}$$

$$e^t = \sum_{k=0}^{\infty} \frac{t^k}{k!} = 1 + t + \frac{t^2}{2} + \frac{t^3}{6} + \dots$$

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} = \underbrace{\mathbb{1} + A + \frac{A^2}{2} + \frac{A^3}{6} + \dots}_{\text{konvergent?}}$$

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}\,t} e^{At} = Ae^{At}}$$

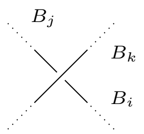
Beispiel 0.10 (Knoten- und ein sehr merkwürdiges Rechnen)**.** Wir rechnen jetzt nicht mehr in $\mathbb{R}$ (oder $\mathbb{Q}$) sondern in einem Bereich $\mathbb{F}_3$ mit nur 3 Zahlen:

$$\underline{0}, \underline{1} \text{ und } \underline{2} \qquad (\text{„modulo 3 rechnen“})$$

$+$	$\underline{0}$	$\underline{1}$	$\underline{2}$
$\underline{0}$	$\underline{0}$	$\underline{1}$	$\underline{2}$
$\underline{1}$	$\underline{1}$	$\underline{2}$	$\underline{0}$
$\underline{2}$	$\underline{2}$	$\underline{0}$	$\underline{1}$

$\cdot$	$\underline{0}$	$\underline{1}$	$\underline{2}$
$\underline{0}$	$\underline{0}$	$\underline{0}$	$\underline{0}$
$\underline{1}$	$\underline{0}$	$\underline{1}$	$\underline{2}$
$\underline{2}$	$\underline{0}$	$\underline{2}$	$\underline{1}$

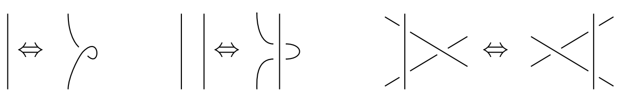
Wir ordnen jetzt einem Knoten ein Gleichungssystem zu:



$$\rightsquigarrow \boxed{2x_k - x_i - x_j = 0}$$

Eine solche Gleichung ist für $x_i, x_j, x_k \in \mathbb{F}_3$ genau dann erfüllt, wenn $x_i = x_j = x_k$ oder x_i, x_j, x_k paarweise verschieden sind.

Problem: Wie ändert sich das Gleichungssystem, wenn man den Knoten deformiert?
Reidemeisterbewegungen:

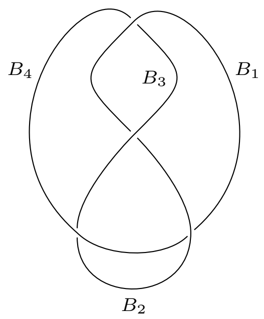


(R1) (R2) (R3)

- mehr Bögen, mehr Kreuzungen $\rightsquigarrow$ mehr Unbekannte, mehr Gleichungen
- aber Lösbarkeit verhält sich gleich!
- Es gibt immer die Lösung $x_1 = x_2 = \ldots = x_k$ (*„diagonale Lösung“*)
- Was zählt sind nicht-diagonale Lösungen.

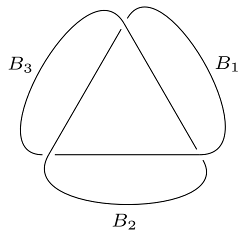
Man nennt einen Knoten *3-färbbar (p-färbbar)*, wenn das Gleichungssystem eine nicht-diagonale Lösung über $\mathbb{F}_3$ ($\mathbb{F}_p$) besitzt.

Achterknoten 4_1



$$
\begin{aligned}
2x_1 - x_3 - x_4 &= 0\\
2x_3 - x_1 - x_2 &= 0\\
2x_4 - x_2 - x_3 &= 0\\
2x_2 - x_1 - x_4 &= 0
\end{aligned}
$$

Kleeblattschlinge 3_1



$$
\begin{aligned}
2x_3 - x_1 - x_2 &= 0\\
2x_2 - x_1 - x_3 &= 0\\
2x_1 - x_2 - x_3 &= 0
\end{aligned}
$$

Lösungen über $\mathbb{R}$ für 3_1

$$
\begin{aligned}
\text{I:}\quad & 2x_3 - x_1 - x_2 && = 0\\
\text{II:}\quad & 2x_2 - x_1 - x_3 && = 0\\
\text{III:}\quad & 2x_1 - x_2 - x_3 && = 0
\end{aligned}
$$

$$
\begin{aligned}
\text{III} + 2\text{I:}\quad & -3x_2 + 3x_3 && = 0 \quad \Longrightarrow \quad x_2 = x_3\\
\text{in I:}\quad & -x_1 + x_2 && = 0 \quad \Longrightarrow \quad x_1 = x_2
\end{aligned}
$$

Also gibt es nur diagonale Lösungen.

Lösungen über $\mathbb{F}_3$ für 3_1

$$\text{III} + 2\text{I:}\quad -3x_2 + 3x_3 \quad = 0 \quad \Longrightarrow \quad x_2, x_3 \text{ beliebig}$$

$$x_2 = \underline{1}, \quad x_3 = \underline{2} \quad \overset{\text{in I}}{\Longrightarrow} \quad x_1 = -\underline{1} + \underline{2} \cdot \underline{2} = \underline{2} + \underline{1} = \underline{0}$$

Jetzt gibt es also auch nicht diagonale Lösungen.

1 Gleichungssysteme

1.1 Rechnen im $\mathbb{R}^n$

1.1.1 Algebraische Aspekte

Definition 1.1.1. Grundkörper $\mathbb{R}$ = *reelle Zahlen*

Definition 1.1.2. Der $\mathbb{R}^n = \mathbb{R} \times \ldots \times \mathbb{R}$ besteht aus Elementen, die man sich auf zwei Weisen vorstellt:

(i) Zahlentupel = Zeilenvektoren: $x = (x_1, \ldots, x_n)$

- x_i = Einträge, Koordinaten, Komponenten
  z.B. Punkte im $\mathbb{R}^3 : P = (x_1, x_2, x_3)$

(ii) Spaltenvektoren: $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

Für beide Auffasungen gibt es nun Operationen, wir notieren sie für einen Spaltenvektor:

- Addition: $x + y = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1+y_1 \\ \vdots \\ x_n+y_n \end{pmatrix}$
- Skalierung: $\lambda \in \mathbb{R} : \lambda \cdot x = \begin{pmatrix} \lambda\cdot x_1 \\ \vdots \\ \lambda\cdot x_n \end{pmatrix}$
- Nullvektor: $0 = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$

Unterschied: Wird bei der Matrizenmultiplikation deutlich werden.

Satz 1.1.3. ***Rechenregeln***
$x + (y + z) = (x + y) + z$
$x + y = y + x$
$x + 0 = x$
$-x = (-1) \cdot x$
$\lambda(\mu x) = (\lambda\mu)x$
$(\lambda + \mu) \cdot x = \lambda x + \mu x$

Definition 1.1.4. Skalarprodukt
$\langle\ ,\ \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$
Für $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ definiert man ein sogenanntes **Skalarprodukt**: $(x, y) \mapsto \langle x, y\rangle := x_1y_1 + \cdots + x_ny_n$
$\langle x, y\rangle$ ist linear in x (bei festem y) und linear in y (bei festem x). Das nennt man bilinear.

Satz 1.1.5. ***Rechenregeln***
$\langle x, y\rangle = \langle y, x\rangle$
$\langle x + x', y\rangle = \langle x, y\rangle + \langle x', y\rangle$
$\langle x, y + y'\rangle = \langle x, y\rangle + \langle x, y'\rangle$
$\langle 0, y\rangle = 0$
$\langle \lambda x, y\rangle = \lambda \cdot \langle x, y\rangle$
$\langle x, \mu y\rangle = \mu \cdot \langle x, y\rangle$
$(\forall y : \langle x, y\rangle = 0) \implies x = 0$

Beweis. Angenommen $x \neq 0 \Rightarrow \exists i : x_i \neq 0$.
Setze $y = \begin{pmatrix} 0 \\ \vdots \\ \mathrm{i} \\ \vdots \\ 0 \end{pmatrix}$, also $y_j = \begin{cases} 1 & j = i \\ 0 & j \neq i \end{cases}$
Dann gilt $0 = \langle x, y\rangle = x_i \cdot y_i = x_i$. Widerspruch. □

Das waren algebraische Aspekte (unabhängig von $\mathbb{R}$, gilt in jedem Körper $\mathbb{K}$ bzw. in seinem Spaltenvektorraum K^n)

1.1.2 Geometrische Aspekte

Definition 1.1.6.

- *Betrag*: $\|x\| := \sqrt{\langle x, x\rangle} = \sqrt{x_1^2 + x_2^2 + \ldots}$ Diese Formel ist die Verallgemeinerung des Satzes des Pythagoras im $\mathbb{R}^n$
- *Winkelmessung*: $\cos(x, y) = \frac{\langle x,y\rangle}{\|x\| \cdot \|y\|}$, $(x, y \neq 0)$
- *Matrix* = rechteckiges Zahlenschema $A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$
  Der erste Index zählt die Zeilen, der zweite Index zählt die Spalten.

  Spalte: $S_j(A)$ j-te Spalte
  Zeile: $Z_i(A)$ i-te Zeile
- *Nullmatrix* alle Einträge 0
- *Skalierung* von Matrizen: $\lambda \cdot A$ = jedes Element mit Lambda multiplizieren
- Addition, Subtraktion von Matrizen erfolgt komponentenweise:

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1n} \\ b_{21} & b_{22} & \ldots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \ldots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \ldots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \ldots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \ldots & a_{mn} + b_{mn} \end{pmatrix}$$

Die Matrix $A \in \mathbb{R}^{mn}$ kann als Riesenspaltenvektor oder Riesenzeilenvektor oder Mittelweg mit n Spalten und m Zeilen aufgefasst werden.

Das Produkt einer $(m \times n)$-Matrix A mit einem Spaltenvektor x liefert einen neuen Spaltenvektor:

$$A \cdot x = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n \end{pmatrix}$$

Satz 1.1.7. *Wir betrachten die Multiplikation mit einem festem A als Funktion als Funktion $T_A : \mathbb{R}^n \to \mathbb{R}^m, x \mapsto Ax$.*

- $T_A(0) = 0$
- $T_A(\lambda x) = \lambda T_A(x)$
- $T_A(x' + x'') = T_A(x') + T_A(x'')$

1.1.3 Lineare Gleichungssysteme

Definition 1.1.8. Ein lineares Gleichungssystem (LGS) ist ein System von m Gleichungen in n Unbekannten $x_1, \dots, x_n$ der Art:

$$\begin{aligned} &(G_1) && a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ &\vdots && \vdots \\ &(G_m) && a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_n \end{aligned}$$

Die a_{ij} mit $i = 1, \dots, n$ und $j = 1, \dots, m$ heißen Koeffizienten die $b_i = 1, \dots, m$ heißen Werte des LGS. Wichtig: alles ist durchnummeriert.

Kompakte Schreibweise
$A = (a_{ij})$ = Matrix der Koeffizienten
$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ = Spaltenvektor der Unbekannten
$b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ Spaltenvektor der Werte
$A \cdot x = b$

Notation:
Gegeben A und b
$(A|b)$ = erweiterte Matrix

Definition 1.1.9. Ein $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, der simultan alle Gleichungen löst, heißt Lösung des LGS.

$$\mathcal{L}(A|b) := x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n | Ax = b \subseteq \mathbb{R}^n$$

Bemerkung 1.1.10.
1. $A \in \mathbb{R}^{n \cdot m}, x \in \mathbb{R}^n, b \in \mathbb{R}^m$
2. Alles ist durchnummeriert
3. Es ist möglich, dass alle Koeffizienten $= 0$ sind: $A = 0$ ("triviales" LGS)
$\mathcal{L}(0|b) = \begin{cases} \mathbb{R}^n, & b = 0 \\ \emptyset, & b \neq 0 \end{cases}$
4. Es ist möglich, dass eine Zeile von A null ist.
$b_j = 0 \implies$ Diese Gleichung ist überflüssig, $b \neq 0 \Rightarrow \mathcal{L}(A|b) = \emptyset$
5. Es kann sein, dass eine Spalte von A null ist
$a_{1i} = a_{2i} = ... = 0 \implies x_i$ "überflüssig"

Definition 1.1.11. Ist $b = 0$, so heißt das LGS $(A|0)$ homogen . Ist $b \neq 0$ so heißt das LGS inhomogen.

Bemerkung 1.1.12. Was hat das alles mit $T_A : \mathbb{R}^n \to \mathbb{R}^m$ zu tun?
$x \mapsto y := A \cdot x$

1. $(A|b)$ hat eine Lösung (d.h. $\mathcal{L}(A|b) \neq \emptyset$) genau dann, wenn $b \in \text{Bild}(T_A) = T_A(\mathbb{R}^n) =$ Wertemenge gibt, es also ein $x \in \mathbb{R}^n$ mit $T_A(x) = b$ gibt.
2. $0 \in \mathcal{L}(A|0)$ (d.h. $\mathcal{L}(A|0) \neq \emptyset$)

Ein homogenes System hat immer mindestens eine Lösung, nämlich die triviale Lösung $x = 0$.

1.1.4 Struktur der Lösungsmenge

Satz 1.1.13. *(homogener Fall)*

(i) $(A|0)$ *hat immer die triviale Lösung*

(ii) $x \in \mathcal{L}(A|0) \implies \lambda x \in \mathcal{L}(A|0)$

(iii) $x, x' \in \mathcal{L}(A|0) \implies x + x' \in \mathcal{L}(A|0)$

Satz 1.1.14. *(inhomogener Fall)*

(i) $x, x' \in \mathcal{L}(A|b) \implies x - x' \in \mathcal{L}(A|0)$

(ii) Sei $\overline{x} \in \mathcal{L}(A|b)$

Dann gibt es zu jeder anderen Lösung $x \in \mathcal{L}(A|b)$ *genau eine Lösung* ζ *des homogenen Systems mit* $x = \overline{x} + \xi$

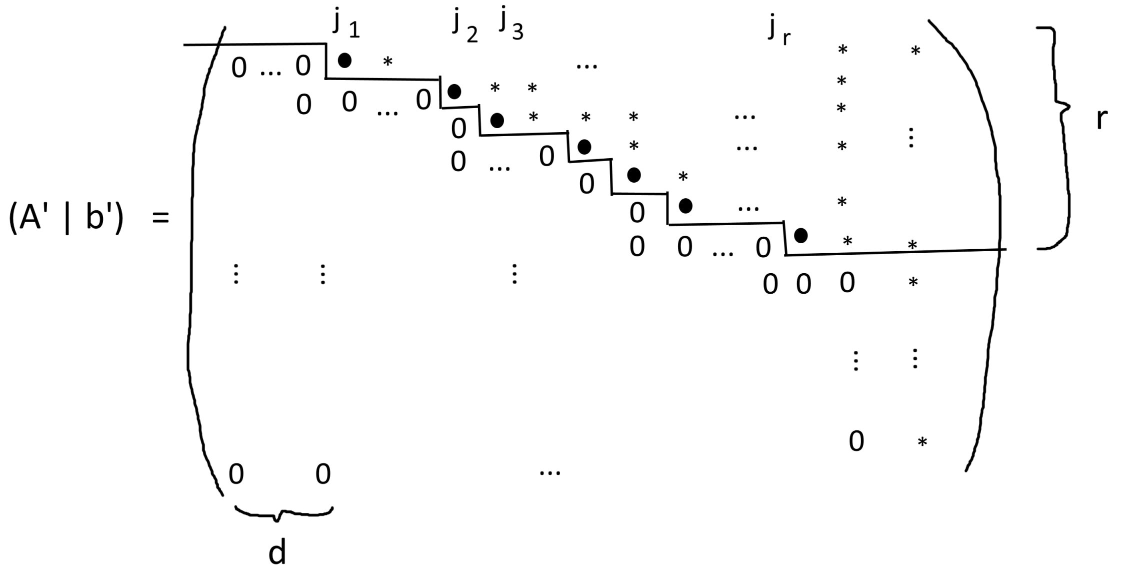
Beweis. (i) $x, x' \in \mathcal{L}(A|b) \implies Ax = b, Ax' = b$
$Ax - Ax' = b - b = 0$
$A(x - x') = 0$
Also $x - x' \in \mathcal{L}(A|0)$

(ii) Es gilt $A\overline{x} = b$
Ist $x \in L(A|b)$, also $Ax = b$, so haben wir nach (i) $x - \overline{x} = \xi \in \mathcal{L}(A|0)$
Also: $x = \overline{x} + \xi$
Und ξ ist eindeutig bestimmt: $x - \overline{x} = \xi$
$x - \overline{x} = \xi'$
$\implies \xi = \xi'$

□

1.1.5 Gauß-Algorithmus

Das Ziel des Gauß-Algorithmus ist es, ein LGS $Ax = b$ durch Zeilenumformungen in die sogenannte Zeilenstufenform zu bringen; d.h. am Ende soll unser LGS folgendermaßen aussehen:



Die Einträge • sind von Null verschieden, $*$ ist beliebig und alles unter der Treppenlinie ist Null. Die Größen d und r werden wichtige Charakteristiken sein.
Weil wir nur Zeilenumformungen verwenden werden, gilt:
$\mathcal{L}(A|b) = \mathcal{L}(A'|b')$
Und wir erhalten rekursiv eine Darstellung der Lösungsmenge.

Wir definieren zunächst den Begriff der Treppenfunktion.

Definition 1.1.15. Eine Treppenfunktion τ ist eine Funktion, die Spaltennummern $0, 1, ..., n$ auf Zeilennummern $0, 1, ..., m$ abbildet, sodass folgende Eigenschaften gelten:

- $\tau(0) = 0$
- $\tau(j) \leq \tau(j+1)$ $(1 \leq j \leq n)$ monoton
- $\tau(j) = \tau(j-1) oder \tau(j-1) + 1$ $(1 \leq j \leq n)$ Stufenhöhe ≤ 1

Eine Matrix $A \in Mat_{m,n}(\mathbb{R})$ ist in Zeilenstufenform, wenn es eine treppenfunktion τ gibt, sodass gilt:

- (ZSF 1) $a_{ij} = 0$ für $i > \tau(j) (1 \leq i \leq m, 1 \leq j \leq n)$
- (ZSF 2) $a_{ij} \neq 0$ für $i = \tau(j)$, falls $\tau(j) > \tau(j-1)$ $(1 \leq j \leq n)$

Bemerkung 1.1.16. Wir werden zwar den Gauß-Algorithmus auf die erweiterte Matrix $(A|b)$ anwenden, aber die Treppenfunktion bezieht sich nur auf A; wie werden sagen, $(A|b)$ sei in Zeilenstufenform, wenn A in Zeilenstufenform ist.

Bemerkung 1.1.17.
- Wir nennen die Stellen $1 \leq j_1 < ... < j_r \leq n$, an dem τ springt (also $\tau(j_k) = \tau(j_k - 1) + 1$ die Sprungstellen.
- Die Anzahl der Sprungstellen nennen wir den Rang $rg(A) = r$ der Matrix (oder auch Zeilenrang); es ist immer $0 \leq r \leq m, n$

Gauß-Algorithmus
Eingabe $(A|b)$
$A \in Mat_{m,n}(A), b \in \mathbb{R}^m$
Start: Setze $p \to 0, q \to 0$ (Hilfsvariablen)

1. Suche in den Spalten $S_j(A)$ in der Reihenfolge $j = q+1, ..., n$ einen Eintrag $a_{ij} \neq 0$ in der Reihenfolge $i = p+1, ..., m$.
   - Gibt es kein solchen a_{ij}, so gehe zu ENDE.
   - Sei $a_{ij} \neq 0$ mit j minimal unter $q < j \leq n$ und i minimal unter $p < i \leq m$
2. Vertausche Zeile i mit Zeile $p+1$
3. Subtrahiere von Zeile $i+1$ das $\frac{a_{p+1,q}}{a_{p,q}}$-fache der Zeile p.
   Subtrahiere von Zeile $i+2$ das $\frac{a_{p+2,q}}{a_{p,q}}$-fache der Zeile p.
   ...
   Subtrahiere von Zeile m das $\frac{a_{m,q}}{a_{p,q}}$-fache der Zeile p.
4. Setze $p \to p+1, q \to j$.
   - Ist $p = m$ oder $q = n$, gehe zu ENDE
   - Sonst gehe zu 1.

ENDE.
Ausgabe: $(A'|b')$ in Zeilenstufenform.

Bemerkung 1.1.18. Man nennt das a_{ij} in 1. ein Pivotelement .

Satz 1.1.19. *Jedes LGS $(A|b)$ wird durch den Gauß-Algorithmus in ein äquivalentes LGS $(A'|b')$ mit gleicher Lösungsmenge überführt.*

Beweis.
1. Zunächst ist offensichtlich, dass für $A = 0$ das LGS $(0|b)$ bereits in Zeilenstufenform ist, und zwar für die konstante Treppenfunktion $\tau(0) = \tau(1) = ... = \tau(n) = 0$. Wie führen den Beweis durch Induktion über m.
2. Für $m = 1$ und $A \neq 0$ sei $a_{11} = ... = a_{1j-1} = 0$ und $a_{1j} \neq 0$. Dann ist A bereits in Zeilenstufenform mit $\tau(0) = ... = \tau(j-1) = 0$ und $\tau(j) = ... = \tau(n) = 1$.
3. Für $m > 1$ führt ein erster Durchlauf des Gauß-Algorithmus für $A \neq 0$ an einem ersten Pivotelement $a_{i_1 j_1} \neq 0$. Nach Vertauschen und den Subtraktionen haben wir ein System und eine partiell definierte Treppenfunktion. Der zweite Durchlauf des Gauß-Algorithmus ist eigentlich ein erster Durchlauf auf das neue System. Dieses hat noch $m-1$ Zeilen. Nach Induktion wird es in Zeilenstufenform überführt.

$\square$

2 Körper und Vektorräume

2.1 Körper

2.1.1 Körperaxiome und Unterkörper

Definition 2.1.1. Ein *Körper* ist eine Menge $\mathbb{K}$ mit zwei Verknüpfungen $+$ (Addition),$\cdot$ (Multiplikation), die die folgenden Eigenschaften erfüllen:

- Addition $\mathbb{K} \times \mathbb{K} \to \mathbb{K}$
  1. $a + (b + c) = (a + b) + c$ (Assoziativität)
  2. $a + b = b + a$ (Kommutativität)
  3. Es gibt ein neutrales Element der Addition $\tilde{x} \in \mathbb{K}$, sodass für alle $x \in \mathbb{K}$ gilt $\tilde{x} + x = x$. Dieses bezeichnen wir mit 0.
  4. Zu jedem Element $x \in \mathbb{K}$ gibt es ein Element $-x \in \mathbb{K}$, sodass $x + (-x) = 0$. (Existenz von additiv Inversen)
- Multiplikation $\mathbb{K} \times \mathbb{K} \to \mathbb{K}$
  5. $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (Assoziativität)
  6. $a \cdot b = b \cdot a$ (Kommutativität)
  7. Es gibt ein neutrales Element der Multiplikation $\tilde{x} \in \mathbb{K}$, sodass für alle $x \in \mathbb{K}$ gilt $\tilde{x} \cdot x = x$. Dieses Element bezeichnen wir mit 1.
  8. Zu jedem Element $x \in \mathbb{K}$ gibt es ein Element $x^{-1} \in \mathbb{K}$, sodass $x \cdot x^{-1} = 1$. (Existenz von multiplikativ Inversen)
- Zusammenhang zwischen Addition und Multiplikation
  9. $a \cdot (b + c) = (a + b) \cdot c$ (Distributivität)
  10. $1 \neq 0$

Definition 2.1.2. Eine Teilmenge $\mathbb{K}' \subseteq \mathbb{K}$ eines Körpers $\mathbb{K}$ nennen wir *Unterkörper* , wenn sie die folgenden Eigenschaften erfüllt:

1. $x, y \in \mathbb{K}' \Rightarrow x + y \in \mathbb{K}', x \cdot y \in \mathbb{K}'$ (Abgeschlossenheit bezüglich Addition und Multiplikation)
2. $0, 1 \in \mathbb{K}'$
3. $x \in \mathbb{K}' \Rightarrow -x \in \mathbb{K}'$
4. $x \in \mathbb{K}' \Rightarrow x^{-1} \in \mathbb{K}'$

Bemerkung 2.1.3. Wie wir es aus der Schule gewohnt sind, schreiben wir vereinfachend $x - y$ statt $x + (-y)$ und $\frac{x}{y}$ statt $x \cdot y^{-1}$. Dies sind jedoch erstmal nur Notationen, da wir keine Verknüpfung $-$, oder $/$ definiert haben.

Beispiel 2.1.4. Beispiele für Körper, die uns allen geläufig sind sind die reellen Zahlen $(\mathbb{R}, +, \cdot)$, die rationalen Zahlen $(\mathbb{Q}, +, \cdot)$, oder auch die komplexen Zahlen $(\mathbb{C}, +, \cdot)$, welche wir im nächsten Abschnitt definieren werden. Ein Beispiel für einen Teilkörper von $\mathbb{R}$, der nicht der Körper rationalen Zahlen ist, ist $(\mathbb{Q}\left[\sqrt{2}\right], +, \cdot)$, wobei $\mathbb{Q}\left[\sqrt{2}\right] = \{a + b\sqrt{2} | a, b \in \mathbb{Z}\}$

2.1.2 Komplexe Zahlen

Definition 2.1.5. Der Körper der *komplexen Zahlen* $\mathbb{C}$ besteht aus der Menge $\mathbb{R}^2$ zusammen mit den Verknüpfungen

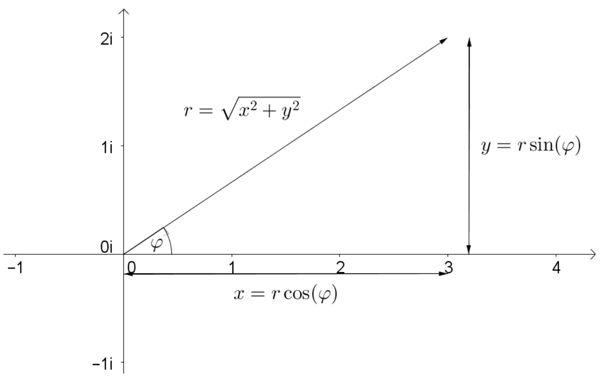
- $(x, y) + (x', y') = (x + x', y + y')$
- $(x, y) \cdot (x', y') = (x \cdot x' - y \cdot y', x \cdot y' + x' \cdot y)$

Es ist:

- $0_{\mathbb{C}} = (0, 0)$
- $1_{\mathbb{C}} = e_1 = (1, 0)$
- $-(x, y) = (-x, -y)$
- $(x, y)^{-1} = (\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2})$

wie sich jeweils leicht nachrechnen lässt. Des weiteren stellen wir fest, dass $(e_2)^2 = (0, 1)^2 = (-1, 0)$ und nennen e_2 die *imaginäre Einheit* i. Wir schreiben $(x, y) = x + iy$.
Dass für Addition und Multiplikation die gewünschten Eigenschaften (Assoziativität, Kommutativität, Distributivität) gelten lässt sich leicht nachrechnen.

Definition 2.1.6. $z = (x, y) \in \mathbb{C}$ kann als Zahl in der Gaußschen Zahlenebene interpretiert werden, die durch einen Winkel und eine Länge definiert ist.



mit $0 \leq \varphi < 2\pi$ und $r = \sqrt{x^2 + y^2}$ ist diese Darstellung eindeutig, wobei

$$x = r\cos(\varphi)$$
$$y = r\sin(\varphi)$$

r und φ nennen wir die *Polarkoordinaten* von z. Wir schreiben $z = (r; \phi)$

2.1.3 Endliche Körper

Bemerkung 2.1.7. Nachdem alle Körper, die uns bis hier hin begegnet sind unendlich viele Elemente hatten wollen wir nun versuchen zu einer Primzahl p einen Körper mit genau p Elementen, also einen endlichen Körper zu finden.

Definition 2.1.8. Man kann zeigen, dass es zu gegebenen $n, z \in \mathbb{Z}, n > 1$ eindeutige Zahlen $a, r \in \mathbb{Z}$ gibt mit $0 \leq r < n$, sodass $z = an+r$. Dies ist die Division durch n mit Rest r. Für ein gegebenes n wollen wir alle Zahlen als gleich auffassen, bei denen r gleich ist und sie in einer Menge zusammenfassen. Wir bezeichnen die Menge

$$[z] = \{z' \in \mathbb{Z} | z' = a'n + r, a' \in \mathbb{Z}\} = \{z' \in \mathbb{Z} | z' - z \text{ ohne Rest durch } n \text{ teilbar}\}$$

als die *Restklasse von z modulo n* . Und nennen $z' \in [z]$ *Repräsentant* der Restklasse .

2.1.4 Inverse in $\mathbb{F}_p$

Satz 2.1.9. *Es seien $a, b \in \mathbb{N}$ mit $0 < b < a$, dann gilt:*

1. *Der $ggT(a,b)$ ist der letzte nicht verschwindende Rest r_n der folgenden Kette von Divisionen-mit-Rest*
$$\begin{array}{lll} a = q_1 b + r_1 & \textit{mit} & 0 \le r_1 < b; q_1, r_1 \in \mathbb{N} \\ b = q_2 r_1 + r_2 & \textit{mit} & 0 \le r_2 < r_1; q_2, r_2 \in \mathbb{N} \\ r_1 = q_3 r_2 + r_1 & & \\ r_{n-2} = q_n r_{n-1} + r_n & \textit{mit} & 0 \le r_n < r_{n-1}; q_n, r_n \in \mathbb{N} \\ r_{n-1} = q_{n+1} r_n & \textit{mit} & q_{n+1} \in \mathbb{N} \end{array}$$

2. *Es gibt $\alpha, \beta \in \mathbb{Z}$ mit der Eigenschaft*
$$ggT(a,b) = \alpha a + \beta b.$$

Beweis.

1. Da die r_i strikt fallend, denn $b > r_1 > r_2 > ... > r_n > r_{n+1} = 0$, bricht der Algorithmus ab. Ist $t \in \mathbb{Z}$ ein Teiler von a und b, so ist t auch Teiler von r_1 und damit auch von r_2 usw. bis schließlich r_n. Umgekehrt gilt: Ist t Teiler von r_n, so auch von r_{n-1}, r_{n-2}, usw. bis a und b.

2. Löse die Gleichung rückwärts auf und setze ein.

□

2.1.5 Charakteristik eines Körpers

Bisher konnte man den Eindruck gewinnen als läge $\lambda \in \mathbb{N}$ in $\mathbb{K}$, da öfters schon so gerechnet wurde. Dies ist nicht so. Zwar gilt $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$, aber $\mathbb{N} \subset \mathbb{Z}$ sind keine Teilmenge von $\mathbb{F}_p$. Daher soll von jetzt an als Vereinbarung gelten, dass für $n \in \mathbb{Z}$ und $x \in \mathbb{K}$ gilt

$$nx := x + x + ... + x \text{ (n-mal)}.$$

Dies ist also nicht die Multiplikation von Körperelementen. Nun das Ganze etwas genauer in folgender

Definition 2.1.10. Für jedes $n \in \mathbb{Z}$ gilt:
$$\begin{aligned} \varepsilon_0 &:= 0 \\ \varepsilon_1 &:= 1 \\ \varepsilon_{n+1} &:= \varepsilon_n + 1 \qquad (n \ge 1) \\ \varepsilon_n &:= -\varepsilon_{-n} \qquad (n \le -1) \end{aligned}$$

Also
$$\varepsilon_n := n \cdot 1$$

Definition 2.1.11. Falls kein $\varepsilon_n = 0$ mit $n > 0$, so hat $\mathbb{K}$ die *Charakteristik* 0, und schreibt

$$char(\mathbb{K}) = 0.$$

Andernfalls nennt man das kleinste $n \ge 1$ mit $\varepsilon_n = 0$ die *Charakteristik* von $\mathbb{K}$, schreibe

$$char(\mathbb{K}) = n.$$

Beispiel 2.1.12.
1. $char(\mathbb{Q}) = char(\mathbb{R}) = char(\mathbb{C}) = 0$
2. $char(\mathbb{F}_p) = p$ (p ist eine Primzahl)

Bemerkung 2.1.13. Ist $char(\mathbb{K}) = 0$, so ist $\mathbb{Q}$ ein Unterkörper.
Denn
$$\begin{aligned} \mathbb{Q} &\longrightarrow \mathbb{K} \\ \frac{n}{m} &\longmapsto \varepsilon_n \cdot \varepsilon_n^{-1}, \end{aligned}$$

so ergibt sich $\varepsilon_a + \varepsilon_b = \varepsilon_{a+b}$ und $\varepsilon_a \cdot \varepsilon_b = \varepsilon_{ab}$ und somit erhält man ein $\mathbb{K}' = \{x - \varepsilon_n \cdot \varepsilon_m^{-1} | n, m \in \mathbb{Z} \wedge m \neq 0\}$ mit $\varepsilon_0 = 0$ und $\varepsilon_1 = 1$.

Definition 2.1.14. Eine Funktion $\phi : \mathbb{K} \longrightarrow \mathbb{K}'$ zwischen zwei Körpern heißt *Körperhomomorphismus* , falls gilt

1. $\phi(x + y) = \phi(x) + \phi(y)$
2. $\phi(x \cdot y) = \phi(x) \cdot \phi(y)$
3. $\phi(0) = 0$ und damit auch $\phi(-x) = -\phi(x)$
4. $\phi(1) = 1$ und damit auch $\phi(x^{-1}) = \phi(x)^{-1}$, falls ϕ injektiv und $x \neq 0$

Beispiel 2.1.15. $\mathbb{R} \hookrightarrow \mathbb{C}$ mit $x \longmapsto (x, 0)$ ist ein *Körperisomorphismus* auf einem Unterkörper.

Satz 2.1.16. *Ist char*$(\mathbb{K}) > 0$*, so ist char*$(\mathbb{K})$ *eine Primzahl.*

Beweis. Sei $char(\mathbb{K}) = n$ mit $n > 0$ und $n = a \cdot b$ mit $a, b \in \mathbb{N}$ sowie $1 < a,\ b < n$. Das heißt n ist keine Primzahl. Dann wäre $0 = \varepsilon_n = \varepsilon_a \cdot \varepsilon_b$ mit $\varepsilon_a, \varepsilon_b \neq 0$ und somit gäb es einen Nullteiler. □

2.2 Vektorräume

2.2.1 Vektorraumaxiome

Im folgenden Abschnitt ist mit $\mathbb{K}$ immer ein Körper gemeint.

Definition 2.2.1. Ein *Vektorraum* V über $\mathbb{K}$ ist eine Menge mit einer Verknüpfung und einer Operation.
Addition:

$$+ : V \times V \longrightarrow V \text{ mit } (x, y) \longmapsto x + y$$

Skalierung:

$$\cdot : \mathbb{K} \times V \longrightarrow V \text{ mit } (\lambda, x) \longmapsto \lambda \cdot x$$

mit den folgenden Eigenschaften:

1. $\forall x, y \in V : \quad x + y = y + x$
2. $\forall x, y, z \in V : \quad x + (y + z) = (x + y) + z$
3. $\exists 0 \in V\ \forall x \in V : \quad x + 0 = x = 0 + x$
4. $\forall x \in V \exists -x : \quad x + (-x) = 0 = (-x) + x$
5. $\forall x \in V,\ \forall \lambda, \mu \in \mathbb{K} : \quad \lambda \cdot (\mu \cdot x) = (\lambda \cdot \mu) \cdot x$
6. $\forall x \in V : \quad 1 \cdot x = x$
7. $\forall \lambda \in K,\ \forall x, y \in V : \quad \lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$
8. $\forall x \in V,\ \forall \lambda, \mu \in \mathbb{K} : \quad (\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$

Proposition 2.2.2. *1.* $0 \cdot x = 0$

2. $\lambda \cdot 0 = 0$

3. $(-1) \cdot x = -x$

4. $\forall \lambda \in \mathbb{K}\ \forall x \in V: \quad \lambda \cdot x = 0 \implies \lambda = 0$ *oder* $x = 0$

Beweis. Übungsaufgabe ☐

Beispiel 2.2.3.

1. $\mathbb{K}^n$ mit $n > 0$ ist ein Vektorraum. Denn die Addition ist definiert durch $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1+y_1 \\ \vdots \\ x_n+y_n \end{pmatrix}$ und die Skalierung durch $\lambda \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda \cdot x_1 \\ \vdots \\ \lambda \cdot x_n \end{pmatrix}$. Die Eigenschaften sind leicht nachzuprüfen.
2. Insbesondere ist $\mathbb{K}^0 = \{0\}$ der *triviale* Vektorraum und $\mathbb{K}^1$ Vektorraum über sich selbst.
3. Auch $Mat_{m,n}(\mathbb{K})$ mit festem $m, n \in \mathbb{N}$ ist ein Vektoraum. Auch hier lassen sich Addition und Skalierung sowie deren Eigenschaften leicht nachweisen.

Bemerkung 2.2.4. Vektorräume über $\mathbb{Q}$, $\mathbb{R}$ bzw. $\mathbb{C}$ nennt man *rational*, *reell* bzw. *komplex*.

2.2.2 Unterraum

Sei V im Folgenden ein $\mathbb{K}$-Vektorraum.

Definition 2.2.5. Eine Teilmenge $U \subset V$ heißt *Unterraum* von V, falls gilt

1. $u, u' \in U \implies u + u' \in U$ (abgeschlossen unter der Addition)
2. $\lambda \in \mathbb{K}\ u \in U \implies \lambda \cdot u \in U$ (abgeschlossen unter Skalierung)

Bemerkung 2.2.6. Sei $u \in U$ beliebig und $\lambda = 0$, dann folgt daraus $0 \cdot u = 0 \in U$.

Beispiel 2.2.7.

1. $U = 0 = \{0\}$ ist Unterraum in jedem Vektorraum.
2. $V = \mathbb{K}^n,\ i = 1, ..., n\ U_i := \{x = (x_1, ..., x_n) \in \mathbb{K}^n | x_i = 0\}$ ist Unterraum von V.
   $U_I := \{x = (x_1, ..., x_n) \in \mathbb{K}^n | x_i = 0 \text{ für } i \in I\}\ I \subset 1, ..., n$ ist Unterraum von V.
3. $V = \mathbb{K}^n \quad U = \{x = (x_1, ..., n) \in \mathbb{K}^n | x_1 + x_2 + ... + x_n = 0\}$ ist Unterraum von V.
4. $v \in V \quad U = \{tv | t \in \mathbb{K}\}$ ist Unterraum von V. Diesen Unterraum bezeichnet man auch als *Spann* bzw. *lineares Erzeugnis* von v. Manchmal nennt man ihn auch *lineare Hülle* von v. Doch dazu später mehr.
5. Sei $u, v, w \in V$, dann ist $U = Span(u, v, w) = \{x = \lambda u + \mu v + \nu w \in V | \lambda, \mu \nu \in \mathbb{K}\}$
6. $(A|0)$ sei homogenes LGS und $A \in Mat_{m,n}(\mathbb{K})$, dann ist $Ł(A|0) \subset \mathbb{K}^n$ ein Unterraum.
7. U_1, U_2 seien Unterräume von V, dann ist auch $U = U_1 \cap U_2$ ein Unterraum von V. Allgemeiner: $U_i \subset V\ i \in I \implies U := \bigcap_{i \in I} U_i$. I muss nicht endlich sein. **Achtung!** Die Vereinigung von Unterräumen sind im Allgemeinen keine Unterräume mehr. Als einfaches Beispiel dafür dienen hier die beiden Koordinatenachsen des $\mathbb{R}^2$ als Unterräume. Addiert man nämlich $(1, 0)$ aus dem x-Achsen-Unterraum zu $(0, 1)$ aus dem y-Achsen-Unterraum, so erhält man $(1, 1)$. $(1, 1)$ liegt aber nicht in der Vereinigung von den beiden Koordinatenachsen. **Achtung!** Auch das mengenmäßige Komplement ist im Allgemeinen kein Unterraum. Das orthogonale Komplement allerdings schon.
   Denn sei $V = \mathbb{R}^n$ und $n = \begin{pmatrix} n_1 \\ \vdots \\ n_n \end{pmatrix}$, dann ist $U = \{x \in \mathbb{R}^n | x_1 n_1 +, ..., +x_n n_n = 0\}$
8.

2.2.3 Körperwechsel

$$\begin{tikzcd}
\mathbb{K} \arrow[d, hook] & \times & \mathbb{L} \arrow[r] \arrow[d, Rightarrow] & \mathbb{L} \arrow[d, equal] \\
\mathbb{L} & \times & \mathbb{L} \arrow[r] & \mathbb{L}
\end{tikzcd}$$

$$\begin{tikzcd}
(\lambda, x) \arrow[r, mapsto] \arrow[d, mapsto] & \lambda \cdot x \arrow[d, equal] \\
(\lambda, x) \arrow[r, mapsto] & \lambda x
\end{tikzcd}$$

2.3 Lineare Abbildungen

Im folgenden bezeichne $\mathbb{K}$ einen Körper; V und W seien Vektorräume über $\mathbb{K}$ (Man sagt auch: $\mathbb{K}$-Vektorräume).

Definition 2.3.1. Eine Funktion $f : V \longrightarrow W$ ist eine *lineare Abbildung* , falls gilt:

(i) $f(0) = 0$

(ii) $f(x + y) = f(x) + f(y)$, mit $x, y \in V$

(iii) $f(\lambda x) = \lambda f(x)$, mit $\lambda \in \mathbb{K}, x \in V$.

Hierbei folgt (i) offensichtlich aus (iii): $f(0) = f(0 \cdot 0) = 0 \cdot f(0) = 0$.

Lineare Abbildungen sind die zentralen Objekte der linearen Algebra; man könnte diese auch als Studium der linearen Abbildungen bezeichnen. Wir betrachten nun einige Beispiele:

Beispiel 2.3.2. 1. Die Multiplikation einer Matrix A mit einem Vektor x, wie wir sie bereits direkt zu Beginn eingeführt haben, ist eine lineare Abbildung: $V = \mathbb{K}^n$; $W = \mathbb{K}^m$; $A \in \mathrm{Mat}_{m,n}(\mathbb{K})$
$f = T_A : \mathbb{K}^n \longrightarrow \mathbb{K}^m$, $x \longmapsto A \cdot x$

2. Für ein $A \in \mathrm{Mat}_{m,n}(\mathbb{K})$ haben wir ein lineares Gleichungssysteme studiert. Ein Ergebnis war $W = \mathcal{L}(A|0) \subseteq \mathbb{K}^n$, die Lösungsmenge im homogenen Fall ist ein Untervektorraum. Mit $r = \mathrm{rg}(A)$ und $l = n - r$ erhalten wir die bekannte Parameterdarstellung des Lösungsraums als lineare Abbildung:

$$\begin{aligned} f : \mathbb{K}^l &\longrightarrow \mathcal{L}(A|0) \\ u &\longmapsto L(u) \end{aligned}$$

Diese lineare Abbildung ist außerdem bijektiv. (Eine solche Abbildung nennt man auch *Isomorphismus.*)

3. Geometrische Beispiele: Drehung, Streckung, Scherung und Spiegelung im $\mathbb{R}^2$ lassen sich als Matrizenmultipliktion darstellen und sind somit lineare Abbildungen.

$$f = T_A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2, x \longmapsto A \cdot x$$

Die zur Drehung um ϕ assoziierte Matrix ist

$$A = \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix}$$

Die Matrix zur Streckung um λ_1 entlang der x-Achse und um λ_2 entlang der y-Achse ist:

$$A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}; \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} \lambda_1 \cdot x \\ \lambda_2 \cdot y \end{pmatrix}$$

Bei der Scherung bleibt eine Achse erhalten, in diesem Beispiel die x-Achse:

$$A = \begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix}; \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} x \\ \alpha x + y \end{pmatrix}$$

Zur Spiegelung an der y-Achse betrachte man

$$A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}; \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} -x \\ y \end{pmatrix}$$

die Spiegelung an der Winkelhalbierenden ist gegeben durch

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} y \\ x \end{pmatrix}$$

Auch die Projektion ist eine lineare Abbildung. Jeder Vektor $v \in V = \mathbb{R}^2$ kann eindeutig geschrieben werden als $v = w + w'$ mit $w \in W$ und einem Vektor w', der auf der Geraden liegt, die von $\tilde{w}'$ aufgespannt wird. Die Projektion $\pi : \mathbb{R}^2 = V \to W, w + w' = v \to w$ ist eine (surjektive!) lineare Abbildung.

4. Analytische Beispiele
Auf dem Vektorraum $V = C^{q+1}(\mathbb{R})$ der $(q+1)$-mal stetig differenzierbaren reellen Funktionen (mit der Addition $(f+g)(x) := f(x) + g(x), f, g \in V$) ist die Ableitung eine lineare Abbildung:

$$\begin{aligned} D : C^{q+1}(\mathbb{R}) &\longrightarrow C^q(\mathbb{R}) \\ f &\longmapsto f' \end{aligned}$$

Hierbei bezeichnet $f'(x) = \frac{d}{dx} f(x)$ die Ableitung von f nach x. Wegen der bekannten Ableitungsregeln, dass die Ableitung der Summe gleich der Summe der Ableitungen ist, und dass die Ableitung einer skalierten Funktion gleich der skalierten Ableitung ist, sind Ableitungen lineare Abbildungen. Insbesondere gilt für beliebig oft differenzierbare Funktionen ($q = \infty$):

$$\begin{aligned} D : C^{\infty}(\mathbb{R}) &\longrightarrow C^{\infty}(\mathbb{R}) \\ f &\longmapsto f' \end{aligned}$$

Auf dem Vektorraum $V = L^1([0,1])$ der über $[0,1]$ integrierbaren Funktionen ist die Integration eine lineare Abbildung.

$$\begin{aligned} I : L^1([0,1]) &\longrightarrow \mathbb{R} \\ f &\longmapsto \int_b^a f(x)dx \end{aligned}$$

mit $a, b \in [0,1]$. Hier gilt die analoge Regel, dass das Integral einer Summe gleich der Summe der Integrale ist.
Betrachten wir als letztes Beispeil $V = \mathcal{F}_{kov}$, den Vektorraum der reellen konvergenten Folgen $a = (a_0, a_1, \ldots)$. Die Grenzwertbildung ist eine lineare Abbildung:

$$\begin{aligned} \lim : V = \mathcal{F}_{konv} &\longrightarrow \mathbb{R} \\ a = (a_n) = (a_0, a_1, \ldots) &\longmapsto \lim_{n\to\infty} a_n \end{aligned}$$

Die Summe zweier konvergenter Folgen ist wieder eine konvergente Folge, also ist die Grenzwertbildung eine lineare Abbildung:

$$\begin{aligned} (a_n) &\longmapsto \lim a_n, \ b_n \longmapsto \lim b_n \\ (a_n) + (b_n) &\longmapsto \lim a_n + \lim b_n. \end{aligned}$$

Proposition 2.3.3. *Es seien U, V, W Vektorräume über $\mathbb{K}$, dann gilt:*

(i) Die Identität $\mathrm{id}_V : V \longrightarrow V, x \longmapsto x$ ist linear.

(ii) Sind $f : U \to V$ und $g : V \to W$ lineare Abbildungen, dann auch die Verknüpfung $g \circ f : U \to W$.

(iii) Die Nullabbildung $0 : V \to W, x \mapsto 0$ ist eine lineare Abbildung.

(iv) Sind f, f' lineare Abbildungen, dann ist auch λf, definiert durch $(\lambda f)(x) = \lambda f(x)$, mit $\lambda \in \mathbb{K}$, wieder eine lineare Abbildung.

(v) *Genauso ist auch* $f + f'$ *mit* $(f + f')(x) = f(x) + f'(x)$ *eine lineare Abbildung. Beides folgt einfach aus der Tatsache, dass* f *und* f' *lineare Abbildungen sind (*f' *ist* nicht *die Ableitung von* f*!).*

(vi) *Sei* $f : V \longrightarrow W$ *eine lineare bijektive Abbildung. Dann ist auch die Umkehrabbildung* $f^{-1} : W \to V$ *linear.*

Beweis von Proposition (vi). Zu zeigen: $f^{-1}(x+y) = f^{-1}(x)+f^{-1}(y)$. Seien $x, y \in W$. Da f bijektiv ist gilt: $\exists! x', y' \in V : x = f(x'), y = f(y')$; „$\exists!$“ bedeutet „es gibt eindeutig“. Dann folgt mit der Linearität von f:

$$f^{-1}(x + y) = f^{-1}(f(x') + f(y')) = f^{-1}(f(x' + y') = x' + y' = f^{-1}(x) + f^{-1}(y)$$

Damit haben wir die Bedingung (ii) aus Definition 2.3.1 verfiziert. Wir zeigen nun Bedingung (iii): Sei $\lambda \in \mathbb{K}$. Zu zeigen: $f^{-1}(\lambda x) = \lambda f^{-1}(x)$. Auch dies folgt mit x' wie oben aus der Linearität von f:

$$f^{-1}(\lambda x) = f^{-1}(\lambda f(x')) = f^{-1}(f(\lambda x')) = \lambda x' = \lambda f^{-1}(x)$$

$\square$

Definition 2.3.4. Sei $f : V \to W$ eine lineare Abbildung zwischen zwei Vektorräumen.

(i) Die Teilmenge von W $\mathrm{im}(f) = f(V) := \{x \in W | x = f(x') \text{ für } x' \in V\} \subseteq W$ heißt *Bild von* f (engl. *image*).

(ii) Die Teilmenge von V $\ker(f) = f^{-1}(0) := \{x \in V | f(x) = 0\} \subseteq V$ heißt *Kern von* f (engl. *kernel*).

Proposition 2.3.5. *Sei* $f : V \to W$ *eine lineare Abbildung zwischen zwei Vektorräumen.*

(i) *Das Bild* $\mathrm{im}(f)$ *ist ein Untervektorraum in* W*.*

(ii) *Der Kern* $\ker(f)$ *ist ein Untervektorraum in* V*.*

(iii) *Ist* $b \in \mathrm{im}(f)$ *und etwa* $f(\xi) = b$*, so ist* $f^{-1}(b)$ *ein affiner Unterraum und zwar* $f^{-1}(b) = \ker(f) + \xi$*.*

Ein Spezialfall von (iii) ist, dass die Lösungsmenge eines linearen Gleichungssystems im inhomogenen Fall ein affiner Unterraum ist.

Beweis. (i) Seien $x, y \in \mathrm{im}(f)$, sodass $x = f(x'), y = f(y')$ für $x', y' \in V$. Daraus folgt, dass $x + y = f(x') + f(y') = f(x' + y')$. Also liegt $x + y$ im Bild von f.

(ii) Seien $x, y \in V$ mit $f(x) = f(y) = 0$, also $x, y \in \ker(f)$. Dann gilt $f(x + y) = f(x) + f(y) = 0$. Damit liegt auch $x + y$ im Kern von f.

(iii) Sei $b \in \mathrm{im}(f)$ mit $f(\xi) = b$. Für $x \in V$ mit $f(x) = b$ gilt:

$$f(x - \xi) = f(x) - f(\xi) = b - b = 0$$

Also liegt $x - \xi$ im Kern von f.

$\square$

Beispiel 2.3.6. Betrachten wir wieder die Matrizenmultipliktion: $f = T_A : \mathbb{K}^n \to \mathbb{K}^m, x \mapsto T_A(x) = A \cdot x$ mit $A \in \mathrm{Mat}_{m,n}(\mathbb{K})$.

- $\ker(T_A) = T_A^{-1}(0) = \mathcal{L}(A|0)$. Der Kern ist der Lösungsraum des assoziierten linearen Gleichungssystems im homogenen Fall.

- $\mathrm{im}(T_A) = \{b \in \mathbb{K}^m | \mathcal{L}(A|b) \neq \}$. Das Bild ist die Menge der Vektoren, für die das assoziierte lineare Gleichungssystem im inhomogenen Fall lösbar ist.

- Falls $b \in \mathrm{im}(T_A)$ mit etwa $T_A(\xi) = b$, so ist $T_A^{-1}(b) = \mathcal{L}(A|b) = \mathcal{L}(A|0) + \xi$, der Lösungsraum des inhomogenen und lösbaren Falls.

Beispiel 2.3.7. Die *komplexe Konjugation* $^{-}$ ist ein Beispiel für eine Abbildung, die $\mathbb{R}$-linear, aber nicht $\mathbb{C}$-linear ist.

$$\begin{gathered} ^{-}:\mathbb{C}\to\mathbb{C} \\ z=x+iy\mapsto\overline{z}=x-iy \end{gathered}$$

Die komplexe Konjugation ist ein Körperisomorphismus, denn

$$\overline{0}=0;\ \overline{1}=1;\ \overline{z_1+z_2}=\overline{z_1}+\overline{z_2};\ \overline{z_1z_2}=\overline{z_1}\cdot\overline{z_2}\,\text{(Multiplikativität)}.$$

$\mathbb{C}$-Linearität würde bedeuten $\overline{z_1z_2}=z_1\cdot\overline{z_2}$. Die Abbildung kann also nicht sowohl ein Körperisomorphismus sein als auch $\mathbb{C}$-linear. Da die komplexe Konjugation multiplikativ ist, ist sie lediglich $\mathbb{R}$-linear:

$$\overline{\lambda z}=\lambda\overline{z}\Leftrightarrow\lambda=\overline{\lambda}\Leftrightarrow\lambda\in\mathbb{R}=\mathrm{Fix}(^{-})$$

Definition 2.3.8. Seien V,W Vektorräume über $\mathbb{K}$. Eine Funktion $f:V\to W$ heißt *affin* , falls es eine linere Abbildung $\varphi:V\to W$ gibt, und ein $b\in W$ mit $f(x)=\varphi(x)+b$.

Beispiel 2.3.9.

1. Die Translation $\tau_b:W\to W$ mit $0\neq b\in W,w\mapsto w+b$ ist immer eine affine Abbildung
2. Die Abbildung $f(x)=mx+b$ mit $m\in\mathbb{K}$ ist eine affine Abbildung. *Achtung!:* In der Schule und der Analysis werde solche Funktionen als lineare Funktionen bezeichnet. In der linearen Algebra sind dies aber im allgemeinen affine Funktionen!
3. Wir betrachten die Parameterdarstellung der Lösungsmenge $\mathcal{L}(A|b)$ eines Linearen Gleichungssystems $(A|b)$ der Größe $m\times n$, wobei $A\in Mat_{m,n}(\mathbb{K}),b\in\mathbb{K}^n,r=rg(A),l=n-r$
   Es sei L^0 die Parameterdarstellung für den homogenen Fall, also
   $$L^0:\mathbb{R}^l\to\mathcal{L}(A|b);U=(U_0,U_1,\dots,U_r)\mapsto L^0(U)=(U_0,L_1,U_1,L_2,U_2,\dots,L_r,U_r)$$
   wobei $L_1=L_1(U_1,U_2,\dots,U_r),L_2=L_2(U_2,U_3,\dots,U_r),\dots,L_r=L_r(U_r)$ Nun suchen wir uns irgend eine Lösung $\xi\in\mathcal{L}A|b$ des inhomogenen Gleichungssystems. Für den Lösungsraum des inhomogenen Gleichungssystems gilt nun:
   $$L=L^0+\xi;L(U)=L^0(U)+\xi$$
4. Es sei f eine affine Abbildung. Offenbar ist $f(0)=b$ die Translationskonstante, während $f(x)-b=\Phi(x)$ der so genannte lineare Anteil ist. Damit hat man für affine Abbildungen folgende Gleichungen:
   $$\begin{aligned} f(\lambda x)-b=&\lambda(f(x)-b) \\ f(x+y)-b=&(f(x)-b)+(f(y)-b)=f(x)+f(y)-2b \end{aligned}$$
5. Ist f eine lineare Abbildung, so ist $Im(f)\subseteq W$ ein linearer Unterraum.
   Ist f eine affine Abbildung, so ist $Im(f)\subseteq W$ ein affiner Unterraum.

Proposition 2.3.10. *Es sei $\mathbb{K}\subseteq\mathbb{L}$ ein Unterkörper von $\mathbb{L}$ (genau wie $\mathbb{Q}\subseteq\mathbb{R}\subseteq\mathbb{C}$) und V ein Vektorraum über Ł. Wir wissen, dass V dann auch ein $\mathbb{K}$-Vektorraum ist. Es kommt aber vor, dass eine Funktion $f:V\to W$ zwar linear ist, wenn man V als $\mathbb{K}$-Vektorraum betrachtet, es jedoch nicht ist, wenn man V als Ł-Vektorraum betrachtet.*

Beispiel 2.3.11. Ein Beispiel für die eben genannten Funktionen ist die komplexe Konjugation über $\mathbb{R}$ und $\mathbb{C}$.

$$\_:\mathbb{C}\to\mathbb{C};x+iy=z\mapsto\overline{z}=x-iy$$

Diese Abbildung ist $\mathbb{R}$-Linear, da man leicht nachrechnen kann, dass $\lambda(x,y)\mapsto(x,-y)$. Sie ist jedoch nicht $\mathbb{C}$-Linear, wie man leicht am Beispiel der Multiplikation mit i sieht. Es ist:

$$i\cdot\overline{(x,y)}=i\cdot(x,-y)=(x,y)\neq(-x,-y)=\overline{(-y,x)}=\overline{i\cdot(x,y)}$$

Abgesehen von der Linearität bei Multiplikation mit einer komplexen Zahl besitzt die komplexe Konjugation aber viele nützliche Eigenschaften. Für $z_1, z_2 \in \mathbb{C}$ gilt nämlich

$$\begin{aligned}
&\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2} \\
&\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2} \\
&\overline{0} = 0, \overline{1} = 1 \\
&z^{-1} = \frac{\overline{z}}{|z|^2}
\end{aligned}$$

2.3.1 Morphismen

Definition 2.3.12. Es sei $f : V \to W$ eine lineare Abbildung, dann heißt f
Monomorphismus, wenn f injektiv
Epimorphismus, wenn f surjektiv
Isomorphismus, wenn f bijektiv Endomorphismus, wenn $V = W$ Automorphismus, wenn $V = W$, f bijektiv

Proposition 2.3.13. *Es seien $f : V \to W$, $g : W - U$ linear. Dann gilt*

- *f ist ein Monomorphismus $\Leftrightarrow$ $ker(f) = 0$*
- *$g \circ f$ ist ein Monomorphismus $\Leftrightarrow$ f ist ein Monomorphismus*
- *$g \circ f$ ist ein Epimorphismus $\Leftrightarrow$ f ist ein Epimorphismus*

3 Basen und Dimension

3.1 Erzeugendensystem

Definition 3.1.1. Es sei V ein $\mathbb{K}$-Vektorraum, $v, e_1, e_2, ..., e_n \in V$, $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$ und

$$v = \lambda_1 e_1 + \ldots + \lambda_n e_n \qquad (*)$$

Wir nennen $(*)$ eine Linearkombination (der $e_1, \ldots, e_n$ über $\mathbb{K}$), oder eine lineare Darstellung von v (durch $e_1, \ldots, e_n$ über $\mathbb{K}$). Ist $v = 0$, so nennt man $(*)$ eine lineare Relation zwischen den $e_1, \ldots, e_n$ (über $\mathbb{K}$). Sind alle $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$, so nennt man $(*)$ eine triviale Linearkombination, oder triviale Darstellung von $v = 0$, oder eine triviale Relation.

Definition 3.1.2. Es sei $\xi \subseteq V$ eine Teilmenge von V. Dann heißt

$$Span(\xi) := \{v \in V | \text{Es gibt endlich viele } e_1, \ldots, e_n \in \xi \text{ und } \lambda_1, ..., \lambda_n \in \mathbb{K} \text{ so dass } v = \lambda_1 e_1 + \ldots + \lambda_n e_n\}$$

Der Spann, oder das lineare Erzeugnis von ξ in V. Wir vereinbaren, dass $Span(\emptyset) = \{0\} = 0$.

Beispiel 3.1.3.
1. Ist $\xi = \{0\}$, so ist $Span(\xi) = 0$
2. Ist $u \in V, \neq 0$, so ist $Span(\{u\})$ die von u aufgespannte Gerade.
3. Sind $u_1, u_2 \neq 0$ nicht kollinear, also weder $u_1 \in Span(u_2)$ noch $u_2 \in Span(u_1)$, so ist $Span(u_1, u_2)$ eine Ebene.
4. Ist $V = \mathbb{K}^n$, so nennen wir $e_i = (0, 0, \ldots, \underset{\text{i-te Stelle}}{1}, \ldots, 0)$ den i-ten Einheitsvektor. Für $0 \leq m \leq n$ ist $\xi_m = \{e_1, \ldots, e_m\}$ und $Span(\xi_m) = \mathbb{K}^m \subseteq \mathbb{K}^n$, wobei einfach die letzten $n - m$ Koordinaten 0 sind.
5. Ist $V = \mathbb{R}$, $\mathbb{K} = \mathbb{Q}$, $\xi = \{1, \sqrt{2}\}$, so ist $Span(\xi) = \{a + b\sqrt{2} | a, b \in \mathbb{Q}\} = \mathbb{Q}\left[\sqrt{2}\right]$
6. Sind $U_i \subseteq V, i \in I$ Untervektorräume eines Vektorraumes V und ist $\xi = \bigcup\limits_{i \in I} U_i$, so nennen wir $Span(\xi) = \underset{i \in I}{+} U_i$ die Interne Summe der U_i

Proposition 3.1.4. *Es Sei $\xi \subseteq V$ eine Teilmenge von V*

1. $0 \in Span(\xi)$
2. $\xi \subseteq Span(\xi)$ *insbesondere ist* $Span(V) = V$
3. $Span(U) = U$, *falls U ein UVR ist*
4. $Span(Span(U)) = Span(U)$
5. $Span(\xi)$ *ist der kleinste UVR von V, der ξ enthält.*
6. $Span(\xi)$ *ist der Durchschnitt aller UVR von V welche ξ enthalten.*

Definition 3.1.5. Eine Teilmenge $\mathcal{E}$ heißt *Erzeugendensystem von V*, falls gilt: $Span(\mathcal{E}) = V$.
V heißt *endlich erzeugt*, falls es ein endliches Erzeugendensystem gibt.

Uns interessiert dabei jedoch nur das kleinste Erzeugendensystem, beziehungsweise das $minimale Erzeugendensystem$.

Lemma 3.1.6. *Ist $\mathcal{E}$ ein EZS von V und $0 \neq \lambda \in \mathbb{K}, e \in \mathcal{E}$, dann ist auch $\mathcal{E}' = (\mathcal{E} - \{e\}) \cup \{\lambda \cdot e\}$ mit*
$\mathcal{E} = \{e_1, \ldots, e_i, \ldots\}$
$\Downarrow \mathfrak{M}_i[\lambda]$
$\mathcal{E}' = \{e_1, \ldots, \lambda e_i, \ldots\}$

Beweis. $v = \lambda_1 e_1 + \cdots + \lambda_n e_n$
$v = \lambda_1 e_1 + \cdots + \frac{\lambda_i}{\lambda}\lambda_i e_i + \cdots + \lambda_n e_n$

□

Lemma 3.1.7. *Sei $\mathcal{E} = \{e_1, e_2, \ldots, e_i, \ldots, e_j, \ldots\}$ ein EZS. Dann ist $\mathcal{E}' = (\mathcal{E} - \{e_i\}) \cup \{e_i + e_j\}$ wieder ein EZS.*
$\mathcal{E} = \{e_1, \ldots, e_i, \ldots, e_j, \ldots\}$
$\Downarrow \mathfrak{A}_j$
$\mathcal{E}' = \{e_1, \ldots, e_i + e_j, \ldots, e_j, \ldots\}$

Beweis. $v = \lambda_1 e_1 + \cdots + \lambda_i e_i + \cdots + \lambda_n e_n$
$= \lambda_1 e_1 + \cdots + \lambda_i(e_i + e_j) + \cdots + (\lambda_j - \lambda_i)e_j + \cdots + \lambda_n e_n$

□

Lemma 3.1.8. *Sei $\mathcal{E}$ ein EZS von V. Es gäbe in $\mathcal{E}$ eine nicht-triviale Relation, dh. es gibt $e_i, \ldots, e_n \in \mathcal{E}, \lambda_1, \ldots, \lambda_n \in \mathbb{K}$ wobei nicht alle $\lambda_i = 0$ und es gelte $0 = \lambda_1 e_1 + \cdots + \lambda_i e_i + \cdots + \lambda_n e_n$ und etwa $\lambda_i \neq 0$. Dann ist $\mathcal{E}' = \mathcal{E} - \{e_i\}$ wieder ein EZS von V.*

Beweis. Es gilt:
$e_i = -\frac{1}{\lambda_i}(\lambda_1 e_1 + \cdots + \underbrace{\widehat{\lambda_i e_i}}_{[*]} + \cdots + \lambda_n e_n)$
$[*] = -\frac{1}{\lambda_i}(\lambda_1 e_1 + \cdots + \lambda_{i-1}e_{i-1} + \lambda_{i+1}e_{i+1} + \cdots + \lambda_n e_n)$
d.h. $e_i \in Span(\mathcal{E}) = Span(\mathcal{E} - e_i)$
Also: wenn immer e_i in einer Darstellung eines $v \in V$ vorkommt gilt:
$v = \alpha_{i_1} e_{i_1} + \cdots + \alpha_{i_k} e_{i_k} + \cdots + \alpha_{i_m} e_{i_m}$
$\downarrow$
$= \alpha_{i_1} e_{i_1} + \cdots + \alpha[*] + \cdots + \alpha_{i_m} e_{i_m}$

□

3.2 Lineare Unabhängigkeit

V = Vektorraum über $\mathbb{K}$
$\mathcal{B} \subseteq V$ Teilmenge

Definition 3.2.1. $\mathcal{B}$ heißt *linear unabhängig* in V, falls gilt:
$b_1, \ldots, b_n \in \mathcal{B}$ sind paarweise verschieden und $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$.
Wenn $\lambda_1 b_1 + \cdots + \lambda_n b_n = 0$ gilt, dann nur deshalb, weil $\lambda_1 = \cdots = \lambda_n = 0$.

Bemerkung 3.2.2. $\mathbb{B}$ ist NICHT "von" etwas unabhängig! Es IST linear unabhängig.

Bemerkung 3.2.3. Oder: Jedes b ist VON ALLEN ANDEREN linear unabhängig.
($\Leftrightarrow b$ kann nicht als Linearkombination der anderen geschrieben werden $\Leftrightarrow b \notin Span(\mathcal{B}, \{b\})$)

Bemerkung 3.2.4. Man sagt: Es gibt keine nicht-triviale Relation unter den Elementen von $\mathcal{B}$.

Bemerkung 3.2.5. Die Eigenschaft linearer Unabhängigkeit kommt in der Menge $\mathcal{B}$ vor, nicht in ihren Elementen.

Bemerkung 3.2.6. v ist linear unabhängig von $\mathcal{B} \Leftrightarrow v \notin Span(\mathcal{B})$.
v ist linear abhängig von $\mathcal{B} \Leftrightarrow v \in Span(\mathcal{B})$.

Bemerkung 3.2.7. $0 \notin \mathcal{B}$, falls $\mathcal{B}$ linear unabhängig.
Denn sonst $0 = \lambda \cdot 0, \lambda \neq 0$ z.B. für $\lambda = 1$

Bemerkung 3.2.8. Wir setzen $\mathcal{B} = \emptyset$ ist linear unabhängig.

Beispiel 3.2.9. $\mathcal{B} = \{b\}$ ist linear unabhängig $\Leftrightarrow b \neq 0$. Jede Relation müsste lauten $\lambda \cdot b = 0$

Beispiel 3.2.10. $\mathcal{B} = \{b_1, b_2\}$ ist linear unabhängig $\Leftrightarrow b_1, b_2 \neq 0$ mit b_1, b_2 nicht kolinear
$b_1 = \lambda_2 b_2$
$b_2 = \lambda_1 b_1$
Beides ist nicht möglich, da $\mu_1 b_1 + \mu_2 b_2 = 0$

Beispiel 3.2.11. $V = \mathbb{K}$, $\mathcal{B} = \{e_1, \dots, e_m\}$, $m \leq n$, Standardeinheitsvektoren $e_i = (0, \dots 0, 1, 0, \dots, 0)$, wobei 1 an i-ter Stelle steht.
$\mathcal{B}$ ist linear unabhängig.

Angenommen $\lambda_1 e_1 + \dots + \lambda_m e_m = 0 \Rightarrow \lambda_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dots + \lambda_m \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 0$

1. Koordinate
$\lambda_1 \cdot 1 + \lambda_2 \cdot 0 + \dots + \lambda_m \cdot 0 = 0 \Rightarrow \lambda_1 = 0$
2. Koordinate
$\lambda_1 \cdot 0 + \lambda_2 \cdot 1 + \dots + \lambda_m \cdot 0 = 0 \Rightarrow \lambda_2 = 0$
$\vdots$
m. Koordinate
$\lambda_1 \cdot 0 + \dots + \lambda_m \cdot 1 = 0 \Rightarrow \lambda_m = 0$

Satz 3.2.12. *$\mathcal{B} \subset V$ ist genau dann linear unabhängig, wenn jedes $v \in Span(\mathcal{B})$ genau eine (oder eine eindeutige) Darstellung als Linearkombination besitzt.*
$v = \lambda_1 b_1 + \dots + \lambda_n$ für ein geeignetes $b_1, \dots, b_n \in \mathcal{B}$, und $\lambda_1, \dots, \lambda_n \in \mathbb{K}$.

Beweis. Angenommen wir hätten zwei Darstellungen
$v = \lambda_{i_1} b_{i_1} + \dots + \lambda_{i_n} b_{i_n}$
$v = \lambda_{j_1} b_{j_1} + \dots + \lambda_{j_m} b_{j_m}$
Dann gilt:
$0 = v - v = \lambda_{i_1} b_{i_1} + \dots + \lambda_{i_n} b_{i_n} - \lambda_{j_1} b_{j_1} + \dots + \lambda_{j_m} b_{j_m}$

Wir fassen zusammen und kürzen weg. Deshalb können wir folgendes annehmen:
alle b_{i_r} und b_{j_r} sind paarweise verschieden und alle λ_{i_r} und $\lambda_{j_r} \neq 0$

Sind die Darstellungen verschieden, so gibt es mindestens ein $b_{i_k} = b_{j_l}$, aber $\lambda_{i_k} \neq \lambda j_l$.
Dann ist die rechte Seite eine nicht triviale Relation. Aber $\mathcal{B}$ war als linear unabhängig angenommen. Widerspruch! $\square$

$V = P_n(\mathbb{K}) \ni p(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n$
$a_i \in \mathbb{K}$ $\quad p'(x) := a_1 + 2a_2 x + \dots + n a_n x^n$ ist die formale Ableitung
$W = P_{n-1}(\mathbb{K})$ $\quad P_\infty(\mathbb{K}) = \mathbb{K}[x]$
$P_\infty(\mathbb{K}) = \mathbb{K}[x]$

$\mathcal{B} = \{1, x, x^2, \dots, x^n\} \subseteq P_n(\mathcal{K})$ sind Stammpolynome. $\mathcal{B}$ ist ein EZS und linear unabhängig.

3.3 Basen

Definition 3.3.1. Eine Teilmenge $\mathcal{B} \subseteq V$ eines VR-V über $\mathbb{K}$ heißt *Basis von V*, falls $\mathcal{B}$ ein EZS und linear unabhängig ist.

Bemerkung 3.3.2. Eine Basis ist ein minimales EZS und ein maximales linear unabhängiges System.

Beispiel 3.3.3. $V = \mathbb{K}^n$, $\mathcal{B} = \{e_1, \dots, e_n\}$

Beispiel 3.3.4. $V = P(\mathbb{K}) : \mathcal{B} = 1, x, x^2, \dots, x^n$
$V = \mathbb{C}/\mathbb{C} \qquad \mathcal{B} = \{1\}$
$V = \mathbb{C}/\mathbb{R} \qquad \mathcal{B} = \{e_1, e_2\}$
$V = \mathbb{C}/\mathbb{K} = \mathbb{Q} \qquad \mathcal{B} = ?$
$\mathcal{B} = \{\log 2, \log 3, \log 5, \dots\}$ ist ein EZS und somit linear unabhängig.

Beispiel 3.3.5. Anwendung: $(A \mid 0)$ = homogenes LGS, $m \times n$ über $\mathbb{K}$
Annahme: Die Matrix ist bereits in ZSF

$$\begin{pmatrix}
0 & 0 & * & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & * \\
0 & * & * & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & * \\
\vdots & 0 & * & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & * \\
\vdots & \vdots & 0 & \cdots & 0 & * & \cdots & \cdots & \cdots & * \\
\vdots & \vdots & \vdots & \cdots & \vdots & 0 & * & \cdots & \cdots & * \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & 0 & * & \cdots & * \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0
\end{pmatrix}$$

$1.Zeile = a_1$
$2.Zeile = a_2$
$\vdots$
letzte Zeile vor Nullzeilen $= a_r$
$r = rang(A),\ \pi_j$ =Pivotelemte, $j_1, \dots, j_r$ = Sprungstellen der Treppenfunktion
1. Behauptung: Die ersten r Zeilen $a_1, \dots, a_r$ sind linear unabhängige Vektoren in $\mathbb{K}^n$.
2. Behauptung: Die Pivotspalten $j_1, \dots, j_r$ sind linear unabhängig in $\mathbb{K}^m$.

Anwenden: $a_1, \dots, a_m \in V = \mathbb{K}^n,\ U = Span(a_1, \dots, a_n)$
$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \Rightarrow$ Zeilenstufenform mit Treppenfunktion
Gesucht ist eine Basis.

Nachtrag:

Lemma 3.3.6. *$\mathcal{E} \subseteq V$. $Span(\mathcal{E})$ ist der kleinste ("kleinste obere Schranke") Untervektorraum von V, der $\mathcal{E}$ enthält.*

Beweis. Alle UVR von V bilden einen "Verband" = Menge X mit einer *Ordnungsrelation*.
D.h. wir haben:
zwischen je zwei $x_1, x_2 \in X$ kann eine Relation gelten $x_1 \leq x_2$ oder nicht. Es muss nicht Vergleichbarkeit herrschen. Es kann sein, dass weder $x_1 \leq x_2$ noch $x_2 \leq x_1$ gilt. □

Beispiel 3.3.7. $X = \mathbb{R}, x_1 \leq x_2$ mit der uns bekannten Bedeutung.
Hier:
$x_1 \leq x_2$ oder $x_2 \leq x_1$ oder beides $(\Rightarrow x_1 = x_2)$
Dies nennt man eine *totale* oder *lineare Ordnung*.

Beispiel 3.3.8. M = Menge, $\mathfrak{X} = Pot(M)$ = Potenzmenge = $\mathcal{P}$ = Menge aller Teilmengen von M.
$M = \{1, 2\}$
$\mathfrak{X} = \emptyset$
$\swarrow \searrow$

$$\begin{tikzcd} \{1\} \arrow[dr] & & \{2\} \arrow[dl] \\ & \{1,2\} & \end{tikzcd}$$

$Ordnungsrelation: \; x_1 \leq x_2 \; :\Leftrightarrow \; x_1 \subseteq x_2$

Beispiel 3.3.9. $\mathfrak{X} = \mathbb{N}, a \leq b \Leftrightarrow a$ ist Teiler von b.
Es müssen folgende Axiome gelten:

1. Reflexivität: $x \leq y$
2. Transitivität: $x \leq y \wedge y \leq y \Rightarrow x \leq z$

Ein maximales Element in $\mathfrak{X}$ ist ein $x \in \mathfrak{X}$, sodass gilt: $x \geq x_0 \Rightarrow x = x_0$

Beispiel 3.3.10. $V = VR$ über $\mathbb{K}, \mathfrak{X} =$ Menge der linear unabhängigen Teilmengen von V, Ordnungsrelation = Inklusion

Lemma 3.3.11. Zorn'sches Lemma
garantiert ein maximales $\mathcal{B} \in \mathfrak{X}$. Es ist dann leicht zu zeigen: $\mathcal{B}$ ist auch ein EZS.

Beweis. Angenommen, das ist nicht so. D.h. es gibt ein UVR W mit $\mathcal{E} \subseteq W \subsetneq Span(\mathcal{E})$.
Dann gibt es ein $x \in Span(\mathcal{E})$ mit $x \notin W$, also gibt es geeignete $e_1, \dots, e_n \in \mathcal{E}$ und $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ mit $x = \lambda_1 e_1 + \dots + \lambda_n e_n$.
Es ist aber:
alle $e_i \in W$
alle $\lambda e_i \in W$
ABER $x = \lambda_1 e_1 + \dots + \lambda_n e_n \in W \Rightarrow$ Widerspruch! □

Beispiel 3.3.12. $Span(\mathcal{E})$ ist der Durchschnitt aller UVR, die $\mathcal{E}$ enthalten.

Beweis. Sei $U_\alpha(\alpha \in I)$: alle UVR in V mit $\mathcal{E} \subseteq U_\alpha$. Nach (1) und (2) ist $Span(\mathcal{E})$ ein solches U_α. Also gilt: $\bigcup\limits_{\infty} U_\alpha \subseteq Span(\mathcal{E})$.
Sei $x = \lambda_1 e_1 + \dots + \lambda_n e_n \in Span(\mathcal{E})$ d.h. $e_i \in \bigcap\limits_{\infty} U_\alpha$. □

Nachtrag:

Lemma 3.3.13. *Ist $\mathcal{B}$ eine Basis, $b_i \in \mathcal{B}, 0 \neq \lambda \in \mathbb{K}$, dann ist $\mathcal{B}' = (\mathcal{B} - \{b_i\} \cup \{\lambda \cdot b_i\})$ wieder eine Basis.*

Beweis. Aus Lemma 3.2 und Lemma 3.17. □

Lemma 3.3.14. *Ist $\mathcal{B}$ eine Basis, $b_i, b_j \in \mathcal{B}$, dann ist $\mathcal{B}' = (\mathcal{B} - \{b_i\} \cup \{b_i + b_j\}$ wieder eine Basis.*

Beweis. Aus Lemma 3.3 und Lemma 3.18. □

Lemma 3.3.15. *Austauschsatz von Steinitz*
Es sei $\mathcal{B} = \{b_1, \dots, b_i, \dots\}$ eine Basis von V und sei $b \in V$ mit einer Darstellung
$b = \beta_1 b_1 + \dots + \beta_i b_i + \dots + \beta_n b_n$ für $\beta_1, \dots, \beta_n \in \mathbb{K}$ und etwa $\beta_i \neq 0$.
Dann ist $\mathcal{B}' = (\mathcal{B} - \{b_i\} \cup \{b\}$ wieder eine Basis.

Beweis. Sei $v \in V$ dargestellt durch $v = \lambda_1 b_1 + \dots + \lambda_i b_i + \dots + \lambda_n b_m$.
O.B.d.A. nehmen an: $n = m$ (Ergänzung durch Nullterme, falls $n \neq m$)
Dann können wir schreiben:
$v = (\lambda_1 - \frac{\beta_1 \lambda_i}{\beta_i}) b_1 + \dots + \hat{b_i} + \dots + (\lambda_n - \frac{\beta_n \lambda_i}{\beta_i}) b_n$
weil $b_i = -\frac{\beta_1}{\beta_i} b_1 - \dots - \frac{\beta_{i-1}}{\beta_i} bi - 1 - \frac{\beta_{i+1}}{\beta_i} bi + 1 - \dots - \frac{\beta_n}{\beta_i} b_n$
D.h. v ist auch darstellbar durch $\mathcal{B}'$. Ist $\alpha_1 b_1 + \dots + \alpha_i b_i + \dots + \alpha_n b_n = 0$ eine Relation in $\mathcal{B}'$, so erhalten wir durch Einsetzten von b daraus eine Realtion in $\mathcal{B}$:
$(\alpha_1 - \alpha_i \beta_1) b_1 + \dots + \alpha_i \beta_i b_i + \dots + (\alpha_n - \alpha_i \beta_n) b_n = 0$
Weil $\mathcal{B}$ linear unabhängig, muss dies eine triviale Relation sein, also
$\alpha_1 - \alpha_i \beta_1 = 0$

$$\begin{array}{ll} \vdots & \\ \alpha_i\beta_i = 0 & \Rightarrow \alpha_i = 0 \rightarrow \text{Also } \alpha_k = 0 \text{ auch für } \alpha_1, \dots, \alpha_n \\ \vdots & \\ \alpha_n - \alpha_i\beta_n = 0 & \end{array}$$

Deshalb ist $\mathcal{B}'$ linear unabhängig. □

Satz 3.3.16. *Existenzsatz von Basen*
Jeder Vektorraum besitzt eine Basis

Bemerkung 3.3.17. Wir zeigen dies hier nur für endlich abzählbar erzeugte VR über $\mathbb{K}$, denn dafür genügt die gewöhnliche Induktion. Für VR. die überabzählbar erzeugt sind, braucht man "transfinite Induktion". Eine Variante ist das *Zorn'sche Lemma*, welches zum Auswahlaxiom und zum Wohlordnungssatz äquivalent ist. (siehe Aufgabe 35)

Beweis. (Im endlichen oder abzählbaren Fall)
Wir wählen ein abzählbares EZS $\mathcal{E} = \{e_1, e_2, \dots, e_k, \dots\}$. (Ist V endlich erzeugt, so läuft k nur bis zu diesem einen endlichen n.) Wir betrachten technisch (algorithmisch) $\mathcal{E}$ als eine geordnete Folge.
Wir setzten $\mathcal{E}_\prime := \emptyset$
$\mathcal{E}_k := \{e_1, \dots, e_k\}$ für $k = 1, \dots, (n), \dots$
Algorithmus (Basisauswahlverfahren)
$\mathcal{B} := \emptyset$

$$\mathcal{B}_k := \begin{cases} \mathcal{B}_{k-1} & \text{falls } b_k \in Span(\mathcal{B}_{k-1}) \\ \mathcal{B}_{k-1} \cup b_k & \text{falls } b_k \notin Span(\mathcal{B}_{k-1}) \end{cases}$$

(Falls V endlich erzeugt wird, so läuft k nur bis n.)

1. $\mathcal{B}_k \subseteq \mathcal{B}_{k-1}$ - nach Konstruktion
2. $Span(\mathcal{B}_k) = Span(\mathcal{E}_k)$ - nach Konstruktion und Induktion
3. $\mathcal{B}_k$ ist linear unabhängig - nach Induktion, mit Lemma 3.19.
   Wir setzten $\mathcal{B} = \begin{cases} \bigcup\limits_{k=0}^{n} & \text{falls } V \text{ endlich erzeugt ist} \\ \bigcup\limits_{k=0}^{\infty} & \text{falls } V \text{ abzählbar unendlich ist} \end{cases}$
4. $\mathcal{B}$ ist ein EZS:
   Ist $v \in V$ durch die ersten $k : e_1, \dots, e_k$ darstellbar, also $v = \lambda_1 b_1 + \dots + \lambda_k b_k$, so ist
   $v \in Span(\mathcal{E}_k) =_2 Span(\mathcal{B}_k) \subseteq_1 Span(\mathcal{B})$
5. $\mathcal{B}$ ist linear unabhngig.
   Gibt es eine Relation zwischen (endlich vielen) Elementen in $\mathcal{B}$, so liegen diese bereits in einem $\mathcal{B}_N$, weil es nur endlich viele sind.
   $$\begin{array}{ll} e_1 \in \mathcal{B}_{k_1} & \\ e_2 \in \mathcal{B}_{k_2} & \\ \vdots & \Rightarrow e_1, \dots, e_l \in \mathcal{B}_N,\ N = max\{k_1, \dots, k_n\} \\ e_l \in \mathcal{B}_{k_l} & \end{array}$$
   Aber nach 3) ist $\mathcal{B}_N$ linear unabhängig.

□

Bemerkung:
Das Basisauswahlverfahren (BAV) aus dem Basisexistenzsatz ist im Wesentlichen ein Algorithmus:

<u>EINGABE:</u> Geordnetes EZS $\mathcal{E} = \{e_1, ..., e_n, ...\}$ (endlich oder abzählbar)

<u>ABFRAGE:</u> Ist $e_{k+1} \in Span(\{e_1, ..., e_k\})$?
Wenn nein füge Ihn zur Basis hinzu.

Wenn ja füge es nicht zur Basis hinzu und betrachte e_{k+2}

AUSGABE: Geordnete Basis $\mathcal{B} \subseteq \mathcal{E}$

Beispiele:
1) $V = Pol(\mathbb{R})$ (Die Polynomfunktionen $f : \mathbb{R} \longrightarrow \mathbb{R}$).
Wähle als EZS die Stammpolynome $\mathcal{E} = \{1, X, X^2, ..., X^n, ...\}$.
Behauptung: $X^{n+1} \notin Span(1, ..., X^n)$ (Dies ist für Polynome trivial aber noch lange nicht für Polynomfunktionen!)
Um dies zu zeigen nutzen wir einen Trick der universal einsetzbar ist:
Gäbe es eine Linearkombination, so dass

$$X^{n+1} = \sum_{k=0}^{n} a_k X^k \tag{3.1}$$

so könnte man (1) n-mal differenzieren (bei Polynomfunktionen wendet man formales Differenzieren an) und erhält

$$(n+1)!X = a_n n! \tag{3.2}$$

Damit wäre aber bereits, wenn X=0 in (2) eingesetz wird

$$0 = a_n n! \Rightarrow a_n = 0 \tag{3.3}$$

setzt man nun (3) in (1) ein und iteriert das Verfahren, erhält man, dass alle $a_i = 0$ sind, womit aber $X^{n+1} = 0$ wäre, was definitiv nicht der Fall ist.

2) Analog sieht man ein, dass $\mathcal{B} = \{e^{a_k x} | k \in \mathbb{N} \ und \ a_k \in \mathbb{R}\}$ linear unabhängig in $C^\infty(\mathbb{R})$ ist.

3) Wiederum Analog sieht man ein, dass $\mathcal{B} = \{sin(a_k x) | k \in \mathbb{N} \ und \ a_k \in \mathbb{R}\}$ linear unabhängig ist.

4) Sei V der Raum stetiger stückweise affiner Funktionen $f : [0; 1] \longrightarrow \mathbb{R}$ mit 0=f(0)=f(1)
So ist $\mathcal{B} = \{Z_\xi \mid Z_\xi(x) : [0;1] \longrightarrow \mathbb{R};\ Z_\xi = \begin{cases} \dfrac{x}{\xi} & 0 \leq x \leq \xi \\ \dfrac{x-1}{\xi-1} & \xi \leq x \leq 1 \end{cases} ; \xi \in \mathbb{R}\}$ eine Basis.
Es ist sehr selten, dass eine überabzählbare Basis explizit angegeben werden kann, da es im Allgemeinen sehr schwer ist mit überabzählbarer Dimension (siehe Abschnitt 3.4) zu arbeiten. So ist für $C^\infty(\mathbb{R})$ keine Hamel-Basis bekannt.

Satz 3.3.18 (Basisergänzungssatz)**.** *Sei V ein endlich oder abzählbar erzeugter K-Vektorraum. So kann jede linear unabhängige Teilmenge $\mathcal{B}' \subseteq V$ zu einer Basis ergänzt werden.*

Beweis. Wähle ein beliebiges geordnetes Erzeugendensystem $\mathcal{E} = \{e_1, ..., e_n, ...\}$, welches nach dem Basisexistenzsatz definitiv existiert und wegen Abzählbarkeit auch geordnet werden kann.
So ist offensichtlich $\mathcal{E}' = \{b_1, ..., b_n, e_1, ..., e_n, ...\}$ erneut ein Erzeugendensystem. Wird nun BAV auf $\mathcal{E}'$ angewandt, so werden die b_i zuerst betrachtet und wegen linearer Unabhängigkeit als Basisvektoren gewählt. Somit gilt für die letztendliche Basis $\mathcal{B}' \subseteq \mathcal{B}$. □

Beispiele:(BAV)
1)$V = K^n;\ \mathcal{E} = \{e_1, e_1 + e_2, e_1 + e_2 + e_3, ..., e_1 + e_2 + ... + e_n\}$
$\mathcal{B}' = \{2e_2\} \longmapsto \mathcal{B} = \{2e_2, e_1, e_1 + e_2 + e_3, e_1 + e_2 + e_3 + e_4, ...\}$

2)$V = \mathbb{Q}(\sqrt{2}) \subset \mathbb{R};\ K = \mathbb{Q};\ \mathcal{E} = \{2, 2\sqrt{2}\}$
$\mathcal{B}'_1 = \{1 + \sqrt{2}\} \longmapsto \mathcal{B}_1 = \{1 + \sqrt{2}, 2\}$
$\mathcal{B}'_2 = \{3\} \longmapsto \mathcal{B}_2 = \{3, 2\sqrt{2}\}$

3.4 Dimension

Unser Ziel wird es sein die Dimension eines Vektorraumes durch die Kardinalität einer Basis des Vektorraumes zu definieren.
Doch dafür müssen wir uns zunächst überzeugen, dass solch ein Dimensionsbegriff wohldefiniert ist. Nach dem Basisexistenzsatz besitzt jeder Vektorraum zumindest eine Basis, dies bereitet uns also keine Probleme. Allerdings haben wir in den Übungen eingesehen, dass ein Vektorraum sehr viele Basen besitzt.
Zwar hat $\mathbb{F}_2$ genau eine Basis, dieser ist aber auch der einzige neben dem trivialen Nullvektorraum mit nur einer Basis. So besitzt $\mathbb{F}_3$ zwei Basen, $(\mathbb{F}_2)^2$ hat drei Basen und $(\mathbb{F}_7)^3$ besitzt bereits 5.630.688 Basen.
Noch deutlicher wird das Problem im $\mathbb{R}^d$. Denn nach dem Basisergänzungssatz kann man eine Basis konstruieren, indem man zunächst einen beliebigen von Null verschiedenen Vektor wählt, und daraufhin immer wieder von den vorher gewählten linear unabhängige Vektoren hinzufügt. So ergibt sich für die Menge aller geordneten Basen des $\mathbb{R}^d$ folgende Darstellung $\prod_{k=1}^{d-1}\{\mathbb{R}^d \setminus \mathbb{R}^k\}$ Womit $\mathbb{R}^d$ sogar überabzählbar viele Basen besitzt.
Also stellt sich die Frage, ob alle Basen gleich mächtig sind.

Proposition 3.4.1. *Sei V ein von n Vektoren (insbesondere endlich) erzeugter Vektorraum, dann sind n+1 Vektoren linear abhängig.*

Beweis. Es sei $\mathcal{E} = \{e_1, e_2, ..., e_n\}$ ein Erzeugendensystem in V. So folgt die Behauptung induktiv:

Induktionsverankerung: Für n=0 ist $\mathcal{E} = \{\}$ womit $V = Span(\{\}) = 0$ ist.
Induktionsschritt: Seien $v_1, ..., v_n, v_{n+1} \in V$ beliebige aber verschiedene Vektoren. So besitzt jedes v_i folgende Darstellung:

$v_i = \sum_{j=1}^{n} \omega_{i,j} e_{i,j}$ mit $\omega_{i,j} \in K$

Zu Zeigen ist, dass eine nicht-triviale Relation $0 = \sum_{k=1}^{n+1} \lambda_k v_k$ existiert, also dass mindestens ein $\lambda_k \neq 0$ ist. Betrachte λ_k als Unbekannte in einem wie folgt konstruierten LGS:

$Es\ gelte: 0 = \sum_{i=1}^{n+1} \lambda_i v_i = \sum_{i=1}^{n+1} \sum_{j=1}^{n} \omega_{i,j} e_{i,j}$
setze $\omega'_{i,j} := \omega_{j,i}$

Womit sich ein LGS des Ranges $r \leq n$ ergibt, also min. eine freie Variable λ_k enthält. Womit wiederum das LGS eine nichttriviale Lösung besitzt. □

Nun können wir folgenden gewünschten Satz formulieren:

Satz 3.4.2 (Dimensionssatz)**.** *Sei V ein endlich erzeugter Vektorraum und $\mathcal{B} = \{b_1, ..., b_n\}$ sowie $\mathcal{B}' = \{b'_1, ..., b'_m\}$ zwei Basen in V, dann sind $\mathcal{B}$ und $\mathcal{B}'$ schon gleich mächtig (bzw. n=m).*

Beweis. $\mathcal{B}$ ist ein Erzeugendensystem in V mit n Vektoren. Nach Proposition 1 sind n+1 Vektoren linear Abhängig, womit $m \leq n$. Analog für $\mathcal{B}'$.
Somit ist n=m. □

Also können wir nun unseren angestrebten Dimensionsbegriff definieren:

Definition 3.4.3 (Dimension)**.** Sei $\mathcal{B}$ eine Basis in V, wobei V ein K-Vektorraum ist. So definiere:

$$dim_K(V) := |\mathcal{B}|$$

Als die Dimension von V.

Beispiele:
1) Sei K ein Körper so ist $dim_K(K^n) = n$

2) $dim_{\mathbb{R}}(\mathbb{C}) = 2$

3) Nach Übung sind $log(p_i)$ für beliebige Primzahlen p_i linear unabhängig in $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum. Also

ist $dim_{\mathbb{Q}}(\mathbb{R}) = \infty$.
Dies ist zwar etwas informell, da es wie bereits erwähnt mehrere Arten der Unendlichkeit gibt. Es wird sich herausstellen, dass jene Dimension sogar überabzählbar ist.
Hier sieht man auch, dass unsere geometrische Vorstellung des Dimensionsbegriffes bei manchen Vektorräumen versagt, da R als Q-Vektorraum so zu sagen unendlich viele Richtungen besitzt, so stellt z.B. jedes $log(p_i)$ und jedes $\sqrt{a}$ für ein $a \neq n^2$ mit $a, n \in \mathbb{N}$ eine Richtung dar.

4) Sei K[X] ein Polynomring über K, so ist $dim_K(K[X]) = |\mathbb{N}|$ also abzählbar.
Denn wähle als Basis $\mathcal{B} = \{1, X, X^2, ..., X^n, ...\}$.

5) Sei $\mathbb{A}$ der Raum stetiger stückweise affiner Funktionen auf dem Intervall [0;1] mit f(0)=0 und f(1)=0. So ist $dim_{\mathbb{R}}(\mathbb{A}) = |R|$ also überabzählbar.

6) Sei $\mathbb{E}$ der Untervektorraum des Vektorraumes der glatten Funktionen über $\mathbb{R}$ aufgespannt von e^{ax}. So ist $dim_{\mathbb{R}}(\mathbb{E}) = |\mathbb{R}|$ also auch überabzählbar.

Bemerkung:
Was wir in der linearen Algebra eine Basis nennen heißt im allgemeinen Hamel-Basis. Denn in anderen Gebieten der Mathematik treten andere Basisbegriffe auf. So arbeitet die Analysis oftmals in bzw. mit Räumen überabzählbarer Dimension, wie $C^\infty(\mathbb{R})$, wobei es für solche im allgemeinen sehr schwer ist eine Hamel-Basis anzugeben. Also betrachten Analytiker lediglich approximative Basen, die sogenannte Hilbert-Basis, bei der sich zwar nicht jedes Element des Vektorraumes als (endliche) Linearkombination der Basisvektoren schreiben lässt, wohl aber beliebig gut durch eine Linearkombination der Basisvektoren approximiert werden kann. Beispiele sind Fourier-Reihen, welche glatte periodische Funktionen durch Überlagerungen von sin und cos approximieren. Und der Weierstraß'sche Approximationssatz garantiert, dass glatte Funktionen durch Polynomfunktionen approximiert werden können.

Für die Algebra reicht eine Hilbertbasis jedoch nicht aus!

Die folgende Propsition mag trivial erscheinen –und bei Vektorräumen ist sie es tatsächlich–, aber wenn man sich beispielsweise mit Gruppen beschäftigt, gilt die analoge Aussage nicht.

Proposition 3.4.4. *Für einen Unterraum $U \subseteq V$ eines endlich-erzeugten Vektorraums gilt immer* $\dim(U) \leq \dim(V)$.

Beweis. Sei $\mathcal{B}$ eine Basis von U, dann ergänze sie zu einer Basis $\mathcal{B}'$ von V, also $\mathcal{B} \subseteq \mathcal{B}'$. Dann ist $\dim(U) = |\mathcal{B}| \leq |\mathcal{B}'| = \dim(V)$ □

Proposition 3.4.5. *Es seien U', U'' Untervektorräume von V (V wieder endlich-erzeugt) mit $U := U' \cap U''$ und $\tilde{U} := U' + U''$. Dann gilt* $\underbrace{\dim(\tilde{U})}_{n+m+k} = \underbrace{\dim(U')}_{k+n} + \underbrace{\dim}_{k+m}(U'') - \underbrace{\dim(U)}_{k}$.

Beweis. Wir wählen eine Basis $\mathcal{B}$ für U. Wir ergänzen einmal zu einer Basis $\mathcal{B}' \supset \mathcal{B}$ von U' und $\mathcal{B}'' \supset \mathcal{B}$ von U''. Außerdem setzen wir $\tilde{\mathcal{B}} = \mathcal{B}' \cup \mathcal{B}''$ Die Beziehungen sind im folgenden *kommutativen Diagramm* dargestellt:

$$\begin{array}{cccccccc} & U'' & \hookrightarrow & \tilde{U} & = U' + U'' & & \mathcal{B}'' & \subseteq & \tilde{\mathcal{B}} \\ & \uparrow & & \uparrow & & & \cup\mathrm{I} & & \cup\mathrm{I} \\ U' \cap U'' = & U & \hookrightarrow & U' & & \mathcal{B}' \cap \mathcal{B}'' = & \mathcal{B} & \subseteq & \mathcal{B}' \end{array}$$

Behauptung: $\tilde{\mathcal{B}}$ ist eine Basis für U'. $\mathcal{B}'$ ist offensichtlich ein Erzeugendensystem für $\tilde{U}$. Betrachte die linearunabhängige Menge $\mathcal{B}'$ und das Erzeugendensystem $\mathcal{E} := \tilde{\mathcal{B}} = \mathcal{B}' \cup \{b_1', ..., b_{m-k}'\}$ von $\tilde{U}$ und wende das Basisauswahlverfahren an:
$b_1'' \notin \mathrm{Span}(\mathcal{B}')$, also Fügen wir b_1' hinzu; $b_2'' \notin \mathrm{Span}(\mathcal{B}' \cup \{b_1''\})$, wir fügen b_2' hinzu, und so weiter. Das heißt, das Basisauswahlverfahren liefert uns hier tatsächlich $\tilde{\mathcal{B}}$. □

Beispiel 3.4.6. 1. Seien $\mathbb{K} \subseteq \mathbb{L}$ Körper. Sei $\Lambda = \{\lambda_1, ..., \lambda_l\}$ eine Basis von $\mathbb{L}$ als $\mathbb{K}$-Vektorraum. Ist V ein $\mathbb{L}$-Vektorraum mit Basis $\mathcal{B}_{\mathbb{L}} = \{b_1, ..., b_n\}$ (als $\mathbb{L}$-Vektorraum!), dann ist

$$\mathcal{B}_{\mathbb{K}} := \Lambda \cdot \mathcal{B}_{\mathbb{K}} = \{\lambda_1 b_1, ..., \lambda_l b_1, \lambda_1 b_2, ..., \lambda_l b_2, ..., \lambda_1 b_n, ..., \lambda_l b_n\}$$

eine Basis von V als $\mathbb{K}$-Vektorraum. Insbesondere gilt:

$$\dim_{\mathbb{K}} V = \underbrace{\dim_{\mathbb{K}} \mathbb{L}}_{l} \cdot \underbrace{\dim_{\mathbb{L}} V}_{n}$$

Konkret heißt das, wenn wir $\mathbb{K} = \mathbb{R}$ und $\mathbb{L} = \mathbb{C}$ betrachten: $\Lambda = \{1, i\}, l = 2$. Also: $\dim_{\mathbb{R}} V = 2 \cdot \dim_{\mathbb{C}} V$

2. Wir betrachten $V = \mathbb{C}^n$ als $\mathbb{C}$- und als $\mathbb{R}$-Vektorraum:

$$\mathcal{B}_{\mathbb{C}} = \{e_1, ..., e_n\};\ e_k = (0, ..., 0, \underset{\boxed{k}}{1}, 0, ..., 0)$$

$$\mathcal{B}_{\mathbb{R}} = \{e_1, ie_1, e_2, ie_2, ..., e_n, ie_n\};\ ie_k = (0, ..., 0, \underset{\boxed{k}}{i}, 0, ..., 0)$$

3. Wir betrachten $V = \mathbb{C}[X]$ der komplexen Polynome in X, wieder zunächst als Vektorraum über $\mathbb{C}$ und dann über $\mathbb{R}$:

$$\mathcal{B}_{\mathbb{C}} = \{1, X, X^2, ..., X^n, ...\} \text{ Stammpolynome}$$

$$\mathcal{B}_R = \{1, i, X, iX, ..., X^n, iX^n, ...\}$$

Beispiel 3.4.7. es sei $\mathbb{L}$ ein *endlicher* Körper (etwa $\mathbb{F}_2, \mathbb{F}_3, \mathbb{F}_4, \mathbb{F}_5, \ldots, \mathbb{F}_p$). Weil $\mathbb{L}$ endlich ist, muss $\mathrm{char}(\mathbb{L}) = p$ (p eine Primzahl) sein. Wie in Definition 2.1.11 setzen wir

$$\varepsilon_0 := 0, \varepsilon_1 := 1, \varepsilon_{n+1} := \varepsilon_n + 1$$

Also gilt $\varepsilon_p = 0$.
Behauptung: $\mathbb{K} := \{\varepsilon_0$, für $\varepsilon_1, ..., \varepsilon_{p-1}\}$ bilden einen Unterkörper, der isomorph ist zu $\mathbb{F}_p = \mathbb{Z}/_{p\mathbb{Z}}$. Zu zeigen: $\mathbb{F}_p \hookrightarrow \mathbb{L}, \underline{r} \mapsto \varepsilon_r,\ 0 \leq r < p$. Es gilt $-\varepsilon_k = \varepsilon_{p-k}, \varepsilon_n + \varepsilon_m = \varepsilon_{n+m}$. Das multiplikative Inverse ε_k^{-1} ist durch den euklidischen Algorithmus eindeutig bestimmt.
$\mathbb{L}$ hat als Vektorraum über $\mathbb{K}$ eine endliche Basis $\mathcal{B} = \{b_1, ..., b_n\}$. Da jedes $v \in \mathbb{L}$ genau eine Darstellung $v = \lambda_1 b_1 + ... + \lambda_n b_n$ hat, sind insgesamt p^n Elemente möglich: Für jeden der n Koeffizienten $\lambda_1 ... \lambda_n$ gibt es jeweils p Wahlmöglichkeiten, da $\lambda_i \in \mathbb{K}$. Daraus folgt, dass

$$|\mathbb{L}| = p^n$$

Es gibt also insbesondere keinen Körper mit etwa sechs Elementen.

3.5 Prinzip der linearen Fortsetzung

Das Prinzip der linearen Fortsetzung besagt, dass eine lineare Abbildung durch die Werte auf einer Basis vollständig bestimmt ist. Dieser „Trick“ist hilfreich, da man statt unendlicher Systeme endliche Basen betrachten kann. Außerdem können wir die Werte auf einer Basis vorgeben, und dann wissen, dass es eine solche lineare Abbildung gibt.

Proposition 3.5.1. *Seien V, W zwei Vektorräume über einem Körper $\mathbb{K}$.*

(i) Für ein Erzeugendensystem $\mathcal{E}$ von V gilt: Sind $f, g : V \to W$ zwei lineare Abbildungen und gilt $f(e) = g(e)$ für alle $e \in \mathcal{E}$, dann gilt $f = g$ (d.h. $\forall v \in V : f(v) = g(v)$).

(ii) Für eine linear-unabhängige Menge $\mathcal{B}$ in V gilt: Ist $\varphi : \mathcal{B} \to W, b \mapsto \varphi(b)$ eine beliebige Funktion, so gibt es eine lineare Fortsetzung *$f : V \to W$, das heißt eine lineare Abbildung f mit der Eigenschaft $f|_{\mathcal{B}} = \varphi$ (d.h. $f(b) = \varphi(b), b \in \mathcal{B}$).*
Im kommutativen Diagramm:

$$\begin{tikzcd} V \arrow[r, dashed, "f"] & W \\ \mathcal{B} \arrow[u, phantom, sloped, "\subseteq"] \arrow[ur, "\varphi"] & \end{tikzcd}$$

(iii) *Aus (i) und (ii) folgt dann: Für eine Basis $\mathcal{B}$ von V gilt: Jede Funktion $\varphi : \mathcal{B} \to W$ besitzt geanu eine lineare Fortsetzung $f : V \to W$ (d.h. $f|_{\mathcal{B}} = \varphi$).*

Korollar 3.5.2. *Es gibt eine Bijektion zwischen den Funktionen von $\mathcal{B}$ nach V und den linearen Abbildungen von V nach W.*

$$\begin{aligned} \mathrm{Funkt}(\mathcal{B}, W) = W^{\mathcal{B}} &\xrightarrow{\cong} \mathrm{lin}_{\mathbb{K}}(V, W) = \mathrm{Hom}_K(V, W) \\ \varphi &\longmapsto f \\ \varphi := f|_{\mathcal{B}} &\longleftarrow\!\shortmid f \end{aligned}$$

Beweis. (i) Sei v dargestellt durch $\mathcal{E}$, etwa

$$v = \lambda_1 e_1 + ... + \lambda_n e_n$$

Dann rechnet man einfach nach:

$$\begin{aligned} f(v) &= f(\lambda_1 e_1 + ... + \lambda_n e_n)) \\ &= \lambda_1 \underbrace{f(e_1)}_{\shortparallel} + ... + \lambda_n \underbrace{f(e_n)}_{\shortparallel} \\ &= \lambda_1 \overbrace{g(e_1)} + ... + \lambda_n \overbrace{g(e_n)} \\ &= g(\lambda_1 e_1 + ... + \lambda_n e_n) \\ &= g(v) \end{aligned}$$

(ii)Auf $\mathcal{B} = \{b_1, ..., b_n\}$ seien Werte $w_i = \varphi(b_i)$ vorgegeben. wir ergänzen $\mathcal{B}$ zu einer Basis $\mathcal{B}' \supseteq \mathcal{B}$ von v:

$$\mathcal{B}' = \{b_1, ..., b_n, \underbrace{b'_1, ..., b'_m}_{\text{ergänzt}}\}$$

Wir schreiben $v = \lambda_1 b_1 + ... + \lambda_n b_n + \lambda'_1 b'_1 + ... + \lambda'_m b'_m$. Wir wählen $w_{n+1}, ..., w_{n+m}$ beliebig in W (z.B. $= 0$). Dann definieren wir:

$$f(v) = \lambda_1 w_1 + ... + \lambda_n w_n + \lambda'_1 w_{n+1} + ... + \lambda'_m w_{n+m}$$

Wir stellen fest:

1. Damit ist $f : V \to W$ für jedes v definiert, denn für jedes v gibt es eine Darstellung.
2. f ist wohldefiniert, weil die Darstellung von v eindeutig ist.
3. f ist linear.
4. $f|_{\mathcal{B}} = \varphi$ auf $\mathcal{B}$

□

V Sei ein $\mathbb{K}$-Vektorraum. $\mathcal{B} \subseteq V$

$$\begin{tikzcd} V \arrow[r, dashed, "f"] & W \\ \mathcal{B} \arrow[u, dashed, "\subseteq"] \arrow[ur, "\phi"] & \end{tikzcd}$$

f ist linear, ϕ entspricht dann einer beliebigen Funktion.
$\mathrm{Funkt}(\mathcal{B}, W) = W^{\mathcal{B}} \longrightarrow \mathrm{Lin}_{\mathcal{B}}(V, W) = \mathrm{Hom}_{\mathcal{B}}(V, W)$; $\phi \longleftrightarrow f$

1. $\mathcal{E}$ ist Erzeugendensystem
   $f, g : V \longrightarrow W$
   $f_{|\mathcal{E}} = g_{|\mathcal{E}} \implies f = g$
2. $\mathcal{B}$ ist linear-unabhängig $\forall \phi : \mathcal{B} \longrightarrow W \exists f : V \longrightarrow W$ linear $f_{|\mathcal{B}} = \phi$
3. $\mathcal{B}$ ist eine Basis $\forall \phi : \mathcal{B} \longrightarrow W \exists! f : V \longrightarrow W$ linear $f_{|\mathcal{B}} = \phi$

Beispiel 3.5.3. $V = \mathbb{C}$, $\mathbb{K} = \mathbb{C}$

3.6 Koordinaten

V ist endlich erzeugter Vektorraum über $\mathbb{K}$.
$\mathcal{B} = (b_1, ..., b_n)$ ist geordnete Basis.
Sei $V \ni v = \lambda_1 b_1 + ... + \lambda_n b_n$ und $v = \lambda_1' b_1 + ... + \lambda_n' b_n$, dann ist $\lambda_1 = \lambda_1', ..., \lambda_n = \lambda_n'$ (eindeutige Darstellung).

Definition 3.6.1. Falls $\mathcal{B}$ eine fest gewählte Basis ist, dann gibt es *Koordinaten* bzgl. $\mathcal{B}$. Definiert durch

$$K_{\mathcal{B}} : V \longrightarrow \mathbb{K}^n$$

$$v \longmapsto (\lambda_1, ..., \lambda_n),$$

falls $v = \lambda_1 b_1 + ... + \lambda_n b_n$.
$K^i_{\mathcal{B}}(v) = \lambda_i$ ist die i-te Koordinate.
$K_{\mathcal{B}}(v) = (K^1_{\mathcal{B}}(v), ..., K^n_{\mathcal{B}}(v))$

Satz 3.6.2. *$K_{\mathcal{B}}$ ist ein Isomorphismus.*

Beweis.

1. zz. $K_{\mathcal{B}}$ ist bijektiv:
   $(\lambda_1, ..., \lambda_n) \longmapsto v = \lambda_1 b_1 + ... + \lambda_n b_n$ Rückrichtung ist klar nach Konstruktion (siehe Definition).

2. zz. $K_{\mathcal{B}}$ ist linear:
   $K_{\mathcal{B}}(0) = (0, ..., 0) = 0$
   Seien $v, v' \in V$.
   Dann gilt $K_{\mathcal{B}}(v) + K_{\mathcal{B}}(v') = K_{\mathcal{B}}(v + v')$, denn$(\lambda_1 b_1 + ... + \lambda_n b_n) + (\lambda_1' b_1 + ... + \lambda_n' b_n) = (\lambda_1 + \lambda_1')b_1 + ... + (\lambda_n + \lambda_n')b_n$
   Sei $\alpha \in \mathbb{K}$ und $v \in V$. $K_{\mathcal{B}}(\alpha v) = \alpha K_{\mathcal{B}}(v)$, denn $\lambda_1 \alpha b_1 + ... + \lambda_n \alpha b_n = \alpha(\lambda_1 b_1 + ... + \lambda_n b_n)$

□

Beispiel 3.6.3. 1. Wenn $V = \mathbb{K}^n$ und $\mathcal{B} = (b_1, ..., b_n)$ Standardbasis, dann ist $K_{\mathcal{B}}(v) = v$, also $K_{\mathcal{B}} = id$. Denn $v = (v_1, ..., v_n) = v_1 e_1 + ... + v_n e_n$.

2. Wenn $V = \mathbb{R}^2$ und $\mathcal{B} = (b_1, b_2) = ((1, 0), (1, 1))$, dann ist $v = (1, 1) = 1b_1 + 1b_2 = 1(1, 0) + 1(1, 1)$

3. Wenn $V = P_n(\mathbb{K}) \widehat{=}$ Vektorraum der Polynome vom Grad $\leq n$ und $\mathcal{B} = \{1, X, X^2, ..., X^n\}$. Dann ist $p(X) = a_0 + a_1 X + ... + a_n X_n$, $K_{\mathcal{B}} = (a_0, ..., a_n)$

Es stellt sich nun die Frage, was tun, wenn V nicht endlich erzeugt ist?

- wähle auch hier Basis $\mathcal{B}$
- was wäre auf der rechten Seite $\mathbb{K}^{\mathcal{B}}$ oder $\mathbb{K}^{|\mathcal{B}|}$ $W^{\mathcal{B}}$ ist die Menge aller Funktionen $\phi : \mathcal{B} \longrightarrow W$
- $\mathbb{K}^{\mathcal{B}}$ ist in der Tat zu groß für unsere Zwecke

Besser:Teilraum $\mathbb{K}(\mathcal{B})$ ist frei von $\mathcal{B}$ erzeugter Vektorraum. $\mathbb{K}^{\mathcal{B}} \widehat{=}$ Menge aller Funktionen von $\mathcal{B}$ nach $\mathbb{K}$ ist das kartesische Produkt aller Faktoren $\mathbb{K}$, Indexmenge $\mathcal{B}$
$\Pi \mathbb{K}_b \widehat{=}$ alle mit $\mathcal{B}$ indizierten Familien $(\lambda_b \mid b \in \mathcal{B}) \longleftrightarrow \phi : \mathcal{B} \longrightarrow \mathbb{K}; b \longmapsto \lambda_b$. Im Falle $\mathcal{B}$ abzählbar, so wäre dies dann die folge $(\lambda_1, \lambda_2, ...,)$
Unterraum $\mathbb{K}(\mathcal{B}) \subseteq \mathbb{K}^{\mathcal{B}}$ aller Funktionen $\phi : \mathcal{B} \longrightarrow \mathbb{K}$ mit $\lambda_i \in \mathbb{K}$ und $\phi(b) = 0$ fr fast alle $b \in \mathcal{B}$, d.h. die Menge supp$(\phi) \widehat{=}$ Menge der Träger von $\phi = \{b \in \mathcal{B} \mid \phi(b) \neq 0\}$ endlich.
$\mathrm{supp}(\alpha\phi) = \mathrm{supp}(\phi), \alpha \neq 0$
$\mathrm{supp}(\phi_1 + \phi_2) \leq \mathrm{supp}(\phi_1) \cup \mathrm{supp}(\phi_2)$
$\mathbb{K}(\mathcal{B})$ ist eine $\mathbb{K}-$VR.
Basis? Standardbasis?

Definition 3.6.4. Für jedes $b \in \mathcal{B}$ ist die *charakteristische Funktion* definiert durch:

$$\begin{aligned} \mathcal{X}_b : \mathcal{B} &\longrightarrow \mathbb{K} \\ b' &\longmapsto \delta_{bb'} = \begin{cases} 1, & b' = b, \\ 0, & b' \neq b. \end{cases} \end{aligned}$$

$\phi \in \mathbb{K}(\mathcal{B})$
$\text{supp}(\phi) = \{b_1, ..., b_n\}$
$\phi = \phi(b_1)\mathcal{X}_{b_1} + ... + \phi(b_n)\mathcal{X}_{b_n}$ $X = \{\mathcal{X}_b \mid b \in \mathcal{B}\}$ ist ein Erzeugendensystem und linear-unabhängig $\alpha_1 \mathcal{X}_{\lfloor_\infty} + ... + \alpha_n \mathcal{X}_{b_n} = 0$
Einsetzen $b = b_1, b_2$, dann $\alpha_1 \mathcal{X}_{b_1}(b_1) = 1$ $\mathcal{B} \longleftarrow X$
$b \longmapsto \mathcal{X}_b$
Dann $\mathcal{B} \widehat{=} X \subseteq \mathbb{K}(\mathcal{B})$ und $\dim_{\mathbb{K}} \mathbb{K}(\mathcal{B}) = |\mathcal{B}|$

Beispiel 3.6.5. $\mathcal{B} = \{1, ..., n\}$, $\mathbb{K}(\mathcal{B}) \cong \mathbb{K}^n$; $e_i \longmapsto \mathcal{X}_i$
Koordinaten: $K_{\mathcal{B}} : V \longrightarrow K(\mathcal{B})$, $v = \lambda_1 b_1 + ... + \lambda_n b_n$ eine Darstellung durch gewisse $b_1, ..., b_n \in \mathcal{B}$.(das ist nicht abgezählt)
$v \mapsto \mathcal{X}_v = \lambda_1 \mathcal{X}_{b_1} + ... + \lambda_n \mathcal{X}_{b_n}$, $\qquad \text{supp}(\mathcal{X}_v) \subseteq \{b_1, ..., b_n\}$
$\phi \longmapsto v' = \phi(b'_1)b'_1 + ... + \phi(b'_m)b'_m$, $\qquad \text{supp}(\phi) = \{b'_1, ..., b'_m\}$

Satz 3.6.6. $\mathbb{K}_{\mathcal{B}} : V \longrightarrow \mathbb{K}(\mathcal{B})$ *ist ein Isomorphismus.*

A ist Menge, $\mathbb{K}(A)$ frei von A erzeugter $\mathbb{K}$-VR, $\phi : A \longrightarrow \mathbb{K}$, $\text{supp}(\phi)$ endlich, $\Psi \circ l_A = \Psi_{|A} = l_B \circ \psi$. Ψ soll lineare Abbildung werden mit der Eigenschaft

$$\begin{tikzcd} A \arrow[r, hook, "l_A"] \arrow[d, "\psi"] \arrow[dr, "\phi"] & \mathbb{K}(A) \arrow[d, "l_b \circ \psi"] \\ B \arrow[r, hook, "l_B"] & \mathbb{K}(B) = W \end{tikzcd}$$

3.7 Isomorphieinvarianten

Definition 3.7.1. Zwei Vektorräume V und W (über $\mathbb{K}$) heißen *isomorph*, wenn es einen Isomorphismus $f : W \to W$ gibt. Wir schreiben $V \cong W$

Satz 3.7.2. *Isomorphie ist eine Äquivalenzrelation.*

Beweis.
- $f = id_V$ ist ein Isomorphismus von V nach V, sodass für alle Vektorräume V gilt $V \cong V$
- Ist $f : V \to W$ ein Isomorphismus, so ist auch $f^{-1} : W \to V$ ein Isomorphismus. Also gilt für Zwei Vektorräume V, W: $V \cong W \Leftrightarrow W \cong V$
- Sind $f : V \to W$ und $g : W \to X$ zwei Isomorphismen, so ist auch $g \circ f$ ein Isomorphismus. Für V, W, X gilt also $V \cong W, W \cong X \Rightarrow V \cong X$

□

Was ist eine Isomorphieinvariante? Das ist eine "Größe"γ (z.B. eine Zahl, ein Zahlenpaar, eine Angabe Ja/Nein), welche man einem Vektorraum zuordnen kann, so dass gilt:

$$V \cong W \Rightarrow \gamma(V) = \gamma(W)$$

Gilt auch die Umkehrung, so nennt man γ vollständig.

Beispiel 3.7.3.
- Betrachten wir alle Rechtecke in der Ebene und als Äquivalenzrelation die Kongruenz. Da $R \equiv R' \Leftrightarrow$ Seitenlängen$(a, b) = (a', b')$ ist (a, b) eine vollständige Kongruenzinvariante.
- Wählen wir statt der Seitenlängen (a, b) den Flächeninhalt vol, oder die Diagonallänge γ_2 wählt, so ist dies eine Kongruenzinvariante, aber keine vollständige.

- Betrachten wir als Äquivalenzrelation die Ähnlichkeit, so ist (a, b) keine Invariante. Wohl aber $\gamma_0'(\mathbb{R}) = \frac{a}{b}$

Wir werden für $\gamma = dim_{\mathbb{K}}$ wählen und sogar mehr zeigen.

Proposition 3.7.4. *Sei $f : V \to W$ linear*

1. *Ist $\xi \subseteq V$ ein Erzeugendensystem und f epimorph, so ist $\xi' = f(\xi)$ ein Erzeugendensystem von W.*
2. *Ist $\mathcal{B}$ linear unabhängig und f monomorph, so ist $\mathcal{B}' = f(\mathcal{B})$ linear unabhängig.*

Beweis. 1. Es sei $w \in W$, also etwa $w = f(v)$ für ein $v \in V$. Wir können v schreiben als

$$v = \lambda_1 e_1 + \ldots + \lambda_n e_n$$

für geeignete $e_1, \ldots, e_n \in \xi, \lambda_1, \ldots \lambda_n \in \mathbb{K}$ denn ξ ist EZS für V. Dann gilt $w = f(v) = \lambda_1 f(e_1) + \ldots + \lambda_n f(e_n)$ also ist $w \in Span(\xi')$

2. Angenommen, wir hätten eine nicht-triviale Relation zwischen Elementen in $\mathcal{B}'$, also etwa

$$\lambda_1 f(b_1) + \ldots + \lambda_n f(b_n) = 0$$

mit nicht alle

$$\lambda_1, \ldots, \lambda_n$$

null. Dann hätten wir auch

$$f(\lambda_1 b_1 + \ldots + \lambda_n b_n) = 0$$

und weil f monomorph ist, auch

$$\lambda_1 b_1 + \ldots + \lambda_n b_n = 0$$

$\mathcal{B}$ war aber als linear unabhängig vorausgesetzt

□

Korollar 3.7.5. *Es sei $f : V \to V'$ ein Isomorphismus. Ist $\mathcal{B} \subseteq V$ $\begin{cases} \text{Erzeugendensystem} \\ \text{linear unabhäangig} \\ \text{eine Basis} \end{cases}$, so ist auch $\mathcal{B}' = f(\mathcal{B}) \subseteq V'$ $\begin{cases} \text{Erzeugendensystem} \\ \text{linear unabhängig} \\ \text{eine Basis} \end{cases}$*

Beweis. Beweis folgt aus der Proposition. □

Satz 3.7.6. *$V \cong V' \Leftrightarrow dim_{\mathbb{K}}(V) = dim_{\mathbb{K}}(V')$*

Beweis. "⇒": Ist $f : V \to V'$ ein Isomorphismus und $\mathcal{B}$ irgendeine Basis von V, so ist nach dem Korollar $f(\mathcal{B})$ eine Basis von V'. Da f bijektiv ist, sind $\mathcal{B}$ und $f(\mathcal{B}) = \mathcal{B}'$ gleich mächtig, also $dim_{\mathbb{K}}(V) = |\mathcal{B}| = |\mathcal{B}'| = dim_{\mathbb{K}}(V')$.

"⇐": Aus $dim_{\mathbb{K}}(V) = dim_{\mathbb{K}}(V')$ folgt, dass je zwei Basen $\mathcal{B}$ von V bzw. $\mathcal{B}'$ von V' gleich mächtig sind. Wir wählen eine Bijektion $\varphi : \mathcal{B} \to \mathcal{B}'$ und definieren $\Phi = \mathbb{K}(\varphi) : \mathbb{K}(\mathcal{B}) \to \mathbb{K}(\mathcal{B}')$ durch $\Phi(x) = x \circ \varphi : \mathcal{B}' : \mathcal{B}' \xrightarrow[\cong]{\varphi^{-1}} \mathcal{B} \xrightarrow{x} \mathbb{K}$ einen Isomorphismus Φ. Insgesamt haben wir dann

$$\begin{tikzcd} V \arrow[r, "\mathbb{K}_{\mathcal{B}}\ \cong"] \arrow[d, "f", "\cong"'] & \mathbb{K}(\mathcal{B}) \arrow[d, "\Phi=\mathbb{K}(\varphi)", "\cong"'] \\ V' \arrow[r, "\cong"] & \mathbb{K}(\mathcal{B}') \end{tikzcd}$$

Nun definieren wir $f := K_{\mathcal{B}}^{-1} \circ \mathbb{K}(\varphi) \circ K_{\mathcal{B}}$. Dies ist ein Isomorphismus; er macht das obige Diagramm *kommutativ*.

□

Bemerkung 3.7.7. Dieses f ist nicht *kanonisch*, was so viel heißt wie: es ist nicht der einzige Isomorphismus, er hängt von vielen Wahlen ab:

- Wahl von $\mathcal{B}$
- Wahl von B'
- Wahl von φ

Rang: $rg(A) = rg(T_A) = rg(L_{\mathfrak{B}\mathfrak{A}}(A))$
$T_A : \mathbb{K}^n \to \mathbb{K}^m$
$f = L_{\mathfrak{B}\mathfrak{A}}(A) : V \to W$ $\qquad\qquad dim(V) = n \qquad\qquad dim(W) = m$

Satz 3.7.8. $rg(A) = rg(T_A) = rg(L_{\mathfrak{B}\mathfrak{A}}(A))$

Beweis. Für die erste Gleichheit wenden wir die Dimensionsformel an:
$dim(W(A)) = dim(Im(T_A)) = n - dim(\mathcal{L}(A|0)) = r = rg(A)$
Für die zweite Gleichheit beachte man

$$\begin{tikzcd} V \arrow[rr, "f = L_{\mathfrak{B}\mathfrak{A}}"] \arrow[d, "K_{\mathfrak{A}}"', "\cong"] & & W \arrow[d, "\cong"', "K_{\mathfrak{B}}"] \\ {} \arrow[rr, "T_A"] & & {} \end{tikzcd}$$

Wobei $f = L_{\mathfrak{B}\mathfrak{A}} = K_{\mathfrak{B}}^{-1} \cdot T_A \cdot K_{\mathfrak{A}}$
Aber
1) $K_{\mathfrak{B}}(Im(f)) = Im(T_A)$
2) $K_{\mathfrak{B}} : Im(f) \overset{\cong}{\rightrightarrows} Im(T_A)$
Beweis fertig mit Lemma 1(1).

□

Allgemein:

Lemma 3.7.9. *Es sei*
$$\begin{array}{ccccccc} Ker(f) & \subseteq & V & \xrightarrow{f} & W & \supseteq & Im(f) \\ \varphi\downarrow\cong & & \downarrow\varphi & & \downarrow\psi & & \downarrow\cong \\ Ker(f') & \subseteq & V' & \xrightarrow{f} & W' & \supseteq & Im(f') \end{array}$$
ein kommutatives Diagramm,
dh. $\psi \circ f = f' \circ \psi$ sind φ und ψ Isomorphismen, dann gilt:

1. *$Im(f)$ wird von ψ isomorph auf das $Im(f')$ abgebildet.*
2. *$Ker(f)$ wird von φ isomorph auf den $Ker(f')$ abgebildet.*

Beweis. zu 2.
$v \in V, f(v) = 0 \Leftrightarrow \psi(f(v)) = 0 = f'(\varphi(v))$
$v \in Ker(f) \Rightarrow \varphi(v) \in Ker(f')$
Umgekehrt $v' \in Ker(f')$ also $f'(v') = 0$.
Weil φ surjektiv ist, ist $v' = \varphi(v)$ für ein $v \in V$.
Dh. $0 = f'(\varphi(v)) = \psi(f(v))$.
Weil ψ monomorph ist, folgt $f(v) = 0$, also $v \in Ker(f)$.

□

Lemma 3.7.10.
$$\begin{array}{ccccc} V & \xrightarrow{f} & W & \underset{\psi}{\xrightarrow{\cong}} & W \\ \varphi\uparrow\cong & & & & \\ V & & & & \end{array}$$

1. *$Ker(f) = Ker(\psi \circ f)$ (aber i.A. $Im(f) \neq Im(\psi \circ f)$)*
2. *$Im(f \circ \varphi) = Im(f)$ (aber i.A. $Ker(f) \neq Ker(f \circ \varphi)$)*

Bemerkung 3.7.11. zu 4.2
a) V fest, $\mathbb{K}$ fest

$$\begin{array}{lccc} & \overset{f}{\nearrow} W & \to & Hom_{\mathbb{K}}(V, W) \\ V & \varphi :\downarrow & & \downarrow \varphi_* = Hom_{\mathbb{K}}(V, \varphi) \\ & \underset{\varphi \circ f}{\searrow} Y & \to & Hom_{\mathbb{K}}(V, Y) \end{array}$$

Wieder erhält man ein Funktor.
kovariant: $W = Y : (id_W)_* = id_{Hom_{\mathbb{K}}(V,W)}$ (induzierte Abbildung)
$id_* = id$ (aber nicht die gleiche Identität!)
b) W fest, $\mathbb{K}$ fest

$$\begin{tikzcd}
 & X \arrow[dl, "f\circ\psi"'] \arrow[dd, "\psi"] \arrow[r] & Hom_{\mathbb{K}}(X,W) \ni f\circ\psi = \psi^*(f) \\
W & & \\
 & V \arrow[ul, "f"] \arrow[r] & Hom_{\mathbb{K}}(V,W) \ni f \arrow[uu, "\psi^* = Hom_{\mathbb{K}}(\psi,W)"']
\end{tikzcd}$$

Ein Funktor $VR_{/\mathbb{K}} \to VR_{/\mathbb{K}}$
Kontravariant

1. $X = V \qquad (id_V)^* = id_{Hom_{\mathbb{K}}(V,W)}$
2. $(\varphi_2 \circ \varphi_1)^* = \varphi_1^* \circ \varphi_2^*$

=======

4 Matrizen

4.1 Rechnen mit Matrizen

Mit $\mathbb{K}$ bezeichnen wir wie immer einen Körper, mit $M = \mathrm{Mat}_{m,n}(\mathbb{K})$ die Menge der $m \times n$ Matrizen über $\mathbb{K}$. Für eine Matrix schreiben wir $A = ({}_{ij})$, mit dem *Zeilenindex* $i = 1, ..., m$ und dem *Spaltenindex* $j = 1, ..., n$.

Addition

$$\begin{aligned} + : M \times M &\longrightarrow M \\ (A, B) = ((a_{ij}), (b_{ij})) &\longmapsto C = (c_{ij}) = (a_{ij} + b_{ij}) \end{aligned}$$

Die Matrix mit allen Einträgen gleich Null heißt *Nullmatrix* $0_{m,n} = 0$.
Es gilt:

1) $A + B = B + A$
2) $A + (B + C) = (A + B) + C$
3) $A + 0 = A = 0 + A$
4) zu jedem $A \in M$ gibt es ein Negatives $-A = (-a_{ij})$ mit $A + (-A) = 0$

Skalierung

$$\begin{aligned} \mathbb{K} \times M &\longrightarrow M \\ (\lambda, A) &\longmapsto \lambda A = (\lambda a_{ij}) \end{aligned}$$

Es gilt:

5) $\lambda(\mu A) = (\lambda\mu)A$
6) $\lambda(A + B) = \lambda A + \lambda B$ $(\lambda + \mu)A = \lambda A + \mu A$
7) $1 \cdot A = A;\ (-1) \cdot A = -A$
8) $0 \cdot A = 0;\ \lambda \cdot 0 = 0$

Einheitsmatrizen und Elementarmatrizen

Matrizen der Form $\tilde{E}^{kl} = (\tilde{e}^{kl}_{ij})$ $(i = 1, ..., m; j = 1, ..., n; k = 1, ..., m; l = 1, ..., n))$ mit

$$\tilde{e}^{kl}_{ij} = \begin{cases} 1 & \text{falls } (i, j) = (k, l) \\ 0 & \text{falls } (i, j) \neq (k, l) \end{cases}$$

heißen *Einheitsmatrizen*.

$$\tilde{E}^{kl} = \begin{array}{cc} & \boxed{l} \\ \boxed{k} & \begin{pmatrix} 0 & \vdots & \\ \cdots & 1 & \\ 0 & & 0 \end{pmatrix} \end{array}$$

Matrizen der Form $E^{kl} = \mathbb{1} + \tilde{E}^{kl} (k \neq l)$ heißen *Elementarmatrizen*.

Satz 4.1.1. *(i)* $\mathrm{Mat}_{m,n}(\mathbb{K})$ *ist ein* $\mathbb{K}$*-Vektorraum.*

(ii) $\mathcal{B} = \{\tilde{E}^{kl} | k = 1, ..., m; l = 1, ..., n\}$ *ist eine Basis.*

(iii) $\dim_{\mathbb{K}}(\mathrm{Mat}_{m,n}(\mathbb{K})) = m \cdot n$

Beweis. (i) folgt aus 1) bis 8).

(ii)

$$A = (a_{ij}) = \begin{matrix} a_{11}\tilde{E}^{11} & +a_{12}\tilde{E}^{12} & +... & +a_{1n}\tilde{E}^{1n} \\ +a_{21}\tilde{E}^{21} & +... & & \\ \vdots & & & \\ +a_{m1}\tilde{E}^{m1} & +... & & +a_{mn}\tilde{E}^{mn} \end{matrix} = \sum_{i,j} a_{ij}\tilde{E}^{ik}$$

Also ist $\mathcal{B}$ ein Erzeugendensystem. Für die lineare Unabhängigkeit von $\mathcal{B}$ schreibe man eine mutmaßliche Relation

$$0 = \begin{matrix} \lambda_{11}\tilde{E}^{11} & +... & +\lambda_{1n}\tilde{E}^{1n} \\ \vdots & & \\ +\lambda_{m1}\tilde{E}^{m} & +... & +\lambda_{mn}\tilde{E}^{mn} \end{matrix}$$

mit $m \cdot n$ Termen. Also erhalten wir $m \cdot n$ Gleichungen für jede Stelle (i, j):

$$0 = \lambda_{ij} \cdot \tilde{e}_{ij}^{ij} = \lambda_{ij}$$

Also sind alle $\lambda_{ij} = 0$.

(iii) folgt aus (ii).

□

Multiplikation von Matrizen

$$\begin{aligned} \mathrm{Mat}_{m,k}(\mathbb{K}) \times \mathrm{Mat}_{k,n}(\mathbb{K}) & \stackrel{\cdot}{\longrightarrow} \mathrm{Mat}_{m,n}(\mathbb{K}) \\ A,B & \longmapsto A \cdot B = AB =: C \\ (a_{ij}),(b_{jl}) & \longmapsto (c_{il}) \end{aligned}$$

mit $c_{il} := \sum\limits_{s=1}^{k} a_{is}b_{sl};\ i = 1, ..., m; l = 1, ..., n.$
Es gilt:

1) $A \cdot (B \cdot C) = (A \cdot B) \cdot C$
2) $A \cdot (B + C) = AB + AC$
   $(A + B) \cdot C = AC + BC$
3) $\lambda(A \cdot B) = (\lambda A)B = A(\lambda B)$
4) $\underset{m,k}{0} \cdot \underset{k,n}{A} = \underset{m,n}{0};\ \underset{m,k}{A} \cdot \underset{k,n}{0} = \underset{m,n}{0}$

Lemma 4.1.2. *i) Ist die* i*-te Zeile* $Z_i(A) = 0$*, so auch die* i*-te Zeile* $Z_i(AB)$*, für alle* B*.*

ii) Ist die j*-te Spalte* $S_j(B) = 0$*, so auch* $S_j(AB)$*, für alle* A*.*

Achtung!!! Das Matrizenprodukt ist im Allgemeinen **nicht** kommutativ:

$$AB \neq BA$$

.

Beispiel 4.1.3.

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \text{ Spiegelung an der } y\text{-Achse}$$

$$B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \text{ Spiegelung an der } x\text{-Achse}$$

$$AB = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ Spiegelung an der Hauptdiagonalen}$$

$$BA = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \text{ Spiegelung an der Nebendiagonalen}$$

Spezielle Produkte

a) $m = 1, n = 1, k$ beliebig: Skalarprodukt $\langle A, B \rangle$

$$\underbrace{(a_1, ..., a_k)}_{A} \cdot \underbrace{\begin{pmatrix} b_1 \\ \vdots \\ b_k \end{pmatrix}}_{B} = \underbrace{a_1 b_1 + ... + a_k b_k}_{C}$$

b) $k = 1; m, n$ beliebig:

$$\begin{array}{cccc} \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} & (b_1, ..., b_n) & = & \begin{pmatrix} a_1 b_1 & ... & a_1 b_n \\ \vdots & \ddots & \vdots \\ a_m b_1 & ... & a_m b_n \end{pmatrix} \\ A & B & = & C \end{array}$$

mit $\mathrm{rg}(C) = 1$

c) $n = 1; m, k$ beliebig, $A \in \mathrm{Mat}_{m,k}(\mathbb{K})$

$$A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} = b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

Lineare Gleichungssysteme mit m Gleichungen und k Unbekannten.

Bemerkung 4.1.4. Eine $(n \times n)$-Matrix heißt *quadratisch*.

Einsmatrix

Die quadratische $(n \times n)$-Matrix der Form

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

mit den Einträgen 1 entlang der Diagonale und 0 sonst, heißt *Einsmatrix* $\mathbb{1} = \mathbb{1}_n$. Es gilt (für $A \in \mathrm{Mat}_{mn}(\mathbb{K})$):

1)
$$\begin{array}{ccccc} \mathbb{1}_m & \cdot & A & = & A \\ m \times m & & m \times n & & m \times n \end{array}$$

2)
$$\begin{array}{ccccc} A & \cdot & \mathbb{1}_n & = & A \\ m \times n & & n \times n & & m \times n \end{array}$$

Also ist $\mathbb{1}$ ein beidseitiges neutrales Element.

Achtung! Nicht jede $(n \times n)$-Matrix besitzt ein Inverses!

Definition 4.1.5. Sei $A \in \mathrm{Mat}_{m,n}(\mathbb{K})$.

1) Eine Matrix $B \in \mathrm{Mat}_{n,m}(\mathbb{K})$ heißt *Rechtsinverses* zu A, falls gilt: $A \cdot B = \mathbb{1}_m$.
2) Eine Matrix $C \in \mathrm{Mat}_{n,m}(\mathbb{K})$ heißt *Linksinverses* zu A, falls gilt: $C \cdot A = \mathbb{1}_n$.
3) Eine quadratische Matrix $A \in \mathrm{Mat}_{n,n}(\mathbb{K})$ heißt *invertierbar*, falls es ein $\tilde{A} \in \mathrm{Mat}_{n,n}(\mathbb{K})$ gibt mit $A \cdot \tilde{A} = \tilde{A} \cdot A = \mathbb{1}_n$.

Bemerkung 4.1.6. Existiert im Fall 3) ein Inverses, so ist es eindeutig bestimmt. (Aber Rechts- und Linksinverse sind *nicht* eindeutig bestimmt.)

Beweis. Seien $\tilde{A}, \tilde{\tilde{A}}$ zwei Inverse zu A, also $A\tilde{A} = \mathbb{1}$, $A\tilde{\tilde{A}} = \mathbb{1}$. Nach Distributivität gilt:

$$A(\tilde{A} - \tilde{\tilde{A}}) = A\tilde{A} - A\tilde{\tilde{A}} = 0$$

Wenn wir diese Gleichung nun von links mit $\tilde{A}$ multiplizieren, erhalten wir:

$$0 = \tilde{A} \cdot 0 = \tilde{A}A(\tilde{A} - \tilde{\tilde{A}} = \mathbb{1} \cdot (\tilde{A} - \tilde{\tilde{A}})$$

. Damit folgt $\tilde{A} = \tilde{\tilde{A}}$. □

Notation Ist A invertierbar, so nennen wir das (eindeutig bestimmte) Inverse A^{-1}.

Lemma 4.1.7. *i) Ist A invertierbar, so auch A^{-1} und $(A^{-1})^{-1} = A$.*

ii) Sind A und B invertierbar, so auch AB und $(AB)^{-1} = B^{-1}A^{-1}$. (Man beachte die Reihenfolge!)

iii) Ist A invertierbar und $\lambda \in \mathbb{K}, \lambda \neq 0$, so auch λA und $(\lambda A)^{-1} = \frac{1}{\lambda}A^{-1}$.

Beweis. klar. □

Beispiel 4.1.8. $(A_1 \cdots A_r)^{-1} = A_r^{-1} \cdots A_1^{-1}$

Satz 4.1.9. *Eine $(n \times n)$-Matrix A ist invertierbar genau dann, wenn ihr Rang $rg(A) = n$, also maximal.*

Definition 4.1.10 (Transponieren)**.**

$$\begin{aligned} {}^{\top} : \mathrm{Mat}_{m,n}(\mathbb{K}) &\longrightarrow \mathrm{Mat}_{n,m}(\mathbb{K}) \\ (a_{ij}) = a &\longmapsto A^{\top} := (a_{ji}) \end{aligned}$$

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \longmapsto \begin{pmatrix} a_{11} & \dots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \dots & a_{nm} \end{pmatrix}$$

Es gilt:

1) $(A^{\top})^{\top} = A$
2) $(A + B)^{\top} = A^{\top} + B^{\top}$
3) $(\lambda A)^{\top} = \lambda \cdot A^{\top}$
4) $0^{\top} = 0$
5) $\mathbb{1}_n^{\top} = \mathbb{1}_n$
6) $(AB)^{\top} = B^{\top}A^{\top}$
7) Ist A invertierbar, so auch $A^{\top}$ mit $(A^{\top})^{-1} = (A^{-1})^{\top}$.

Beispiel 4.1.11.

$$
\begin{aligned}
x &= (x_1, ..., x_n) \in \mathbb{K}^n && \text{Zeilenvektor} \\
x^\top &= \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} && \text{Spaltenvektor} \\
x \cdot y^\top &= \langle x, y \rangle && x, y \text{ Zeilenvektoren} \\
x^\top \cdot y &= \langle x, y \rangle && x, y \text{ Spaltenvektoren}
\end{aligned}
$$

4.2 Matrizen und lineare Abbildungen

Wir haben bereits gesehen, dass die Multiplikation einer $(m \times n)$-Matrix C mit einem $(n \times 1)$-Spaltenvektor $x \in K^n$ einen $(m \times 1)$-Spaltenvektor ergibt und dass dies eine lineare Abbildung ist:

$$
\begin{aligned}
T_C : K^n &\to K^m \\
x &\mapsto Cx
\end{aligned}
$$

Rechenregeln:

1. $T_C(0) = 0$
2. $T_C(\lambda x) = \lambda T_C(x)$
3. $T_C(x + x') = T_C(x) + T_C(x')$

Sei V ein K-Vektorraum mit geordneter Basis $A = (a_1, \ldots, a_n)$, $\dim V = n$, W ein K-Vektorraum mit geordneter Basis $B = (b_1, \ldots, b_m)$, $\dim W = m$ und $f : V \to W$ eine lineare Abbildung.

Dann ist f durch $f(A)$ vollkommen bestimmt, also durch $f(a_1), \ldots, f(a_n) \in W$.

Jedes $f(a_j)$ ist bestimmt durch seine Koeffizienten bzgl. der Basis B:

$$f(a_j) = \sum_{i=1}^{m} \lambda_{ij} b_i$$

Also ist f durch die λ_{ij} vollkommen bestimmt. Diese λ_{ij} fasse man zu einer Matrix $\Lambda = (\lambda_{ij}) \in \mathrm{Mat}_{m,n}(K)$ zusammen und notiere:

1. $K_B(f(a_j)) = (\lambda_{1j}, \ldots, \lambda_{mj}) = S_j(\Lambda) = \Lambda e_j$ (j-te Spalte)
2. $a_j = K_A^{-1}(e_j)$

Damit haben wir $K_B \circ f = T_\Lambda \circ K_A$:

$$
\begin{tikzcd}
V \arrow[r, "f"] \arrow[d, "K_A"'] & W \arrow[d, "K_B"] \\
K^n \arrow[r, "T_\Lambda"'] & K^m
\end{tikzcd}
\qquad
\begin{tikzcd}
a_j \arrow[r, mapsto] \arrow[d, mapsto] & f(a_j) \arrow[d, mapsto] \\
e_j \arrow[r, mapsto] & S_j(\Lambda)
\end{tikzcd}
$$

$$
\begin{aligned}
\implies f &= K_B^{-1} \circ T_\Lambda \circ K_A \\
T_\Lambda &= K_B \circ f \circ K_A^{-1}
\end{aligned}
$$

Bei festgewählten Basen A und B bestimmt f eine Matrix Λ und damit T_Λ und umgekehrt bestimmt Λ das T_Λ und damit f.

Definition 4.2.1. Wir definieren zwei Funktionen:

$$\begin{aligned} M_{BA} : \mathrm{Hom}_K(V,W) &\to \mathrm{Mat}_{m,n}(K) \\ f &\mapsto C = (c_{ij}) \end{aligned}$$

mit $c_{ij} = K_B(f(a_j))$ und

$$\begin{aligned} L_{BA} : \mathrm{Mat}_{m,n}(K) &\to \mathrm{Hom}_K(V,W) \\ C &\mapsto f = K_B^{-1} \circ T_C \circ K_A \end{aligned}$$

Beispiel 4.2.2.

1. $f = 0: \ M_{BA}(0) = 0$
   $C = 0: \ L_{BA}(0) = 0$
2. $V = W, \ A = B$:
   $f = \mathrm{id}: \ M_{AA}(\mathrm{id}) = \mathbb{1}$
   $C = \mathbb{1}: \ L_{AA}(\mathbb{1}) = \mathrm{id}$
3. $V = W = \mathbb{R}^2, \ A = (e_1, e_2), \ B = (e_1, e_2 + e_1), \ f = \mathrm{id}:$
   $M_{BA}(\mathrm{id}) = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$
4. $V = W = \mathbb{R}^2, \ A = (e_1, e_2), \ B = (e_1, e_2 + e_1), \ C = \mathbb{1}:$
   $L_{BA}(\mathbb{1}) = ((x, y) \mapsto (x, -x + y)$ (Scherung)
5. $V = K^n, \ W = K^m, \ A = (e_1, \ldots, e_n), \ B = (e_1, \ldots, e_m):$
   $M_{BA}(T_C) = C, \ L_{BA}(C) = T_C$, denn $K_A = \mathrm{id}, \ K_B = \mathrm{id}$

Satz 4.2.3. *Sei A bzw. B eine Basis für V bzw. W. M_{BA} und L_{BA} sind zueinander invers und linear, d.h. sie sind Isomorphismen von Vektorräumen.*

$$\begin{tikzcd} \mathrm{Hom}(V,W) \arrow[r, shift left, "M_{BA}"] & \mathrm{Mat}_{m,n}(K) \arrow[l, shift left, "L_{BA}"] \end{tikzcd}$$

Seien $V = K^n$, $W = K^m$ und A und B die Standardbasen von V, W. Nun betrachte man die Funktionen

$$\begin{aligned} L_{BA} = T: \ \mathrm{Mat}_{n,m}(K) &\to \mathrm{Hom}(K^n, K^m) \\ C &\mapsto K_B^{-1} \circ T_C \circ K_A = T_C \end{aligned}$$

Lemma 4.2.4. *T ist linear und multiplikativ, d.h. $T_0 = 0$, $T_{\lambda C} = \lambda T_C$, $T_{A+B} = T_A + T_B$, sowie für $n = m$ gilt $T_{\mathbb{1}} = \mathrm{id}$ und $T_{A^{-1}} = T_A^{-1}$, als auch $T_{AB} = T_A T_B$ für passende A und B.*

Die Beweise reduzieren sich auf Rechenregeln für Matrizen.

Es sei $\varphi : W \to Y$ eine feste lineare Abbildung. Dann haben wir die Postkomposition

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[rr, bend right, "\varphi \circ f"'] & W \arrow[r, "\varphi"] & Y \end{tikzcd}$$

und können dies als Funktion

$$\begin{aligned} \varphi_* = \mathrm{Hom}_K(V, \varphi): \ \mathrm{Hom}_K(V,W) &\to \mathrm{Hom}_K(V,Y) \\ f &\mapsto \varphi \circ f \end{aligned}$$

betrachten.

Entsprchend sei $\psi : X \to V$ eine feste lineare Abbildung. Dann haben wir die Präkomposition

$$\begin{tikzcd} X \arrow[r, "\psi"] \arrow[rr, bend right, "f \circ \psi"'] & V \arrow[r, "f"] & W \end{tikzcd}$$

und können dies als Funktion

$$\begin{aligned} \psi^* = \mathrm{Hom}_K(\psi, W): \ \mathrm{Hom}_K(V, W) &\to \mathrm{Hom}_K(X, W) \\ f &\mapsto f \circ \psi \end{aligned}$$

betrachten.

Lemma 4.2.5. $\varphi_*: \ \mathrm{Hom}_K(V, W) \to \mathrm{Hom}_K(V, Y)$ *ist linear:*

1. $\varphi_*(0) = 0$

2. $\varphi_*(\lambda f) = \lambda \varphi_*(f)$

3. $\varphi_*(f + g) = \varphi_*(f) + \varphi_*(g)$

1'. $\varphi \circ 0 = 0$

2'. $\varphi \circ (\lambda f) = \lambda(\varphi \circ f)$

3'. $\varphi \circ (f + g) = (\varphi \circ f) + (\varphi \circ g)$

Lemma 4.2.6. $\psi^*: \ \mathrm{Hom}_K(V, W) \to \mathrm{Hom}_K(V, Y)$ *ist linear:*

1. $\psi^*(0) = 0$

2. $\psi^*(\lambda f) = \lambda \psi^*(f)$

3. $\psi^*(f + g) = \psi^*(f) + \psi^*(g)$

1'. $0 \circ \psi = 0$

2'. $(\lambda f) \circ \psi = \lambda(f \circ \psi)$

3'. $(f + g) \circ \psi = (f \circ \psi) + (g \circ \psi)$

Beweis. (Lemma 4.2.5).

$$\varphi_*(0) = \varphi \circ 0 = 0$$

$$\begin{aligned} (\varphi_*(\lambda f))(x) &= (\varphi \circ (\lambda f))(x) \\ &= \varphi(\lambda f(x)) = \lambda \varphi(f(x)) \\ &= \lambda((\varphi \circ f)(x)) = \lambda(\varphi_*(f)(x)) \end{aligned}$$

$$\begin{aligned} (\varphi_*(f + g))(x) &= (\varphi \circ (f + g))(x) = \varphi((f + g)(x)) \\ &= \varphi(f(x) + g(x)) = \varphi(f(x)) + \varphi(g(x)) \\ &= (\varphi \circ f)(x) + (\varphi \circ g)(x) = \varphi_*(f)(x) + \varphi_*(g)(x) \end{aligned}$$

□

Beweis. (Satz 4.2.3). Sei $M = M_{BA}$, $L = L_{BA}$, $f \in \mathrm{Hom}(V, W)$ und $x \in V$.

$$\begin{aligned} L(M(f)) &= K_B^{-1} \circ T_{M(f)} \circ K_A \\ &= K_B^{-1} \circ K_B \circ f \circ K_A^{-1} \circ K_A = f \end{aligned}$$

$$\begin{aligned} M(L(C))(x) &= M(K_B^{-1} \circ T_C \circ K_A)(x) \\ &= K_B(K_B^{-1} \circ T_C \circ K_A)K_A^{-1}(x) = T_C(x) = Cx \\ \implies M(L(C)) &= C \end{aligned}$$

Also $L \circ M = \mathrm{id}$ und $M \circ L = \mathrm{id}$.

Es genügt, die Linearität von L zu zeigen. Dies folgt mithilfe der gerade bewiesenen Lemmata:

$$\begin{aligned} L(\lambda C) &= K_B^{-1} \circ T_{\lambda C} \circ K_A \\ &= K_B^{-1} \circ (\lambda T_C) \circ K_A \\ &= \lambda (K_B^{-1} \circ T_C \circ K_A) \\ &= \lambda L(C) \end{aligned}$$

$$\begin{aligned} L(A+B) &= K_B^{-1} \circ T_{A+B} \circ K_A \\ &= K_B^{-1} \circ (T_A + T_B) \circ K_A \\ &= K_B^{-1} \circ T_A \circ K_A + K_B^{-1} \circ T_B \circ K_A \\ &= L(A) + L(B) \end{aligned}$$

□

Basis für $\mathrm{Hom}_K(V, W)$

Da L_{BA} ein Isomorphismus ist, können wir die Basis $\tilde{E}^{kl}$ der Standardeinheitsmatrizen nach $\mathrm{Hom}(V, W)$ schicken und erhalten

$$f^{kl} : V \to W, \; a_j \mapsto \begin{cases} b_k & \text{, falls } j = l \\ 0 & \text{, falls } j \neq l \end{cases}$$

Nach Konstruktion gilt nun

$$\begin{aligned} L_{BA}(\tilde{E}^{kl}) &= f^{kl} \\ M_{BA}(f^{kl}) &= \tilde{E}^{kl} \end{aligned}$$

In Lemma 4.2.4 haben wir für die Funktion T auch schon multiplikative Eigenschaften aufgelistet, daraus folgt

$$\begin{aligned} M_{AA}(\mathrm{id}) &= \mathbb{1} \\ L_{AA}(\mathbb{1}) &= \mathrm{id} \end{aligned}$$

Für die multiplikativen Eigenschaften von M und L betrachten wir drei Vektorräume V, W und Z mit geordneten Basen $A \subseteq V$, $B \subseteq W$ und $C \subseteq Z$ Dann haben wir

$$\begin{aligned} \mathrm{Hom}(W, Z) \times \mathrm{Hom}(V, W) &\to \mathrm{Hom}(V, Z) \\ \mathrm{Mat}_{l,m}(K) \times \mathrm{Mat}_{m,n}(K) &\to \mathrm{Mat}_{l,n}(K) \end{aligned}$$

$$\begin{aligned} f, g \qquad &\mapsto f \circ g \\ A, B \qquad &\mapsto A \cdot B \end{aligned}$$

Satz 4.2.7.

$$\begin{aligned} M_{AA}(\mathrm{id}) &= \mathbb{1} \\ L_{AA}(\mathbb{1}) &= \mathrm{id} \end{aligned}$$

$$\begin{aligned} M_{CA}(f \circ g) &= M_{CB}(f) \cdot M_{BA}(g) \\ L_{CA}(A \cdot B) &= L_{CB}(A) \circ L_{BA}(B) \end{aligned}$$

Beweis. Die ersten beiden Ausgaben wurden bereits bewiesen, bei den anderen genügt es, diese für L zu beweisen.

$$\begin{aligned} L_{CA}(AB) &= K_C^{-1} \circ T_{AB} \circ K_A \\ &= K_C^{-1} \circ T_A \circ T_B \circ K_A \\ &= K_C^{-1} \circ T_A \circ K_B \circ K_B^{-1} \circ T_B \circ K_A \\ &= L_{CB}(A) \circ L_{BA}(B) \end{aligned}$$

□

Korollar 4.2.8. *$f : V \to W$ ist genau dann ein Isomorphismus, wenn $M_{BA}(f)$ invertierbar ist; in diesem Fall ist $M_{BA}(f^{-1}) = M_{BA}(f)^{-1}$.*

$C \in \mathrm{Mat}_{m,n}(K)$ ist genau dann invertierbar, wenn $L_{BA}(C)$ ein Isomorphismus ist; in diesem Fall ist $L_{AB}(C^{-1}) = L_{BA}(C)^{-1}$, und speziell ist $L_{AB}(\mathbb{1}) = L_{BA}(\mathbb{1})^{-1}$.

4.3 Basiswechsel

Angenommen, wir haben bislang einer linearen Abbildung $f : V \to W$ bzgl. der Basen $A \subseteq V$ und $B \subseteq W$ eine Matrix zugeordnet, nämlich $C = M_{BA}(f)$. Wie ändert sich C, wenn wir andere Basen A' und B' nehmen?

Wir benutzen Satz 4.2.7 zweimal:

$$
\begin{tikzcd}
W \arrow[d, "\mathrm{id}"'] & V \arrow[l, "f"'] \\
W & V \arrow[l, "f"] \arrow[u, "\mathrm{id}"']
\end{tikzcd}
\qquad
\begin{tikzcd}
B \arrow[d, "M_{B'B}(\mathrm{id})"'] & A \arrow[l, "M_{BA}(f)"'] \\
B' & A' \arrow[l, "M_{B'A'}(f)"] \arrow[u, "M_{AA'}(\mathrm{id})"']
\end{tikzcd}
$$

$$
\begin{aligned}
C' = M_{B'A'}(f) &= M_{B'A'}(\mathrm{id} \circ f \circ \mathrm{id}) \\
&= M_{B'B}(\mathrm{id}) \cdot M_{BA}(f) \cdot M_{AA'}(\mathrm{id}) \\
&= M_{B'B}(\mathrm{id}) \cdot M_{BA}(f) \cdot M_{A'A}(\mathrm{id})^{-1} \\
&= \Omega_{B'B} \cdot C \cdot \Omega_{A'A}^{-1}
\end{aligned}
$$

Die Matrix $\Omega = \Omega_{A'A}$ heißt Basiswechselmatrix für den Wechsel von A nach A'. Ihre j-te Spalte $S_j(\Omega_{A'A})$ enthält die Koeffizienten, um a_j durch die $a'_1, \ldots, a'_k$ darzustellen:

$$
\begin{aligned}
a_j &= \sum_{k=1}^{n} \omega_{kj} \cdot a'_k \\
K_{A'}(a_j) &= (\omega_{1j}, \ldots, \omega_{nj}) = S_j(\Omega)
\end{aligned}
$$

Proposition 4.3.1.

1. $\Omega_{AA} = 1$
2. $\Omega_{A'A} = \Omega_{AA'}^{-1}$
3. $\Omega_{A''A'} \cdot \Omega_{A'A} = \Omega_{A''A}$
4. *Es sei A eine Basis von V und D eine invertierbare Matrix in $\mathrm{Mat}_n(K)$. Dann git es eine Basis A' von V mit $\Omega_{A'A} = D$.*

Beweis. $\varphi = L_{AA}(D) : V \to V$ ist ein Automorphismus. Also ist $A' = (\varphi^{-1}(a_1), \ldots, \varphi^{-1}(a_n))$ wieder eine Basis, wenn $A = (a_1, \ldots, a_n)$ war. Dann haben wir

$$
\begin{aligned}
D = M_{AA}(\varphi) &= M_{AA'}(\varphi) \cdot M_{A'A}(\mathrm{id}) \\
&= M_{A'A}(\varphi^{-1}) \cdot \Omega_{A'A} \\
&= \mathbb{1} \cdot \Omega_{A'A} = \Omega_{A'A}
\end{aligned}
$$

□

Normalformen

Wir werden noch sehen, dass wir die zu einer linearen Abbildung $f : V \to W$ bzgl. zweier Basen $A \subseteq V$ und $B \subseteq W$ gehörende Matrix $C = M_{BA}(f)$ durch geschickte Wahl von A und B auf eine sehr einfache Form bringen können.

Anders gesagt, multipliziert man eine Matrix C von links und rechts mit invertierbaren Matrizen R und S, dann kann man $C' = RCS$ auf sehr einfache Form

$$C' = \begin{pmatrix} \mathbb{1}_r & 0_{r,n-r} \\ 0_{m-r,r} & 0_{m-r,n-r} \end{pmatrix}$$

bringen, wobei $r = rg(f) = rg(C)$ der Rang von C ist.

Interessanter ist, wenn C eine $n \times n$-Matrix ist, die also zu einem Endomorphismus $f : V \to V$ bzgl. einer Basis A gehört, also $C = M_{AA}(f)$. Dann kann man sich auf simultane Basiswechsel beschränken und versucht, durch geschickte Wahl von A die Matrix C möglichst einfach zu machen.

Anders gesagt, man multipliziert C von links mit R und von rechts mit R^{-1}, also $C' = RCR^{-1}$. Dies nennt man eine Konjugation. Daran kann man nun viel mehr ablesen.

Und als drittes wird man dann noch die zulässigen R einschränken und nur mit speziellen R konjugieren.

Lemma 4.3.2. *Es sei $f : V \to W$ linear und es seien $\psi : V \to V$ und $\phi : W \to W$ Automorphismen. Dann gilt:*

(i) $ker(\phi \circ f \circ \psi) = \psi^{-1}(ker(f))$

(ii) $im(\phi \circ f \circ \psi) = \phi(ker(f))$

Lemma 4.3.3. *Sei $M \in Mat_{m,n}(\mathbb{K}), A \in Mat_{m,m}(\mathbb{K})$ und $B \in Mat_{n,n}(\mathbb{K})$ sowie A und B invertierbar. Dann gilt:*

(i) $\mathcal{L}(AMB|0) = ker(T_{AMB}) = T_B^{-1}(\mathcal{L}(M|0)) = T_{B^{-1}}(\mathcal{L}(M|0)) = B^{-1}(\mathcal{L}(M|0))$
$:= \{B^{-1}x | x \in \mathcal{L}(M|0)\}$

(ii) $\mathcal{W}(AMB) = T_A(\mathcal{W}(M)) = A\mathcal{W}(M) = \{y \in \mathbb{K}^m | \mathcal{L}(AMB|y) \neq \emptyset\}$

Beweis:
Das erste Lemma war eine Übungsaufgabe, das zweite folgt aus dem ersten Lemma mit Isomorphismen

$$M_{\mathfrak{B}\mathfrak{A}} : \quad Hom_{\mathbb{K}}(V, W) \rightleftarrows Mat_{m,n}(\mathbb{K}) \quad : L_{\mathfrak{B}\mathfrak{A}}$$

Korollar 4.3.4 (→ Übungsaufgabe 55). *(i)* $rg(\phi \circ f \circ \psi) = rg(f)$

(ii) $rg(AMB) = rg(M)$

Beweis: Wir wissen bereits aus Lemma 1:
$ker(\phi \circ f \circ \psi) = \psi^{-1}(ker(f))$, also ist $\dim(ker(\phi \circ f \circ \psi)) = \dim(\psi^{-1}(ker(f))) = \dim(ker(f))$, weil ψ^{-1} ein Isomorphismus ist. Weiter gilt nach der Dimensionsformel:

$$n = \dim(V) = \dim(im(f)) + \dim(ker(f)) = \dim(im(\phi \circ f \circ \psi)) + \dim(im(\phi \circ f \circ \psi)),$$

daraus folgt sofort die Behauptung $rg(f) = rg(\phi \circ f \circ \psi)$. □

4.4 Spaltenumformungen

Wir kennen für Zeilen folgende Umformungen:

$$\begin{array}{lll} \mathfrak{M}_i[\lambda]: & Mat_{m,n}(\mathbb{K}) \to Mat_{m,n}(\mathbb{K}) & \textit{Multiplikation der i-ten Zeile mit } \lambda \\ \mathfrak{A}_{ij}: & Mat_{m,n}(\mathbb{K}) \to Mat_{m,n}(\mathbb{K}) & \textit{Addition der j-ten zur i-ten Zeile} \\ \mathfrak{V}_{ij}: & Mat_{m,n}(\mathbb{K}) \to Mat_{m,n}(\mathbb{K}) & \textit{Vertauschen der i-ten und j-ten Zeile} \end{array}$$

Analog gilt für Spaltenumformungen:

$$\begin{array}{lll} \mathfrak{M}'_i[\lambda]: & Mat_{m,n}(\mathbb{K}) \to Mat_{m,n}(\mathbb{K}) & \textit{Multiplikation der i-ten Zeile mit } \lambda \\ \mathfrak{A}'_{ij}: & Mat_{m,n}(\mathbb{K}) \to Mat_{m,n}(\mathbb{K}) & \textit{Addition der j-ten zur i-ten Zeile} \\ \mathfrak{V}'_{ij}: & Mat_{m,n}(\mathbb{K}) \to Mat_{m,n}(\mathbb{K}) & \textit{Vertauschen der i-ten und j-ten Zeile} \end{array}$$

Voller Gaußalgorithmus

Satz 4.4.1. *Jede Matrix $A \in Mat_{m,n}(\mathbb{K})$ mit $rg(A) = r$ kann durch Zeilen- und Spaltenumformungen in eine Form wie $\begin{pmatrix} \mathbb{1} & 0 \\ 0 & 0 \end{pmatrix}$ gebracht werden.*

Beweis:

(1) Der einfache Gauß-Algorithmus, das heißt wenn nur Zeilenumformungen benutzt werden, bringt A zunächst auf die Form

(2) Durch Vertauschen der Spalten $j_1, \dots, j_r$ auf die Plätze $1, \dots, r$, also durch ein "nach vorne ziehen", erhalten wir

(3) Dividieren der Spalte 1 durch π_1, ..., Spalte r durch π_r, erhalten wir

(4) Als letzten Schritt werden alle Spalten -ausgeräumt"

4.5 Spezielle Matrizen

Sei im Folgenden n fest gewählt.

a) **Elementarmatrizen**
Für $1 \leq k \neq l \leq n$ sei

$$E^{kl} = \mathbb{1}_n + \widetilde{E}^{kl} = \begin{pmatrix} 1 & & & & 0 \\ & \ddots & & 1 & \\ & & 1 & & \\ & & & \ddots & \\ 0 & & & & 1 \end{pmatrix}$$

die Elementarmatrix. Dann ist E^{kl} invertierbar mit $(E^{kl})^{-1} = \mathbb{1} - \widetilde{E}^{kl}$, denn es gilt

$$(\mathbb{1} + \widetilde{E}^{kl})(\mathbb{1} - \widetilde{E}^{kl}) = \mathbb{1}^2 - (\widetilde{E}^{kl})^2 = \mathbb{1} - 0 = \mathbb{1}.$$

Etwas allgemeiner gilt für $0 \neq \lambda \in \mathbb{K}: \quad E^{kl}[\lambda] = \mathbb{1} + \lambda \widetilde{E}^{kl}$,

$$E^{kl}[\lambda]: \qquad \begin{aligned} e_k &\longmapsto e_k \\ e_l &\longmapsto e_l + \lambda e_k \\ e_i &\longmapsto e_i (i \neq k, l) \end{aligned}$$

Geometrisch betrachtet erhalten wir Scherungen:

$$n = 2: \qquad E^{12}[\lambda] = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$$

Lemma 4.5.1. *Für die Elementarmatrizen gelten*

(1) $E^{kl}[0] = \mathbb{1}$

(2) $E^{kl}[\lambda] \cdot E^{kl}[\mu] = E^{kl}[\lambda + \mu]$, *insbesondere*
$(E^{kl}[\lambda])^s = E^{kl}[s\lambda]$

(3) $E^{kl}[\lambda]^{-1} = E^{kl}[-\lambda]$

(4) $E^{ij}[\lambda] \cdot E^{kl}[\mu]E \cdot^{ij} [-\lambda] \cdot E^{kl}[-\mu] = \mathbb{1}, \quad i \neq l, j \neq k$

(5) $E^{ij}[\lambda] \cdot E^{jl}[\mu] \cdot E^{ij}[-\lambda] \cdot E^{jl}[-\mu] = E^{il}[\lambda\mu], \quad i \neq l$

(6) $E^{ij}[\lambda] \cdot E^{ki}[\mu] \cdot E^{ij}[-\lambda] \cdot E^{ki}[-\mu] = E^{kj}[-\lambda\mu], \quad j \neq k$

(7) $(E^{ij}[1] \cdot E^{ji}[1]^{-1} \cdot E^{ij}[1])^4 = \mathbb{1}$

Beweis:
Man kann solche Formeln durch Nachrechnen beweisen, oder man kontrolliert, dass beide Seiten auf den Standardeinheitsvektoren $e_1, \ldots, e_n$ das gleiche tun. Eine gute Methode ist auch folgende: Man rechnet das im kleinsten Fall $n = 2$ oder einem anderen kleinen Wert nach, und nennt die kleinen Matrizen kurz $e^{ij}[\lambda]$. Dann sieht man schnell, dass die Gleichungen auch für die durch $\mathbb{1}$ ergänzte große Elementarmatrix

$$E^{ij}[\lambda] = \begin{pmatrix} e^{ij}[\lambda] & 0 \\ 0 & \mathbb{1} \end{pmatrix}$$

gelten. Und zum Schluss benutzt man $E^{ij}[\lambda] = PE^{ij}P^{-1}$ für ein geeignetes P, wenn k, l beliebig ist.□
Bemerkung 4.5.2. Ausdrücke der Form $aba^{-1}b^{-1}$ nennt man einen *Kommutator*.

b) **Streckungsmatrizen**
Sei $k = 1, \ldots, n$ und $\lambda \in \mathbb{K}$. Dann definieren wir die Streckungsmatrix als

$$D_k[\lambda] = \begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & \lambda & & \\ & & & \ddots & \\ 0 & & & & 1 \end{pmatrix}$$

Lemma 4.5.3. *Es gilt*

(1) $D_k[1] = \mathbb{1}$

(2) $D_k[\lambda] \cdot D_k[\mu] = D_k[\lambda\mu]$, *insbesondere*
$D_k[\lambda]^s = D_k[\lambda^s]$

(3) $D_k[\lambda]$ *ist genau dann invertierbar, wenn* $\lambda \neq 0$ *ist, und in diesem Fall ist* $D_k[\lambda]^{-1} = D_k[\frac{1}{\lambda}]$

(4) $D_i[\lambda] \cdot D_j[\mu] = D_j[\mu] \cdot D_i[\lambda]$

c) **Diagonalmatrizen**
Matrizen der Form $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} =: D[\lambda_1, \ldots, \lambda_n]$ heißen Diagonalmatrizen.
Lemma 4.5.4. *Es gilt*

(1) $D[\lambda_1, \ldots, \lambda_n] = D[\lambda_1] \cdot \ldots \cdot D[\lambda_n]$

(2) $D[1, \ldots, 1] = \mathbb{1}$

(3) $D[\lambda_1, \ldots, \lambda_n] \cdot D[\mu_1, \ldots, \mu_n] = D[\lambda_1\mu_1, \ldots, \lambda_n]_m u_n$

(4) $D[\lambda_1, \ldots, \lambda_n]$ *ist genau dann invertierbar, wenn* $\lambda_1, \ldots, \lambda_n \neq 0$, *in diesem Fall ist* $D[\lambda_1, \ldots, \lambda_n]^{-1} = D[\lambda_1^{-1}, \ldots, \lambda_n^{-1}]$

d) **Permutationsmatrizen**
Eine *Permutation* auf den Ziffern $1, \ldots, n$ ist eine Bijektion $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$, zusammen mit einer Abbildung
$$f_\sigma : \quad \mathbb{K}^n \to \mathbb{K}^n, \quad e_k \longmapsto e_{\sigma(k)},$$
die also eine Permutation der Standardbasis darstellt. Die dazugehörige Matrix sei
$$P_\sigma = M_{SS}(F_\sigma).$$
Sie hat in jeder Zeile i genau einen von Null verschiedenen Eintrag, und zwar an der Stelle $j = \sigma^{-1}(i)$ eine 1, und sie hat in jeder Spalte j genau einen von Null verschiedenen Eintrag, und zwar an der Stelle $i = \sigma(j)$ eine 1.

Beispiel 4.5.5.

- $\sigma = id$
- Transposition $(i\ j) : \quad i \longmapsto j, \quad j \longmapsto i, \quad k \longmapsto k$ für $k \neq i, j$

Lemma 4.5.6. *Es gilt*

(1) $P_{id} = \mathbb{1}$

(2) $P_\sigma \cdot P_\tau = P_{\sigma \circ \tau}$

(3) P_σ *ist invertierbar und es ist* $P_\sigma^{-1} = P_{\sigma^{-1}}$

(4) $P_{(i\ j)}1{-}1 = P_{(i\ j)}$ *für eine Transposition* $(i\ j)$

Lemma 4.5.7.

1. $P_\sigma E^{ij}[\lambda] P_\sigma^{-1} = E^{\sigma(i)\sigma(j)}[\lambda]$
2. $P_\sigma D_k[\lambda] P_\sigma^{-1} = P_{\sigma(k)}[\lambda]$
3. $P_\sigma D(\lambda_1, ..., \lambda_k) P_\sigma^{-1} = D(\lambda_{\sigma^{-1}(1)}, ..., \lambda_{\sigma^{-1}(k)})$
4. $P_\sigma P_\varepsilon P_\sigma^{-1} = P_{\sigma\varepsilon\sigma^{-1}}$
5. $D_j[\lambda^{-1}] E^{ij}[\mu] D_i[\lambda] = Eij[\lambda\mu]$

Was bewirken diese Matrizen?

Proposition 4.5.8. *Sei* $A \in Mat_{m,n}(\mathbb{K}), \lambda \neq 0$

1. $E^{ij} A = \mathfrak{A}_{ij}(A)$ *Addition der j-ten Zeile zur i-ten*
2. $D_i[\lambda] A = \mathfrak{M}_i[\lambda](A)$ *Multiplikation der i-ten Zeile mit* λ
3. $P_{ij} A = \mathfrak{V}_{ij}(A)$ *Vertauschen der i-ten und j-ten Zeile*

Proposition 4.5.9. *Sei* $A \in Mat_{m,n}(\mathbb{K}), E^{ij}, D_i[\lambda], P_{\sigma(j)} \in Mat_{n,m}(\mathbb{K}) \lambda \neq 0$

1. $AE^{ij} = \mathfrak{A}'_{ij}(A)$ *Addition der j-ten Zeile zur i-ten*
2. $AD_i[\lambda] = \mathfrak{M}'_i[\lambda](A)$ *Multiplikation der i-ten Zeile mit* λ
3. $AP_{\sigma(j)} = \mathfrak{V}'_{ij}(A)$ *Vertauschen der i-ten und j-ten Zeile*

Man stellt also fest:

1. Der vollständige Gauß-Algorithmus, angewandt auf A, ist nichts anderes als sukzessives Multiplizieren von links oder rechts mit Elementarmatrizen, Streck- oder Vertauschmatrizen. Gauß macht aus der Matrix A die Matrix $A' = ZAS$, wobei Z das Produkt aller verwendeten Elementar-, Streck- und Vertauschmatrizen ist, in der richtigen Reihenfolge sortiert nach Zeilen- und Spaltenumformungen. Hier ist $r = rg(A)$.

2. A sei quadratisch und invertierbar.
   In diesem Fall kommt man ohne Spaltenumformungen aus. A wird durch Linksmultiplikation mit Z zu $A' = \mathbb{1}\ Z$ ist gleich A^{-1}. **Wichtig:** Man muss nicht wissen, ob A invertierbar ist, der Algorithmus bricht vorher ab.

Beispiel 4.5.10. $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
Linksmultiplikation mit $E21[-3]$ liefert:
$\begin{pmatrix} 1 & 2 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix}$
Linksmultiplikation mit $E^{12}[1]$ liefert:
$\begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ -3 & 1 \end{pmatrix}$
Linksmultiplikation mit $D_2[-\frac{1}{2}]$ liefert schließlich das Inverse A^{-1}. Die Inverse Matrix steht nun da, wo ursprünglich die Einheitsmatrix stand und an Stelle von A steht nun die Einheitsmatrix.
$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix} = A^{-1}$

4.6 Normalformen

Sei im Folgenden $f : V \longrightarrow W$ lineare Abbildung zwischen zwei Vektorräumen V mit $dim(V) = n$ und W mit $dim(W) = m$, sei weiter der $rg(f) = r$. Setze $V' := Ker(f)$ mit Basis $\mathcal{A}'' = (a''_1, ..., a''_{n-r})$. Ergänze nun $\mathcal{A}''$ zu eine Basis $\mathcal{A}$ von V, also $\mathcal{A} = (\mathcal{A}', \mathcal{A}'') = (a'_1, ..., a'_r, a''_1, ..., a''_{n-r})$. Dann gilt: (1) $f(a''_i) = 0$ mit $i = 1, ..., n-r$. Das ist klar, denn Elemente im Kern werden auf Null geschickt.
(2) Setze $\mathcal{B}' := f(A') = (f(a'_1), ..., f(a'_r))$. $\mathcal{B}'$ ist eine geordnete Basis von $im(f) \subseteq W$.
Warum ist das so? $0 = \lambda_1 f(a'_1) + ... + \lambda_r f(a'_r) = f(\lambda_1 a'_1 + ... + \lambda_r a'_r)$, da gleich Null muss $v' = (\lambda_1 a'_1 + ... + \lambda_r a'_r) \in Ker(f)$. v' ist mit der Basis des Kerns darstellbar $v' = \mu_1 a''_1 + ... + \mu_{n-r} a''_{n-r}$. Daraus folgt, dass $0 = (-\lambda_1) a'_1 + ... + (-\lambda_r) a'_r + \mu_1 a''_1 + ... + \mu_{n-r} a''_{n-r}$ und somit gelten muss, dass alle $\lambda_1 = ... = \lambda_r = 0$ und $\mu_1 = ... = \mu_{n-r} = 0$. Ergänze zu einer Basis $\mathcal{B} = (\mathcal{B}', \mathcal{B}'') = (f(a'_1), ..., f(a'_r), b_1, ..., b_n - r)$. Dann ist

$$M_{\mathcal{B},\mathcal{A}}(f) = \left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right)$$

Satz 4.6.1. *Für jede lineare Abbildung $f : V \longrightarrow W$ zwischen Vektorräumen gibt es Basen $\mathcal{A}$ in V und $\mathcal{B}$ in W, so dass gilt*

$$M_{\mathcal{B},\mathcal{A}}(f) = \left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right) \text{ mit } r = rg(f)$$

Insbesondere gilt: Wenn f ein Isomorphismus ist und $m = n$ (d.h. V und W haben gleiche Dimension), dann ist $M_{\mathcal{B},\mathcal{A}}(f) = 1_n$

Beweis. nach obiger Konstruktion. □

Satz 4.6.2. *Für jede Matrix $M \in Mat_{m,n}(\mathbb{K})$ gibt es invertierbare Matrizen A, B mit*

$$BMA = \left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right)$$

Insbesondere gilt: Wenn M invertierbar ist, so ist $BMA = \mathbb{1}$

Beweis.

1. Gauß-Algorithmus mit $B = Z$ Produkt der Zeilenumformungen und $A = S$ Produkt der Spaltenumformungen.

2. $M_{\mathcal{B},\mathcal{A}} : Hom(V, W) \longleftrightarrow Mat_{m,n}(\mathbb{K}) : L$. $\mathcal{A}$ ist Basiswechselmatrix in V und $\mathcal{B}$ ist Basiswechselmatrix in W.

□

5 Gruppen

Wir haben schon etliche Gruppen gesehen und wollen diesen Begriff nun axiomatisieren.

Definition 5.0.3. Eine *Gruppe* G ist eine nicht-leere Menge mit einer Verknüpfung, genannt *Multiplikation*

$$\begin{array}{rccl} \mu: & G \times G & \longrightarrow & G \\ & (x,y) & \longmapsto & \mu(x,y) = x \cdot y = xy, \end{array}$$

so dass folgende drei Axiome erfüllt sind:

(G1) *Assoziativgesetz*: Für alle $x, y, z \in G$

$$x \cdot (y \cdot z) = (x \cdot y) \cdot z.$$

(G2) *Neutrales Element*: Es gibt ein Element $1 \in G$ so dass für alle $x \in G$

$$x \cdot 1 = 1 \cdot x = x.$$

(G3) *Inverse Elemente*: Für jedes $x \in G$ gibt es ein $x^{-1} \in G$, so dass

$$x \cdot x^{-1} = x^{-1}x = 1.$$

Gilt zusätzlich das

(G4) *Kommutativgesetz*: Für alle $x, y \in G$ gilt $x \cdot y = y \cdot x$.

so nennt man G *kommutativ* oder *abelsch*.

Notation

1) Das Produkt wird meist ohne Punkt xy geschrieben und das Inverse x^{-1}. Wir werden gleich sehen, dass es zu jedem x *nur ein Inverses* gibt. Das neutrale Element wird oft auch mit e bezeichnet.
2) In einer abelschen Gruppe schreibt man das Produkt oft (aber nicht immer) als Addition $x + y$, das neutrale Element als 0, und das Inverse als $-x$. Die Axiome lauten dann

$$\begin{aligned} x + (y + z) &= (x + y) + z \\ x + 0 &= 0 + x = x \\ x + (-x) &= (-x) + x = 0 \\ x + y &= y + x \end{aligned}$$

Lemma 5.0.4. *Gilt*

$$\begin{aligned} &x \cdot y_1 = y_1 \cdot x = 1 \quad \textit{und} \\ &x \cdot y_2 = y_2 \cdot x = 1, \end{aligned}$$

so ist $y_1 = y_2$.

Beweis. Wir benötigen nur

$$\begin{aligned} &(1)\ x \cdot y_1 = 1 \ \text{ und} \\ &(2)\ y_2 \cdot x = 1, \end{aligned}$$

denn

$$y_1 \overset{(G2)}{=} 1 \cdot y_1 \overset{(2)}{=} (y_2 x) y_1 \overset{(G1)}{=} y_2 (x y_1) \overset{(1)}{=} y_2 \cdot 1 \overset{(G2)}{=} y_2$$

□

Lemma 5.0.5.

$$1^{-1} = 1 \tag{1}$$
$$(x^{-1})^{-1} = x \tag{2}$$
$$(xy)^{-1} = y^{-1}x^{-1} \tag{3}$$

Beispiel 5.0.6. 1) $G = \mathbb{Z}$ mit der Addition (abelsch).

2) $G = \mathbb{K}$ mit der Addition (abelsch)

3) $G = \mathbb{K}^* := \mathbb{K} \setminus \{0\}$ mit der Multiplikation (abelsch)

4) $G = V$ Vektorraum mit Addition (abelsch)

5) $G = \mathbb{Z}/n$ mit der Addition (abelsch)

6) *Allgemeine lineare Gruppe* $G = \mathrm{GL}_n(\mathbb{K})$ Gruppe der invertierbaren $n \times n$- Matrizen über $\mathbb{K}$ (nicht-abelsch für $n \geq 2$)

$$A \in \underbrace{\mathrm{GL}_n(\mathbb{C})\, \mathrm{GL}_n(\mathbb{C}) \supseteq \mathrm{GL}_n(\mathbb{R}) \supseteq \mathrm{GL}_n(\mathbb{Q})}_{A^{-1} \text{ hat wieder Einträge im Unterkörper } \mathbb{K}' \subseteq \mathbb{K}} \supseteq \underbrace{\mathrm{GL}_n(\mathbb{Z})}_{\text{Warum ist } A^{-1} \text{ wieder ganzzahlig?}}$$

7) $G = \mathrm{GL}(V) = \mathrm{Aut}_{\mathbb{K}}(V) = \{f : V \to V \mid f \text{linear und bijektiv}\}$ *Automorphismen* von V (nicht abelsch für $\dim V > 1$).

8) *Diedergruppen* $G = \{\text{Drehungen und Spiegelungen eines regelmäßigen n-Ecks}\} = \mathrm{D}_{2n}$

9) Drehungen (ohne Spiegelungen) der 5 platonischen Körper (Tetraeder, Würfel, Oktaeder, Dodekaeder, Ikosaeder)
Die *Ikosaedergruppe* hat 60 Elemente:

12 Ecken
30 Kanten
20 Flächen (Dreiecke)
Euler-Zahl $\chi = 12 - 30 + 20 = 2$

- 6 Geraden durch antipodische Ecken: Drehungen um $72°, 144°, 216°, 288° \rightsquigarrow 6 \cdot 4 = 24$
- 10 Geraden durch antipodische Flächenmittelpunkte: Drehungen um $120°, 240° \rightsquigarrow 10 \times 2 = 20$
- 15 Geraden durch antipodische Kantenmittelpunkte: Drehungen um $180° \rightsquigarrow 15 \cdot 1 = 15$
- identische Abbildung $\rightsquigarrow 1$

5.1 Symmetrische Gruppen

Definition 5.1.1 (Symmetrische Guppe)**.** Für eine beliebige Menge X nennt man

$$\mathrm{Sym}(X) = \{f : X \to X \text{ bijektiv}\}$$

die *symmetrische Gruppe von* X . Für $X = \{1, \ldots, n\}$ schreibt man $\mathrm{S}_n = \mathfrak{S}_n = \mathrm{Sym}(\{1, \ldots, n\})$. Es gilt $|\mathfrak{S}_n| = n!$.
Spezielle Elemente:

- Die *Transposition* $(ij) \in \mathfrak{S}_n$ vertauschen nur zwei Elemente $i \neq j, 1 \leq i, j \leq n$ und lässt alle anderen fix.
- $(ij)^{-1} = (ij) = (ji)$.
- *Zykel*: $i_1, \ldots, i_k$ seien paarweise verschiedene Elemente in $\{1, \ldots, n\}$. Wir definieren $(i_1 i_2 \ldots i_k) \in \mathfrak{S}_n$ durch $i_1 \mapsto i_2 \mapsto i_3 \mapsto \ldots \mapsto i_k \mapsto i_1$ und $i \mapsto i$ für alle anderen Elemente.
- $k = 1$: $(i_1) = 1$.
- $(i_1 i_2 \ldots i_k)^{-1} = (i_k i_{k-1} \ldots i_2, i_1)$.
- $(i_1 i_2 \ldots i_k) = (i_2 i_3 \ldots i_k i_1) = (i_3 i_4 \ldots i_k i_1 i_2) = \ldots$ (Jeder Zykel der Länge k hat k Notationen.)

- $\mathfrak{S}_n$ ist nicht abelsch für $n \geq 3$:

$$(12)(23) = (123)$$
$$(23)(12) = (132)$$

Lemma 5.1.2. *Jedes $\pi \in \mathfrak{S}_n$ lässt sich eindutig als Produkt von disjunkten Zykeln schreiben:*

$$\pi = (a_1 \dots a_k)(b_1 \dots b_l)(c_1 \dots c_m) \cdots (x_1 \dots x_r)$$

Dabei ist die Reihenfolge der r Zykel nicht bestimmt, weil disjunkte Zykel vertauschen; die Reihenfolge der Elemente in einem Zykel ist bis auf zyklische Vertauschung eindeutig bestimmt.

Beweis. Wir schreiben $X = \{1, \dots, n\}$.
Wir beginnen mit dem Element 1 und verfolgen seine Bahn

$$1 = a_1 \overset{\pi}{\mapsto} a_2 \overset{\pi}{\mapsto} \dots \overset{\pi}{\mapsto} a_{k+1} = 1;$$

weil wir nur enndlich-viele Möglichkeiten haben, muss für ein $k \geq 1$ wieder $1 = a_{k+1}$ erscheinen; k sei minimal mit dieser Eigenschaft. Ist $k = 1$, so ist 1 ein Fixpunkt von π und wir brauchen den Zykel (a_1) gar nicht zu notieren; ansonsten haben wir einen ersten Zykel $(a_1 \dots a_k)$.
Wir setzen $\mathfrak{A} = \{a_1, \dots, a_k\}$ und betrachten die Menge $X \setminus \mathfrak{A}$. Ist diese Menge leer, sind wir fertig. Sonst wählen wir das kleinste Element b_1 und verfolgen seine Bahn

$$b_1 \overset{\pi}{\mapsto} b_2 \overset{\pi}{\mapsto} \dots \overset{\pi}{\mapsto} b_{l+1} = b_1.$$

Diese kann die Bahn von a_1 nicht kreuzen, liegt also ganz in $X \setminus \mathfrak{A}$. Ist $l = 1$, so brauchen wir den Zykel (b_1) nicht zu notieren; sonst haben wir den Zykel $(b_1 \dots b_n)$. Wir schreiben $\mathfrak{B} = \{b_1, \dots, b_l\}$.
Wir machen so weiter, als nächstes mit $X \setminus (\mathfrak{A} \cup \mathfrak{B})$, bis ganz X ausgeschöpft ist. □

Definition 5.1.3. Wir definieren den *Zykeltyp* von $\pi \in \mathfrak{S}_n$ als

$$\mathrm{Typ}(\pi) = (t_1, \dots, t_n),$$

wobei t_i die Anzahl der Zykel der Länge i und t_1 die Anzahl der Fixpunkte von π ist. Wir haben dann

$$n = 1t_1 + 2t_2 + \dots nt_n.$$

Beispiel 5.1.4. 1) $\pi = (ij)$ Transposition hat $\mathrm{Typ}(\pi) = (n-2, 1, 0, \dots, 0)$.
2) $\pi = 1$ hat $\mathrm{Typ}(1) = (n, 0, 0, \dots, 0)$.
3) $\pi = (1375462)(8)(910)$ hat $\mathrm{Typ}(\pi) = (1, 1, 0, 0, 0, 0, 1, 0)$.

Lemma 5.1.5. *Jeder Zykel ist ein Produkt von Transpositionen:*

$$(i_1 \dots i_k) = (i_1 i_2)(i_2 i_3) \dots (i_{k-2} i_{k-1})(i_{k-1} i_k)$$

Korollar 5.1.6. *Jede Permutation ist ein Produkt von Transpositionen.*

Lemma 5.1.7. $\sigma(i_1 \dots i_k)\sigma^{-1} = (\sigma(i_1)\sigma(i_2) \dots \sigma(i_k)$.

5.2 Untergruppen

Definition 5.2.1. Eine Teilmenge $H \subseteq G$ einer Gruppe G heißt *Untergruppe*, wenn sie folgende Eigenschaften hat:

(U1) $1 \in H$.
(U2) $g \in H \Leftrightarrow g^{-1} \in H$.
(U3) $g_1, g_2 \in H \Rightarrow g_1 g_2 \in H$.

Wir notieren Untergruppen mit $H \leq G$.

Beispiel 5.2.2.

1) $H = 1 =: 1 \leq G$ ist die *triviale* Untergruppe.

2) $n\mathbb{Z} = \{a \in \mathbb{Z} \mid a \text{ ist durch } n \text{ teilbar}\} \leq \mathbb{Z}$

3) $\mathbb{S}^0 = \{\pm 1\} = \{x \in \mathbb{R}^\times \mid |x| = 1\} \leq \mathbb{R}^\times$

4) $\mathbb{S}^1 = \{z \in \mathbb{C}^\times \mid |z| = 1\} \leq \mathbb{C}^\times$.

5) $\mathbb{S}^3 = \{z \in \mathbb{H} \cong \mathbb{R}^4 \mid |z| = 1\} \cong \mathbb{H}^\times$, wobei $\mathbb{H}$ die Gruppe der Quaternionen bezeichnet.

6) $\mu_n = \{z \in \mathbb{C}^\times \mid z^n = 1\} \leq \mathbb{C}^\times$ Gruppe der n-ten Einheitswurzeln

7) $H = \{1, (ij)\} \leq \mathfrak{S}_n$ für $n \geq 2$; $H' = \{\sigma \in \mathfrak{S}_n \mid \sigma(k) = k\} \leq \mathfrak{S}_n$

8) $\mathrm{GL}_n(\mathbb{Z}) \leq \mathrm{GL}_n(\mathbb{Q}) \leq \mathrm{GL}_n(\mathbb{R}) \leq \mathrm{GL}_n(\mathbb{C})$

9) $\mathrm{SL}_2(\mathbb{K}) = \left\{ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{K}) \mid ad - bc = 1 \right\} \leq \mathrm{GL}_2(\mathbb{K})$ für $\mathbb{K}$ Körper oder $\mathbb{K} = \mathbb{Z}$

10) $H = \{A \in GL_n(\mathbb{K}) \mid A^\top = A^{-1}\}$ ist eine Untergruppe, denn $\mathbb{1}^\top = \mathbb{1}^\top$ und $(AB)^\top = B^\top A^\top = B^{-1}A^{-1} = (AB)^{-1}$.

11) Sei V ein Vektorraum und sei $0 \neq v \in V$. Dann sind

$$\begin{aligned} H &= \{f \in \mathrm{GL}(V) \mid f(v) = v\} && \text{und} \\ H' &= \{f \in \mathrm{GL}(V) \mid \exists \lambda \in \mathbb{K}^\times : f(u) = \lambda u\} \end{aligned}$$

Untergruppen von $\mathrm{GL}(V)$. Ein $f \in H$ lässt das Erzeugnis von v ($\langle v \rangle$) *punktweise* fest. Ein $f \in H'$ lässt hingegen nur den von v erzeugten Unterraum invariant, denn $f(\langle v \rangle) = \langle v \rangle \Leftrightarrow \exists \lambda \neq 0 : f(v) = \lambda v$, wobei letzteres eine *Eigenwertgleichung* ist. Wir haben also $H \leq H'$.
Allgemeiner betrachten wir für einen Untervektorraum $U \subseteq V$ die Untergruppen

$$\begin{aligned} H &= \{f \in GL(V) \mid \forall u \in U : f(u) = u\} && \text{und} \\ H' &= \{f \in GL(V) \mid f(U) = U\} \end{aligned}$$

von $\mathrm{GL}(V)$, für die gilt $H \leq H'$.

12) Die Menge der Diagonalmatrizen $\mathcal{D} = \{D[\lambda_1, ..., \lambda_n] \mid \lambda_1, ..., \lambda_n \neq 0\}$ sind mit der Matrizenmultiplikation eine Gruppe. Die Menge aller *zentralen Matrizen* $\mathcal{Z} = \{D[\lambda, ..., \lambda] | \lambda \neq 0\}$ eine Untergruppe: $\mathcal{Z} \leq \mathcal{D}$. Die zentralen Matrizen sind besonders wichtig, da sie das *Zentrum* der Gruppe $GL_n(\mathbb{K})$ sind, das heißt, sie vertauschen mit allen Matrizen.

13) Die Menge der Permutationsmatrizen $\mathcal{P} = \{P_\sigma | \sigma \in \mathfrak{S}_n\}$ ist eine Untergruppe der $GL_n(\mathbb{K})$: $P_1 = \mathbb{1}$, $P_{\alpha\beta} = P_\alpha P_\beta$.

14) Die Menge der *oberen Dreiecksmatrizen* $\mathcal{T}^+$, also der Matrizen der Form

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

ist abgeschlossen bezüglich der Matrixmultiplikation und daher eine Untergruppe der GL_n. Die entsprechende Aussage gilt auch für die Menge der *unteren Dreiecksmatrizen* $\mathcal{T}^-$.

15) Für den Tetraeder $\mathcal{T}$ ist die Gruppe der Rotationen eine Untergruppe der Symmetrien (also Rotationen *und* Spiegelungen):

$$\mathrm{Rot}(\mathcal{T}) = \mathcal{T} \leq \mathcal{T}' = \mathrm{Sym}(\mathcal{T}')$$

Die analoge Aussage gilt für alle platonischen Körper.

Definition 5.2.3. (i) Die Kardinalität einer Gruppe G heißt ihre *Ordnung* $|G|$.

(ii) Sei $g \in G$. Die kleinste natürliche Zahl s mit $g^s = 1$ nennt man die *Ordnung* von g, falls so ein s existiert; falls nicht, so heißt g von *unendlicher Ordnung*.
Notation: $\mathrm{ord}(g) = 1, 2, ..., \infty$.

Bemerkung 5.2.4. • Die Kardinalität von G kann überabzählbar sein (z.B. $G = \mathbb{S}^1$).

Beispiel 5.2.5.

1. In der Übungsaufgabe 56 haben wir gefunden:

$$A = \begin{pmatrix} \cos\frac{2\pi}{s} & -\sin\frac{2\pi}{s} \\ \sin\frac{2\pi}{s} & \cos\frac{2\pi}{s} \end{pmatrix} \qquad \mathrm{ord}(A) = s$$

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \qquad \mathrm{ord}(A) = \infty \text{ für } \frac{2\pi}{\alpha} \notin \mathbb{Q}$$

2. Für ein festes $g \in G$ betrachten wir die Untergruppe aus allen ganzzahligen Potenzen von g:

$$H := \{1, g^{\pm 1}, g^{\pm 2}, ..., g^{\pm n}, ...\} \leq G$$

   Es gilt $|H| = \mathrm{ord}(g)$.

3. Wir betrachten die symmetrische Gruppe $\mathfrak{S}_n$. Die Ordnung eines k-Zykels $g = (i_1 i_2 ... i_k)$ ist $ord(g) = k$. Wenn g das Produkt disjunkter k- und l-Zykel ist, gilt:

$$g = (i_1 ... i_k)(j_1 ... j_l) \Rightarrow \mathrm{ord}(g) = \mathrm{kgV}(k, l).$$

   Produkte nicht-disjunkter Zykel muss man erst als Produkte disjunkter Zykel ausdrücken.

Definition 5.2.6. Für jede Teilmenge $\mathcal{E} \subseteq G$ einer Gruppe sei $H := \langle \mathcal{E} \rangle$ die Menge aller endlichen Produkte von Elementen und ihren Inversen in $\mathcal{E}$

$$g = e_1^{l_1} e_2^{l_2} \cdots e_i^{l_i} \cdots e_n^{l_n} \qquad l_i \in \mathbb{Z},$$

wobei die e_i *nicht* verschieden sein müssen! (Nur wenn G abelsch ist, kann man verlangen, dass die e_i verschieden sind.) H heißt die von $\mathcal{E}$ *erzeugte* Untergruppe. Ist $H = \langle \mathcal{E} \rangle = G$, so heißt $\mathcal{E}$ ein *Erzeugendensystem* für G.

Beispiel 5.2.7.

1. $G = \mathbb{Z}_{/n} = \{\underline{0}, \underline{1}, ..., \underline{n-1}\}, \mathcal{E} = \{\underline{1}\}$, wir können jede Restklasse als Summe $\underline{1} + \underline{1} + ...$ schreiben. Ein anderes Erzeugendensystem wäre zum Beispiel $\mathcal{E} = \{\underline{2}, \underline{3}\}$, da die Differenz ± 1 ist.
2. $G = \mathfrak{S}_n, \mathcal{E} = \{(ij) \mid 1 \leq i < j \leq n\}$, denn jede Permutation lässt sich als Produkt von Transpositionen schreiben. Es reicht sogar $\{(i\,(i+1)) \mid 1 \leq i < n\}$, da $\sigma\,(i\,(i+1))\,\sigma^{-1} = (\sigma(i)\sigma(i+1)) = (kl)$ für ein σ mit $\sigma(i) = k, \sigma(i+1) = l$.
3. $G = GL_n(\mathbb{K})$: Ein Erzeugendensystem der allgemeinen linearen Gruppe ist die Vereinigung von Elementarmatrizen, Streckungsmatrizen und Permutationsmatrizen.

Bemerkung 5.2.8. Beim linearen Erzeugnis $\mathrm{span}(\mathcal{E})$ in einem Vektorraum wird außer der Addition und dem Negativen auch die Skalierung benutzt.
Sei $\mathfrak{B} = \{b_1, \ldots, b_n\} \subseteq V$ Basis des $\mathbb{R}$-Vektorraums V. Dann ist

$$\langle \mathfrak{B} \rangle = \{v = k_1 b_1 + \ldots k_n b_n \mid k_1, \ldots k_n \in \mathbb{Z}\}$$

nur eine additive Untergruppe von V, aber *kein* Untervektorraum; man nennt $\langle \mathfrak{B} \rangle$ auch ein *Gitter*.

Nun noch ein paar wichtige Untergruppen:

Definition 5.2.9. Das *Zentrum* $\mathrm{Zen}(G)$ einer Gruppe G ist die Untergruppe

$$\mathrm{Zen}(G) := \{z \in G \mid \forall g \in G : zg = gz\},$$

also die Elemente, die mit allen kommutieren.
Die *Kommutatoruntergruppe* $\mathrm{Kom}(G) = [G, G]$ ist die von allen *Kommutatoren* $[x, y] := xyx^{-1}y^{-1}$ ($x, y \in G$ beliebig) erzeugte Untergruppe. Beachte:

- $[x, y]^{-1} = (xyx^{-1}y^{-1})^{-1} = yxy^{-1}x^{-1} = [y, x]$
- $[x, y][a, b]$ ist in der Regel *kein* Kommutator (deshalb nimmt man das Erzeugnis).
- Abgeschlossen bezüglich Konjugieren:

$$\gamma[x, y]\gamma^{-1} = \gamma xyx^{-1}y^{-1}\gamma^{-1} = (\gamma x\gamma^{-1})(\gamma y\gamma^{-1})(\gamma x^{-1}\gamma^{-1})(\gamma y^{-1}\gamma^{-1}) = [\gamma x\gamma^{-1}, \gamma y\gamma^{-1}]$$

Beispiel 5.2.10.

1. G abelsch: $\mathrm{Zen}(G) = G$, denn die Gruppe ist kommutativ; $\mathrm{Kom}(G) = 1$, da $xy = yx \Leftrightarrow [x, y] = 1$.
2. Für $G = GL_n(\mathbb{K})$ ist $\mathrm{Zen}(G) = \{D[\lambda, ..., \lambda] \mid 0 \neq \lambda \in \mathbb{K}\} \cong \mathbb{K}^\times$. Dies Matrizen nennt man *zentrale Matrizen*, vergleiche Übungsaufgabe 63.
3. $G = \mathfrak{S}_n, n \geq 3 : \mathrm{Zen}(G) = 1$
4. Für drei verschiedene Indizes $1 \leq i, j, k \leq n$ gilt
$$[E^{ii}[\lambda], E^{jk}[\mu]] = E^{jk}[\lambda\mu]$$
Für $n \geq 3$ erzeugen die Kommutatoren also alles: $\mathrm{Kom}(\mathrm{GL}_n(\mathbb{K})) = \mathrm{GL}_n(\mathbb{K})$

5.3 Homomorphismen

Definition 5.3.1. Eine Funktion $f : G \to G'$ zwischen zwei Gruppen G, G' heißt *Homomorphismus*, falls gilt:

1. $f(1) = 1$,
2. $f(xy) = f(x)f(y)$ für alle $x, y \in G$.

Bemerkung 5.3.2.

- Es gilt $f(x^{-1}) = f(x)^{-1}$
- $\mathrm{ord}(f(x))$ teilt $\mathrm{ord}(x)$, falls $\mathrm{ord}(x)$ endlich; insbesondere hat $f(x)$ endliche Ordnung, wenn x endliche Ordnung hat.
- Isomorphismus: f bijektiv $\xrightarrow{\cong}$
Epimorphismus: f surjektiv $\twoheadrightarrow$
Monomorphismus: f injektiv $\rightarrowtail$
Endomorphismus: $G = G'$
Automorphismus: $G = G'$ und f Isomorphismus

Beispiel 5.3.3.

1. Für festes n: $G = \mathbb{Z} \to G' = \mathbb{Z},\ f(x) = nx$
$n \neq 0$: mono, $n = \pm 1$: iso
2. $\rho_n : \mathbb{Z} \to \mathbb{Z}/n, \rho_n(x) = \underline{x}$ Restklasse modulo n, epi
3. $\exp : (\mathbb{R}, +) \rightleftarrows (\mathbb{R}_{>0}, \cdot) : \ln$
4. $\exp : (\mathbb{R}, +) \to (\mathbb{S}^1, \cdot),\ t \mapsto e^{2\pi i t}$
5. $f : V \to V' f$ linear, insb. gilt: $f(x + y) = f(x) + f(y)$
6. Die Homomorphismen
$$G = \mathrm{GL}(V) \underset{L_{\mathcal{A}\mathcal{A}}}{\overset{M_{\mathcal{A}\mathcal{A}}}{\rightleftarrows}} G' = \mathrm{GL}_n(\mathbb{K}) \qquad (\dim_{\mathbb{K}}(V) = n, \mathcal{A} \subset V \text{ Basis}),$$
sind beide Isomorphismen.
7. $\varepsilon : \mathbb{Z} \to \mathbb{K}, \varepsilon(n) = \varepsilon_n = (1 + ... + 1)$ n-mal
$\varepsilon(0) = 0, \varepsilon(n + m) = \varepsilon(n) + \varepsilon(m)$
8. $\mathbb{Z}/n \to \mu_n,\ k \mapsto \zeta_n^k,\ \zeta_n = exp(\frac{2\pi i}{n})$
9. $(\mathbb{R}, +) \xrightarrow{\exp} (\mathbb{R}_{>0}, \cdot)$ $\exp(x + y) = \exp(x)\exp(y)$
Umkehrfunktion: $\log(xy) = \log(x) + \log(y)$
10. $P : \mathfrak{S}_n \to \mathbb{P}_n =$ Gruppe der Permutationsmatrizen, $\sigma \mapsto P_\sigma$

11. $D : (\mathbb{K}^\times)^n \to \mathbb{D}$ = Gruppe der Diagonalmatrizen, $(\lambda_1, ..., \lambda_n) \mapsto D(\lambda_1, ..., \lambda_n)$
12. Für $g \in G$ haben wir einen Homomorphismus $f_g : \mathbb{Z} \to G,\ k \mapsto g^k$. Falls $\mathrm{ord}(g) = n$, so ist auch $\mathbb{Z}/n \to G, \bar{k} \mapsto g^k$ ein Homomorphismus.

Definition 5.3.4 (Signum)**.** Auf den symmetrischen Gruppen definieren wir eine wichtige Funktion

$$\mathrm{sign} : \mathcal{S}_n \to \{\pm 1\}, \mathrm{sign}(\sigma) = \prod_{1 \le i < j \le n} \frac{\sigma(j) - \sigma(i)}{j - i}$$

1. Warum ist $\mathrm{sign}(\sigma) = \pm 1$?
2. Ein Paar (i, j) mit $1 \le i < j \le n$ und $\sigma(i) > \sigma(j)$ nennt man einen **Fehlstand** von σ.
3. Hat σ genau k Fehlstände ($0 \le k \le \frac{n(n-1)}{2}$), so gilt

   $$\prod_{1 \le i < q \le n} (\sigma(j) - \sigma(i)) = (-1)^k \prod_{1 \le i < j \le n} |\sigma(j) - \sigma(i)| = (-1)^k \prod_{1 \le i < j \le} |j - i| = (-1)^k \prod_{1 \le i < j \le} (j - i),$$

   weil in den beiden mittleren Produkten die gleichen Faktoren in eventuell anderer Reihenfolge vorkommen. Also ist $\mathrm{sign}(\sigma) = (-1)^k$ bestimmt durch die Anzahl der Fehlstände.

Beispiel 5.3.5.

1. $\sigma = (ij)$ Transposition $\mathrm{sign}(\sigma) = -1$
2. $\sigma = (123)$; $\mathrm{sign}(\sigma) = \frac{3-2}{2-1} \frac{1-2}{3-1} \frac{1-3}{3-2} = +1$

Lemma 5.3.6. sign *ist ein Homomorphismus:*

$$\mathrm{sign}(1) = 1,$$
$$\mathrm{sign}(\pi \circ \sigma) = \mathrm{sign}(\pi) \cdot \mathrm{sign}(\sigma)$$

Beweis.

$$\begin{aligned}
\mathrm{sign}(\pi \circ \sigma) &= \prod_{i<j} \frac{\pi(\sigma(j)) - \pi(\sigma(i))}{j - i} \\
&= \prod_{i<j} \frac{\pi(\sigma(j)) - \pi(\sigma(i))}{\sigma(j) - \sigma(i)} \cdot \frac{\sigma(j) - \sigma(i)}{j - i} \\
&= \prod_{k<l} \frac{\pi(l) - \pi(k)}{l - k} \cdot \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i} \\
&= \mathrm{sign}(\pi) \cdot \mathrm{sign}(\sigma)
\end{aligned}$$

□

Beispiel 5.3.7.

1. $\sigma = (i_1\ i_2\ ...\ _k)$ k-Zykel
   $= (i_1\ i_2)(i_2\ i_3)...(i_{k-1} i_k)$ $k - 1$ Faktoren
   $\mathrm{sign}(\sigma) = (-1)^{k-1}$
2. $\mathrm{sign}(\sigma^{-1}) = \mathrm{sign}(\sigma)$
3. $\mathrm{Typ}(\sigma) = (t_1, t_2, ..., t_n)$:
   $\mathrm{sign}(\sigma) = (-1)^{t_1 \cdot 0} \cdot (-1)^{t_2 \cdot 1} ... (-1)^{t_n \cdot n} = (-1)^{n-(t_1+t_2+...+t_n)}$

Definition 5.3.8. Für einen Homomorphismus $f : G \to G'$ definieren wir

1. $\ker(f) = \bar{f}(1) = \{x \in G \mid f(x) = 1\} \subseteq G$, genannt der *Kern* von f.
2. $\mathrm{im}(f) = f(G) = \{y \in G' \mid \exists_{x \in G} f(x) = y\} \subseteq G'$ genannt das *Bild* von f.

Lemma 5.3.9. *1)* $\ker(f)$ *ist eine Untergruppe; f ist genau dann ein Monomorphismus, wenn* $\ker(f) = 1$ *ist.*

2) $\operatorname{im}(f)$ *ist eine Untergruppe; f ist genau dann ein Epimorphismus, wenn* $\operatorname{im}(G) = G'$ *ist.*

Beweis. 1) $1 \in \ker(f)$, wegen $f(1) = 1$.
$g_1, g_2 \in \ker(f) \Rightarrow f(g_1) = 1, f(g_2) = 1 \;\Rightarrow f(g_1g_2) = f(g_1)f(g_2) = 1 \cdot 1 = 1$
Ist f ein Monomorphismus, so kann nur für $x = 1$ die Bedingung $f(x) = 1$ gelten. Angenommen $f(g_1) = f(g_2)$; dann ist $f(g_1g_2^{-1}) = f(g_1)f(g_2^{-1}) = f(g_1)f(g_2)^{-1} = 1$; also folgt aus $\ker(f) = 1$ nun $g_1g_2^{-1} = 1$, also $g_1 = g_2$.
2) $1 = f(1)$, also $1 \in im(f)$.
Ist $y_1 = f(x_1)$ und $y_2 = f(x_2)$, so ist $f(x_1x_2) = y_1y_2$. □

Beispiel 5.3.10.

1. $f : \mathbb{Z} \twoheadrightarrow \mathbb{Z}/n,\; x \mapsto \bar{x}$ Restklasse mod n
   $\ker(f) = n\mathbb{Z} \subseteq \mathbb{Z}$

2. $\varepsilon\mathbb{Z} \to \mathbb{K},\; \varepsilon(n) = n1$
   $\ker(\varepsilon) = 0 \Leftrightarrow \operatorname{char}(\mathbb{K}) = 0$
   $\ker(\varepsilon) = p\mathbb{Z} \Leftrightarrow \operatorname{char}(\mathbb{K}) = p$ (Primzahl)

3. $L_g : \mathbb{Z} \to G,\; f(n) = g^n \quad (g \in G$ fest)
   $\ker(L_g) = 0 \Leftrightarrow \operatorname{ord}(g) = \infty$
   $\ker(L_g) = n\mathbb{Z} \Leftrightarrow \operatorname{ord}(g) = n$

4. $L_{a,b} : (\mathbb{R}, +) \to (\mathbb{C}^\times, \cdot) \quad a, b \in \mathbb{R}, a, b > 0$
   $L_{a,b}(t) = e^{(a+ib)t}$
   $\operatorname{im}(L_{a,b})$ = Spirale (logarithmische oder Bernoulli'sche)

5. X Menge, G Gruppe: $\Gamma = \operatorname{Funkt}(X, G)$ ist eine Gruppe:
   $(\gamma_1\gamma_2)(x) := \gamma_1(x)\gamma_2(x),\; 1(x) := 1,\; \gamma^{-1}(x) := \gamma(x)^{-1}$

6. Ist X eine Gruppe, so ist $\operatorname{Hom}(X, G)$ i.A. keine Untergruppe:

$$(\gamma_1 \cdot \gamma_2)(xy) = \gamma_1(xy)\gamma_2(xy) = \gamma_1(x)\gamma_1(y)\gamma_2(x)\gamma_1(y)$$
$$(\gamma_1 \cdot \gamma_2)(x)(\gamma_1 \cdot \gamma_2)(y) = \gamma_1(x)\gamma_2(x)\gamma_1(y)\gamma_2(y)$$

Lemma 5.3.11. *Für einen Homomorphismus $\phi : G \to G'$ gilt:*

1. $\phi(\operatorname{Kom}(G)) \subset \operatorname{Kom}(G')$

2. $\operatorname{Kom}(G) \subset \ker(\phi)$*, falls G' abelsch ist.*

Beweis. 1): Ist $g = [x, y] = xyx^{-1}y^{-1}$ ein Kommutator, so ist $\phi(g) = \phi(xyx^{-1}y^{-1}) = \phi(x)\phi(y)\phi(x)^{-1}\phi(y)^{-1} = [\phi(x), \phi(y)]$ wieder ein Kommutator.
$[x, y]^{-1} = xyx^{-1}y^{-1} = yxy^{-1}x^{-1} = [y, x]$ auch Und ist g ein Produkt von Kommutatoren (und Inversen von Kommutatoren), so auch $\phi(g)$.

2): Ist G' abelsch, so ist $\operatorname{Kom}(G') = 1$, also $\phi(\operatorname{Kom}(G)) = 1$, also $\operatorname{Kom}(G) \subset \ker(\phi)$:
Oder auf Elementen:
$\phi[x, y] = \phi(x)\phi(y)\phi(x)^{-1}\phi(y)^{-1} = \phi(x)\phi(x)^{-1}\phi(y)\phi(y)^{-1} = 1$. □

5.3.1 Vorwärtsschicken/ Rückwärtsschicken von Gruppenstrukturen

Ist $(G, \cdot)$ eine Gruppe, $X = G'$ eine bloße Menge und $\phi : G \to X$ eine Bijektion, so kann man X derart zu einer Gruppe $(G', *)$ machen, dass ϕ ein Gruppenhomomorphismus ist.

$$\phi : (G, \cdot) \to X = G' \text{ Bijektion}$$

vorwärtsgeschobene Gruppenstruktur auf G':

$x * y := \phi(\phi^{-1}(x) \cdot \phi^{-1}(y))$
$1^* := \phi(1)$
$x^{-1*} := \phi(\phi^{-1}(x)^{-1})$

1) $\phi^{-1}(x * y) = \phi^{-1}(x) \cdot \phi^{-1}(y)$
2) $\phi^{-1}(1^*) = 1$
3) $\phi^{-1}(x^{-1*}) = \phi^{-1}(x)^{-1}$
$\phi^{-1} : (G', *) \xrightarrow{\sim} (G, \cdot)$

1') $\phi(ab) = \phi(a) * \phi(b)$
2') $\phi(1) = 1^*$
3') $\phi(a^{-1}) = \phi(a)^{-1}$
$\phi : (G, \cdot) \xrightarrow{\sim} (G, *)$

Gruppenaxiome folgen aus (1-3) oder (1'-3'); man kann sie auch direkt nachprüfen:

$$\begin{aligned} x * (y * z) &= x * \phi(\phi^{-1}(y) \cdot \phi^{-1}(z)) \\ &= \phi(\phi^{-1}(x) \cdot \phi^{-1}\phi(\phi^{-1}(y) \cdot \phi^{-1}(z))) \\ &= \phi(\phi^{-1}(x) \cdot (\phi^{-1}y \cdot \phi^{-1}(z))) \\ &= \phi((\phi^{-1}(x) \cdot \phi^{-1}(y)) \cdot \phi^{-1}(z)) \\ &= \phi(\phi^{-1}(\phi(\phi^{-1} \cdot \phi^{-1}(y))) \cdot \phi^{-1}(z)) \\ &= \phi(\phi^{-1}(x * y) \cdot \phi^{-1}z) \\ &= (x * y * z) \end{aligned}$$

Genau analog: Zurückziehen

$X = G' \xrightarrow{\psi} (G, \cdot)$ Bijektion

$x * y := \psi^{-1}(\psi(x) \cdot \psi(y))$
$1^* := \psi^{-1}(1)$
$x^{-1*} := \psi^{-1}(\psi(x)^{-1})$

Beispiel 5.3.12.

1. **Tangens**: $(\mathbb{R}, +) \xrightarrow{\tan} (\mathbb{R}, *)$
   (nicht für $n\frac{\pi}{2}$ definiert!)

$$\begin{aligned} \tan(a + b) &= \frac{\tan(a) + \tan(b)}{1 - \tan(a)\tan(b)} \\ x * y :&= \frac{x + y}{1 - xy} \end{aligned}$$

2. $G = \mathbb{R}$ mit Addition
   $G' = \mathbb{R}_{>0}$ mit Multiplikation

- Das geometrische Mittel ist das mit $\psi = \ln$ zurückgezogene arithmetisches Mittel:
$$\exp(\frac{\ln(x)+\ln(y)}{2}) = \exp(\frac{\ln(xy)}{2}) = \exp(\ln\sqrt{xy}) = \sqrt{xy}$$
- Das harmonische Mittel ist das mit $\psi = \frac{1}{t}$ zurückgezogene arithmetische Mittel
$$G' = \mathbb{R}_{>0}, (\frac{x^{-1}+y^{-1}}{2})^{-1} = \frac{2}{\frac{1}{x}+\frac{1}{y}} = \frac{2}{\frac{y+x}{xy}} = \frac{2xy}{x+y}$$

3. $G = (\mathbb{R}, +) \xrightarrow{\psi=(\ )^3} G' = \mathbb{R}$
$x * y = (\sqrt[3]{x} + \sqrt[3]{y})^3 = x + 3\sqrt[3]{x^2y} + 3\sqrt[3]{xy^2} + y$
$x^{-1*} = -x$
$0^* = 0$ „Deformation der Addition"

5.3.2 Einige wichtige Isomorphismen

1. $\mathbb{Z}/n \xrightarrow{\cong} \mu_n \xrightarrow{\cong} \mathrm{Rot}_n \subset \mathrm{D}_{2n}$ Diedergruppe
$\bar{k} \mapsto \zeta_n^k \mapsto \mathcal{A}_n = \begin{pmatrix} \cos\frac{2\pi}{n} & -\sin\frac{2\pi}{n} \\ \sin\frac{2\pi}{n} & \cos\frac{2\pi}{n} \end{pmatrix}$
2. Gruppe mit zwei Elementen:
$$\begin{array}{ccccc} \mathbb{Z}/2 & \cong & \mathbb{S}^0 & \cong & \mathfrak{S}_2 \\ 0 & \mapsto & 1 & \mapsto & \mathrm{id} \\ 1 & \mapsto & -1 & \mapsto & (12) \end{array}$$
3. Die folgenden Isomorphismen hängen von einer Basis $\mathcal{A}$ von V ab ($\dim V = n$):
$$\mathrm{GL}(V) \underset{L_{\mathcal{A}\mathcal{A}}}{\overset{M_{\mathcal{A}\mathcal{A}}}{\rightleftarrows}} \mathrm{GL}_n(\mathbb{K})$$
4. $\mathcal{S}_n \xrightarrow{\cong} P_n$ Permutationsmatrizen
$\sigma \mapsto P_\sigma$
5. $\mathrm{Zen}(\mathrm{GL}_n(V)) \cong \mathrm{Zen}(\mathrm{GL}_n(\mathbb{K})) \cong \mathbb{K}^\times,\ n \geq 3$
Diagonalmatrizen $D_n \to (\mathbb{K}^\times)^n$
$D[\lambda_1, ..., \lambda_n] \mapsto (\lambda_1, ..., \lambda_n)$
6. Diedergruppen: $D_2 \cong \mathcal{S}_2,\ D_4 \cong \mathbb{Z}/2 \times \mathbb{Z}/2,\ D_6 \simeq \mathcal{S}_3$
7. Platonische Gruppen:
   - Tetraeder: $\mathfrak{T} \cong \mathfrak{A}_4$
   - Würfel und Oktaeder: $\mathfrak{W} \cong \mathfrak{Oc} \cong \mathfrak{S}_4$
   - Dodekaeder und Ikosaeder: $\mathfrak{Do} \cong \mathfrak{J} \cong \mathfrak{A}_5$
8. **Automorphismengruppen**
   a) $\mathbb{Z} \xrightarrow{\mathrm{id}} \mathbb{Z},\ -\mathrm{id} : \mathbb{Z} \to \mathbb{Z},\ n \mapsto -n$
   $\mathrm{Aut}(\mathbb{Z}) = \{id, -id\} \simeq \mathbb{S}^0$
   b) $\mathbb{Z}/2 \to \mathbb{Z}/2$: nur die Identität
   $\bar{0} \mapsto \bar{0}$
   $\bar{1} \mapsto \bar{1}$
   Aut $(\mathbb{Z}/2) = \bar{1}$
   c) $\mathbb{Z}/3 \to \mathbb{Z}/3$: außer der Identität noch $\theta = -id$.
   $\theta : \bar{0} \mapsto \bar{0}$
   $\bar{1} \mapsto \bar{2}$
   $\bar{2} \mapsto \bar{1}$
   $\mathrm{Aut}(\mathbb{Z}/3) \simeq \mathbb{S}^0$

5.4 Produkte

Definition 5.4.1. Gegeben seien zwei Gruppen G_1 und G_2. Wir definieren eine neue Gruppe $G = G_1 \times G_2$ mit der Multiplikation

$$\begin{aligned}(x_1, x_2) \cdot (y_1, y_2) &:= (x_1 \cdot y_1, x_2 \cdot y_2)\\ (x_1, x_2)^{-1} &:= (x_1^{-1}, x_2^{-1})\\ 1 &:= (1, 1)\end{aligned}$$

Offenbar ist das wieder eine Gruppe und wird *direktes Produkt* genannt.

Beispiel 5.4.2.

1. $V = \mathbb{K}^n = \mathbb{K}^{n_1} \times \mathbb{K}^{n_2} \qquad n_1 + n_2 = n$
2. $(\mathbb{R}^\times, \cdot) \cong (\mathbb{S}^0, \cdot) \times (\mathbb{R}_{>0}, \cdot)$
   $x \mapsto (\frac{x}{|x|}, |x|)$
3. $\mathbb{C}^\times \cong \mathbb{S}^1 \times \mathbb{R}_{>0}$
   $z \mapsto (\frac{z}{|z|}, |z|)$
4. $V \supset U_1, U_2$ Untervektorräume,
   $V = U_1 + U_2,\ U_1 \cap U_2 = 0$
   $V \cong U_1 \times U_2 =: U_1 \oplus U_2$ interne direkte Summe

   $(u_1, u_2) \mapsto u_1 + u_2$
5. Sei V wie in 4.

$$\begin{tikzcd}
\mathrm{GL}(V) & \geq & G = \{f \in \mathrm{GL}(V) \mid f(U_1) \subseteq U_1, f(U_2) \subseteq U_2\} \arrow[d] & \ni & f \arrow[d, mapsto] \\
 & & \mathrm{GL}(U_1) \times \mathrm{GL}(U_2) & \ni & (f_1, f_2)
\end{tikzcd}$$

wobei $f_i = f|_{U_i}^{U_i} : U_i \to U_i$.

Eigenschaften

- $$x_1 \xmapsto{i_1} (x_1, 1); (1, x_2) \xleftarrow{i_2} x_2 \qquad \text{Inklusionen}$$

  $$\begin{tikzcd}
G_1 \arrow[r, bend left, "i_1"] & G_1 \times G_2 \arrow[l, bend left, "\pi_1"] \arrow[r, bend right, "\pi_2"'] & G_2 \arrow[l, bend right, "i_2"']
\end{tikzcd}$$

  $$x_1 \xleftarrow{\pi_1} (x_1, x_2) \xmapsto{\pi_2} x_2 \qquad \text{Projektionen}$$

- G_1, G_2 abelsch $\implies$ $G_1 \times G_2$ abelsch.
- Im abelschen Falle sagt man *direkte Summe* und schreibt auch $G = G_1 \oplus G_2$.
  Bsp.: $\mathbb{Z}/6 \cong \mathbb{Z}/2 \oplus \mathbb{Z}/3$

Proposition 5.4.3 (Universelle Eigenschaft)**.** *Zu je zwei Homomorphismen $f_1 : \Gamma \to G_1, f_2 : \Gamma \to G_2$ gibt es genau einen Homomorphismus $f = (f_1, f_2) : \Gamma \to G_1 \times G_2$ mit $f_1 = \pi_1 \circ f, f_2 = \pi_2 \circ f$:*

$$\begin{tikzcd}
\Gamma \arrow[dr, dotted, "f"] \arrow[drr, bend left, "f_1"] \arrow[ddr, bend right, "f_2"'] & & \\
 & G \arrow[r, "\pi_1"'] \arrow[d, "\pi_2"] & G_1 \arrow[d] \\
 & G_2 \arrow[r] & 1
\end{tikzcd}$$

Wir erhalten eine Bijektion von Mengen

$$\begin{aligned}\mathrm{Hom}(\Gamma, G) &\overset{\cong}{\longleftrightarrow} \mathrm{Hom}(\Gamma, G_1) \times \mathrm{Hom}(\Gamma, G_2)\\ f &\longmapsto (\pi_1 \circ f, \pi_2 \circ f)\\ (f : \gamma \mapsto (f_1(\gamma), f_2(\gamma))) &\longleftarrow (f_1, f_2)\end{aligned}$$

5.5 Direkte Summen von Vektorräumen

5.5.1 Externe direkte Summe

Definition 5.5.1. Sind V_1, V_2 zwei Vektorräume, so ist $V := V_1 \times V_2$ mit

Addition	$(v_1, v_2) + (v_1', v_2') := (v_1 + v_1', v_2 + v_2')$
Null	$0 := (0, 0)$
Negative	$-(v_1, v_2) = (-v_1, -v_2)$
Skalierung	$\lambda(v_1, v_2) = (\lambda v_1, \lambda v_2)$

wieder ein Vektorraum. Wir schreiben $V = V_1 \oplus V_2$.

- Es gibt Monomorphismen (Inklusionen)

$$\begin{array}{cccc} i_1: & V_1 & \to & V, \\ & v_1 & \mapsto & (v_1, 0) \end{array} \qquad \begin{array}{cccc} i_2: & V_2 & \to & V \\ & v_2 & \mapsto & (0, v_2) \end{array}$$

  und Epimorphismen (Projektionen)

$$\begin{array}{cccc} \pi_1: & V & \to & V_1, \\ & (v_1, v_2) & \mapsto & v_1 \end{array} \qquad \begin{array}{cccc} \pi_2: & V & \to & V_2 \\ & (v_1, v_2) & \mapsto & v_1 \end{array}$$

- $\dim(V_1 \oplus V_2) = \dim(V_1) + \dim(V_2)$
- $\pi_1 \circ i_2 = 0, \pi_2 \circ i_1 = 0$
- $\pi_1 \circ i_1 = \mathrm{id}_{V_1}, \pi_2 \circ i_2 = \mathrm{id}_{V_2}$
- $(i_1 \circ \pi_1)^2 = (i_1 \circ \pi_1), (i_1 \circ \pi_2)^2 = (i_2 \circ \pi_2)$

Beispiel 5.5.2.
- $V = \mathbb{K}^n = \mathbb{K}^{n_1} \times \mathbb{K}^{n_2}$ für $n = n_1 + n_2$
- $0 \oplus V \cong V \cong V \oplus 0$
- $V_1 \oplus (V_2 \oplus V_3) \cong (V_1 \oplus V_2) \oplus V_3$
- $V_1 \oplus V_2 \cong V_2 \oplus V_1$

Satz 5.5.3 (Universelle Eigenschaft („Produkt“)).
- *Zu je zwei linearen Abbildungen $f_1 : W \to V_1, f_2 : W \to V_2$ gibt es genau eine lineare Abbildung $f : W \to V_1 \oplus V_2$ mit $f_1 = \pi_1 \circ f, f_2 = \pi_2 \circ f$:*

$$\begin{tikzcd} W \arrow[dr, dotted, "f"] \arrow[drr, bend left, "f_1"] \arrow[ddr, bend right, "f_2"'] & & \\ & V_1 \oplus V_2 \arrow[r, "\pi_1"'] \arrow[d, "\pi_2"] & V_1 \\ & V_2 & \end{tikzcd}$$

  Anders gesagt:

$$\begin{array}{cccc} \Pi: & \mathrm{Hom}(W, V_1 \oplus V_2) & \longrightarrow & \mathrm{Hom}(W, V_1) \oplus Hom(W, V_2) \\ & f & \longmapsto & (\pi_1 \circ f, \pi_2 \circ f) \end{array}$$

  ist eine Bijektion von Mengen.

- *Darüber hinaus ist* Π *sogar linear, d.h. ein Isomorphismus von Vektorräumen.*

Satz 5.5.4 (Universelle Eigenschaft „Summe“). • *Zu je zwei linearen Abbildungen* $f_1 : V_1 \to W$ *und* $f_2 : V_2 \to W$ *gibt es genau eine lineare Abbildung* $f : V_1 \oplus V_2 \to W$ *mit* $f \circ i_1 = f|_{V_1}$ *und* $f \circ i_2 = f|_{V_2} = f_2$

$$\begin{tikzcd}
W & & \\
& V_1 \oplus V_2 \arrow[ul, dotted, "f"] & V_2 \arrow[l, "i_2"] \arrow[ull, bend right, "f_2"'] \\
& V_1 \arrow[u, "i_1"'] \arrow[uul, bend left, "f_1"] &
\end{tikzcd}$$

Anders gesagt:

$$\begin{array}{rccc} \Sigma: & \mathrm{Hom}(V_1 \oplus V_2, W) & \longrightarrow & \mathrm{Hom}(V_1, W) \oplus \mathrm{Hom}(V_2, W) \\ & f & \longmapsto & (f \circ i_1, f \circ i_2) \end{array}$$

ist eine Bijektion von Mengen.

- *Darüber hinaus ist* Σ *sogar linear, d.h. ein Isomorphismus von Vektorräumen.*

5.5.2 Interne direkte Summe

Definition 5.5.5. Sind $U_1, U_2 \subseteq V$ Untervektorräume und gilt $U_1 + U_2 = V$ sowie $U_1 \cap U_2 = 0$, so heißt V *direkte interne Summe* von U_1 und U_2. Wir nennen U_1 und U_2 *komplementär* (in V).

- $i_1 : U_1 \to V, i_2 : U_2 \to V$ sind Monomorphismen.

Proposition 5.5.6 (Universelle Eigenschaft der Summe, intern). *Es sei* V *die interne direkte Summe von* U_1 *und* U_2.

(i) Zu je zwei Homomorphismen $f_1 : U_1 \to W, f_2 : U_2 \to W$ *gibt es genau einen Homomorphismus* $f : V \to W$ *mit* $f \circ i_1 = f|_{U_1} = f_1 : U_1 \to W$ *und* $f \circ i_2 = f|_{U_2} = f_2 : U_2 \to W$.
Anders gesagt:

$$\begin{array}{rccc} \Sigma: & \mathrm{Hom}(V, W) & \longrightarrow & \mathrm{Hom}(U_1, W) \oplus \mathrm{Hom}(U_2, W) \\ & f & \longmapsto & (f \circ i_1, f \circ i_2) \end{array}$$

ist eine Bijektion von Mengen.

(ii) Darüber hinaus ist Σ *sogar linear, d.h. ein Isomorphismus von Vektorräumen.*

Proposition 5.5.7. *Ist* V *die interne direkte Summe von* U_1 *und* U_2*, so ist* V *isomorph zur externen direkten Summe* $V \cong U_1 \oplus U_2$.

Beweis.

$$\begin{array}{rccc} \phi: & U_1 \oplus U_2 & \longrightarrow & V \\ & (u_1, u_2) & \longmapsto & u_1 + u_2 \end{array}$$

ist ein Isomorphismus:

1) linear
2) Epimorphismus wegen $U_1 + U_2 = V$
3) Monomorphismus wegen $U_1 \cap U_2 = 0$.

□

Damit haben wir auch Projektionen $\pi_i' : V \to U_1$ durch $\pi_i' = \pi_i \circ \phi^{-1}$: ist v zerlegt in $v = u_1 + u_2$, so ist $\pi_i'(v) = u_i$.

Proposition 5.5.8. *Ist* $V = V_1 \oplus V_2$ *die externe direkte Summe von* V_1 *und* V_2*, so ist* V *die interne direkte Summe von* $U_1 = i_1(V_1) = \{(v_1, 0) \in V \mid v_1 \in V_1\}$ *und* $U_2 = i_2(V_2) = \{(0, v_2) \in V \mid v_2 \in V_2\}$.

5.6 Quotientenvektorräume

Es sei $U \subseteq V$ ein Untervektorraum. Wir definieren eine Äquivalenzrelation auf V:

Definition 5.6.1. Zwei Vektoren $v, v' \in V$ heißen *kongruent modulo* U genau dann, wenn $v - v' \in U$ gilt. Notation: $v \equiv v' \mod U \iff v - v' \in U$.

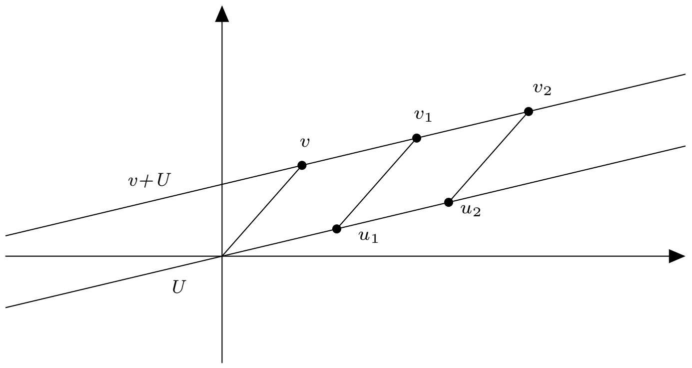
- **Äquivalenzrelation**

  1) $v \equiv v$, da $v - v = 0 \in U$.
  2) $v \equiv v' \iff v' \equiv v$, weil aus $v - v' \in U$ natürlich $v' - v \in U$ folgt.
  3) $v \equiv v', v' \equiv v'' \implies v \equiv v''$, weil aus $v - v' \in U$ und $v' - v'' \in U$ auch $v - v'' = (v - v') + (v' - v'') \in U$ folgt.

- **Äquivalenzklassen**

$$[v]_U := \{v' \in V \mid v \equiv v' \mod U\}$$

Lemma 5.6.2. $[v]_U = v + U$ *(affiner Unterraum)*



Beweis.

$$\begin{array}{rcccl} [v]_U \in & v' & \longmapsto & v + (v - v') & \in v + U \\ & v + u & \longleftarrow & v + u & \end{array}$$

□

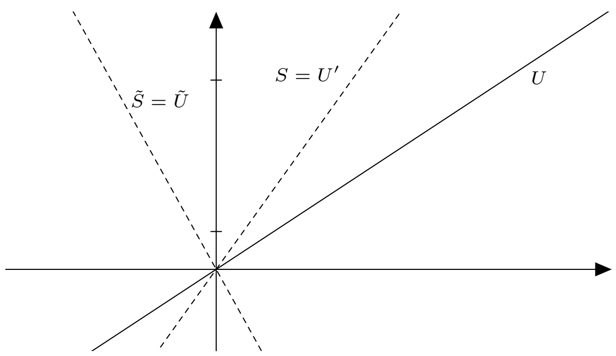
1) $[v]_U \neq \emptyset$, denn $v \in [v]_U$.

2) $[v_1]$ und $[v_2]$ sind gleich ($\Leftrightarrow v_1 \equiv v_2$) oder disjunkt ($\iff v_1 \not\equiv v_2$).

3) $V = \bigcup_{v \in V} [v]_U$.

4) $[0] = U$ ausgezeichnete Klasse

Definition 5.6.3. Ein *„Schnitt“* oder *Repräsentantensystem* ist eine Teilmenge $S \subseteq V$, so dass $S \cap [v]_U$ aus genau einem Element v_s besteht (für alle $v \in V$).



Beispiel 5.6.4. Ist U' komplementärer Unterraum zu U (d.h. $V = U + U', U \cap U' = 0$), so ist $S = U'$ ein Schnitt; V ist in diesem Fall die interne direkte Summe von U und U'.

Für einen Schnitt S haben wir eine Bijektion

$$\begin{array}{rcll} S & \xrightarrow{\cong} & V/U = \{[v]_U \mid v \in V\} & \text{Menge der Äquivalenzklassen} \\ s & \longmapsto & [s]_U & \end{array}$$

Definition 5.6.5. Die Menge V/U wird wie folgt zu einem Vektorraum (über $\mathbb{K}$), welcher *Quotientenvektorraum* heißt:

- Addition:

$$\begin{array}{rcl} V/U \times V/U & \xrightarrow{+} & V/U \\ [v_1]_U + [v_2]_U & := & [v_1 + v_2]_U \end{array}$$

- Null: $0 := [0]_U = U$
- Skalierung $\lambda \cdot [v]_U := [\lambda v]_U$, also $\lambda \cdot (v + U) := \lambda v + U$

Beweis. a) Wohldefiniertheit

- der Addition:

$$\begin{array}{c} v_1 \equiv v_1' \\ v_2 \equiv v_2' \end{array} \implies \begin{array}{c} v_1' = v_1 + u_1 \\ v_2' = v_2 + u_2 \end{array} \implies v_1' + v_2' = v_1 + v_2 + (u_1 + u_2) \implies v_1' + v_2' \equiv v_1 + v_2$$

- der Skalierung:

$$v_1 \equiv v_2 \implies \lambda v_1 \equiv \lambda v_2$$

b) Vektorraumaxiome: nachrechnen.

□

Proposition 5.6.6. *1) $U = 0$: $V/0 \cong V$, $[v] = v + 0 = v$.*

2) $U = V$: $V/V \cong 0$, $[v] = v + V = V$.

3) Sind U, U' komplementäre Untervektorräume von V, so ist $V/U \cong U'$.

Beweis. Wir zeigen 3): Wir haben eine Bijektion

$$\begin{array}{rcl} S = U' & \xrightarrow{\varphi} & V/U \\ u' & \longmapsto & [u']_U \end{array}$$

Diese ist ein Isomorphismus, denn

$$\begin{array}{ccl} u_1' + u_2' & \mapsto & [u_1' + u_2']_U = [u_1']_U + [u_2']_U \\ \lambda u' & \mapsto & [\lambda u']_U = \lambda [u']_U \end{array}$$

□

Proposition 5.6.7. $\dim_{\mathbb{K}} V/U = \dim_{\mathbb{K}} V - \dim_{\mathbb{K}} U$.

Beweis. Es sei $\{b_1, \ldots, b_m\} = \mathfrak{B}'$ eine Basis von U', die wir zu einer Basis $\mathfrak{B} = \{b_1, \ldots, b_m, b_{m+1}, \ldots, b_n\}$ ergänzen.
Dann ist $\overline{\mathfrak{B}} = \{\pi_U(b_{m+1}, \ldots, \pi_U(b_n)\}$ eine Basis für V/U: es ist ein Erzeugendensystem und linear unabhängig. □

Satz 5.6.8 (Universelle Eigenschaft). *(i) Die Abbildung $\pi_U : V \to V/U, v \mapsto [v]_U = v + U$ ist linear und surjektiv.*

(ii) Für jede lineare Abbildung $f : V \to W$ mit $U \subseteq \ker(f)$ gibt es genau eine lineare Abbildung $f' : V/U \to W$ mit $f = f' \circ \pi_U$.

$$\begin{tikzcd} U \subseteq \ker(f) \subseteq & V \arrow[r, "f"] \arrow[d, "\pi_U"'] & W \\ & V/U \arrow[ur, dotted, "f_U"'] & \end{tikzcd}$$

Beweis. (i) π_U ist offenbar surjektiv und nach Konstruktion der Struktur auf V/U auch linear.

(ii) Wir setzen $f_U([v]) = f_U(v+U) = f(v) \in W$.
Dies ist wohldefiniert, denn

$$f(v+u) = f(v) + \underbrace{f(u)}_{=0,\text{ weil } U \subseteq \ker(f)} = f(v)$$

Weil π_U surjektiv ist, ist f_U eindeutig bestimmt.

□

Bemerkung 5.6.9. Der Beweis zeigt, wie man zu einer Basis von V/U kommt:

1) Ist $\{b_1, b_2, \ldots\} = \mathfrak{E}$ ein Erzeugendensystem von V, so ist $\pi_U(\mathfrak{E}) = \{\pi_U(b_1), \ldots\}$ ein Erzeugendensystem von V/U, aus dem man eine Basis auswählen kann.

2) Ist $\mathfrak{B}'$ eine Basis von U und $\mathfrak{B} \supseteq \mathfrak{B}'$ eine ergänzte Basis von V, so ist $\overline{\mathfrak{B}} = \pi_U(\mathfrak{B} \setminus \mathfrak{B}')$ eine Basis von V/U.

Mit dieser 2. Bemerkung sieht man, dass dies auch für unendlich-dimensionale Vektorräume gilt.

Satz 5.6.10. *Jeder Untervektorraum $U \subset V$ besitzt einen komplementären Unterraum U':*

$$(1) \qquad U + U' = V \qquad\qquad (2) \qquad U \cap U' = 0$$

Beweis. Wie oben. □

Satz 5.6.11. *Es sei $f : V \to W$ eine lineare Abbildung. Dann ist*

$$\begin{array}{ccc} V/\ker(f) & \xrightarrow{\cong} & \operatorname{im}(f) \\ [v] & \longmapsto & f(v) \end{array}$$

ein Isomorphismus.

Korollar 5.6.12. *Sei $f : V \to W$ eine lineare Abbildung. Ist U' komplementär zu* $\ker(f) = U$*, so ist*

$$\varphi = f|_{U'}^{W'} : \quad U' \longrightarrow W' := \operatorname{im}(f)$$

ein Isomorphismus.

Beispiel 5.6.13. Wir betrachten ein lineares Gleichungssystem $Ax = b$ mit $A \in \operatorname{Mat}_{m,n}(\mathbb{K}), b \in \mathbb{K}^m$. Sei

$$f = T_A : \mathbb{K}^n \longrightarrow \mathbb{K}^m, x \mapsto Ax,$$

und $V = \mathbb{K}^n, W = K^m$.

- $$\begin{array}{rcccccl} U = & \mathcal{L}(A \mid 0) & = & \ker(T_A) & = & 0 + U & \text{Lösungen des homogenen Systems (linearer Unterraum)} \\ & \mathcal{W}(A) & = & \operatorname{im}(T_A) & \cong & V/\mathcal{L}(A \mid 0) & \end{array}$$

- $v + U = \mathcal{L}(A \mid b)$, falls $f(v) = b$ $(\Leftrightarrow b \in \mathcal{W}(A)$ (affiner Unterraum)
- $\emptyset = \mathcal{L}(A \mid b)$, falls es kein $v \in V$ mit $f(v) = b$ gibt $(\Leftrightarrow b \notin \mathcal{W}(A))$.
- V ist zerlegt in die affinen Unterräume $v + U = \mathcal{L}(A \mid f(v))$ einer davon ist $U = \mathcal{L}(A \mid 0)$.
- Ist U' komplementär zu U so besitzt jeder affine Raum einen eindeutigen Repräsentanten

$$u' + \mathcal{L}(A \mid 0) = \mathcal{L}(A \mid f(u'))$$

5.7 Normale Untergruppen

Sei G eine Gruppe. Wir betrachten für ein festes $\gamma \in G$ die Funktion

$$I_\gamma : \quad G \to G, \qquad I_\gamma(g) = \gamma g \gamma^{-1},$$

welche *Konjugation mit* γ heißt.

(1) I_γ ist ein Homomorphismus von G, denn

$$\begin{aligned} I_\gamma(1) &= \gamma \cdot 1 \cdot \gamma^{-1} = 1, \\ I_\gamma(xy) &= \gamma xy\gamma^{-1} = \gamma x \gamma^{-1}\gamma y \gamma^{-1} = I_\gamma(x) I_\gamma(y) \end{aligned}$$

(2) $I_1 = \mathrm{id}_G$, denn $I_1(x) = 1x1^{-1} = x$.
$I_\alpha \circ I_\beta = I_{\alpha\beta}$, denn $I_\alpha(I_\beta(x)) = \alpha\beta x \beta^{-1}\alpha^{-1} = (\alpha\beta)x(\alpha\beta)^{-1}$.
$I_\alpha^{-1} = I_{\alpha^{-1}}$.

(3) Also ist $I_\gamma : G \to G$ ein Automorphismus für jedes $\gamma \in G$.

(4) Und $I : G \to \mathrm{Aut}(G), \gamma \mapsto I_\gamma$ ist ein Homomorphismus.

(5) Ist $\gamma \in \mathrm{Zen}(G)$, so ist $I_\gamma = \mathrm{id}_G$; und umgekehrt. Also ist $\mathrm{Zen}(G) = \ker(I)$.
Insbesondere ist $I_\gamma = \mathrm{id}$ für alle $\gamma \in G$ wenn G abelsch ist.

Man nennt die I_γ *innere Automorphismen* von G. Sie bilden eine Untergruppe $\mathrm{Inn}(G) \leq \mathrm{Aut}(G)$.
Selbst eine abelsche Gruppe wie $\mathbb{Z}/n$ kann nicht-innere Automorphismen haben:

$$\alpha : \mathbb{Z}/3 = \{\underline{0}, \underline{1}, \underline{2}\} \to \mathbb{Z}/3, \quad \underline{0} \mapsto \underline{0}, \underline{1} \mapsto \underline{2}, \underline{2} \mapsto \underline{1}$$

Es ist $\mathrm{Inn}(\mathbb{Z}/3) = 1$ und $\mathrm{Aut}(\mathbb{Z}/3) = \{\mathrm{id}, \alpha\}$.

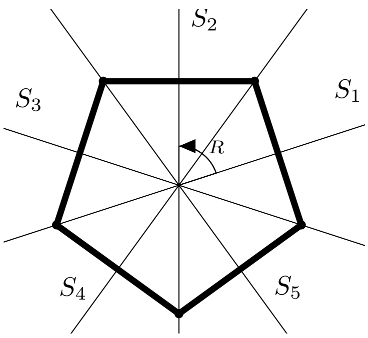
Definition 5.7.1. Eine Untergruppe $H \subset G$ heißt *normal* (oder ein *Normalteiler*, am besten „selbstkonjugiert“), falls $I_\gamma(H) = H$ gilt für alle $\gamma \in G$.

Für eine Untergruppe H von G ist

$$I_\gamma(H) = \gamma H \gamma^{-1} = \left\{\gamma h \gamma^{-1} \mid h \in H\right\} \leq G$$

wieder eine Untergruppe von G, und zwar isomorph zu H, aber vielleicht verschieden von H. Für einen Normalteiler ist hingegen $I_\gamma(H) = H$.

Beispiel 5.7.2. 1) $G = \mathrm{D}_{2n}$ Diedergruppe mit Spiegelungen $S_1, \ldots, S_n$ mit $S_i^2 = 1$ und Rotation R mit $R^n = 1$.



- $H_i := \{1, S_i\}$ ist nicht normal, da $RH_iR^{-1} = H_{i+1}$ wegen $RS_iR^{-1} = S_{i+1}$
- $\mathrm{Rot}_n = \left\{1, R, R^2, \ldots, R^{n-1}\right\}$ ist normal, denn $S_iRS_i^{-1} = R^{-1}$.

2) In einer abelschen Gruppe G ist jede Untergruppe normal.

3) In $G = \mathfrak{S}_n$ ist $H = \{1, (1,2)\}$ nicht normal, denn $(2,3)H(2,3) = \{1, (1,3)\}$.

4) $\mathrm{Inn}(G)$ ist normal in $\mathrm{Aut}(G)$.

5) Die *alternierende Gruppe* $\mathfrak{A}_n := \{\sigma \in \mathfrak{S}_n \mid \mathrm{sign}\,\sigma = 1\} = \ker(\mathrm{sign})$ ist normal in $\mathfrak{S}_n$.

Beweis. Sei $\sigma \in \mathfrak{A}_n$. Dann ist für alle $g \in \mathfrak{S}_n$

$$\begin{aligned} \operatorname{sign}(g\sigma g^{-1}) &= \operatorname{sign}(g)\operatorname{sign}(\sigma)\operatorname{sign}(g^{-1}) \\ &= \operatorname{sign}(g)\cdot 1\cdot \operatorname{sign}(g)^{-1} = 1 \end{aligned}$$

□

Also ist mit σ auch $g\sigma g^{-1}$ wieder in $\mathfrak{A}_n$, für alle $g \in \mathfrak{S}_n$.

Das letzte Beispiel lässt sich verallgemeinern:

Lemma 5.7.3. *Ist $\varphi : G \to G'$ ein Homomorphismus, so ist* $\ker(\varphi)$ *normal in G.*

Beweis. Sei $x \in \ker(\varphi)$, also $\varphi(x) = 1$. Dann gilt für jedes $g \in G$:

$$\begin{aligned} \varphi(gxg^{-1}) &= \varphi(g)\varphi(x)\varphi(g^{-1}) \\ &= \varphi(g)\cdot 1\cdot \varphi(g)^{-1} = 1 \end{aligned}$$

Also ist auch $gxg^{-1} \in \ker(\varphi)$. □

Bemerkung 5.7.4. Es gibt Gruppen, die (außer $H = 1, H = G$) keine normalen Untergruppen haben; diese nennt man *einfache* Gruppen (weil sie in einem gewissen Sinne unzerlegbar sind). Z.B ist die alternierende Gruppe $\mathfrak{A}_n$ einfach für $n \geq 5$.

Definition 5.7.5. (i) Zwei Elemente $g_1, g_2 \in G$ heißen *konjugiert*, wenn es ein $\gamma \in G$ gibt mit $\gamma g_1\gamma^{-1} = g_2$. Notation: $g_1 \sim g_2$.

(ii) Zwei Untergruppen $H_1, H_2 \leq G$ heißen *konjugiert*, wenn es ein $\gamma \in G$ gibt mit $\gamma H_1\gamma^{-1} = H_2$. Notation: $H_1 \sim H_2$.

a) Konjugiertheit ist eine Äquivalenzrelation für Elemente:

- $g \sim g$, denn $I_1(g) = g$.
- $g_1 \sim g_2 \implies \exists\gamma \in G : \gamma g_1\gamma^{-1} = g_2 \Leftrightarrow \exists\gamma' = \gamma^{-1}g_1 = \gamma^{-1}g_2\gamma = \gamma' g_2\gamma'^{-1} \implies g_2 \sim g_1$.
- $g_1 \sim g_2, g_2 \sim g_3 \implies \exists\gamma : \gamma g_1\gamma^{-1} = g_2$, also $\exists\gamma' : \gamma' g_2\gamma'^{-1} = g_3$ und somit $\gamma\gamma' g_1\gamma'^{-1}\gamma^{-1} = (\gamma\gamma')g_1(\gamma\gamma')^{-1} = g_3$, d.h. $g_1 \sim g_2$.

b) Konjugierte Elemente haben gleiche Ordnung. Ist nämlich $g_1^n = 1$ und $\gamma g_1\gamma^{-1} = g_2$, so ist

$$g_2^n = (\gamma g_1\gamma^{-1})^n = \gamma g_1\gamma^{-1}\gamma g_1\gamma^{-1}\ldots\gamma g_1\gamma^{-1} = \gamma g_1^n\gamma^{-1} = \gamma\gamma^{-1} = 1,$$

also ist die Ordnung von g_2 höchstens so groß wie die Ordnung von g_1. Die Rückrichtung ist analog.

c) Ist $\varphi : G \to G'$ ein Homomorphismus und $g_1 \sim g_2$, so auch $\varphi(g_1) \sim \varphi(g_2)$:

$$g_1 \sim g_2 \implies \exists\gamma : \gamma g_1\gamma^{-1} = g_2 \implies \varphi(\gamma g_1\gamma^{-1}) = \varphi(\gamma)\varphi(g_1)\varphi(\gamma)^{-1} = \varphi(g_2)$$

d) Konjugiertheit ist eine Äquivalenzrelation auf der Menge der Untergruppen von G.

e) Konjugierte Untergruppen sind isomorph.

Beispiel 5.7.6. (1) $G = \mathfrak{S}_n$: $\alpha \sim \beta \Leftrightarrow \operatorname{Typ}(\alpha) = \operatorname{Typ}(\beta)$.

(2) $G = D_{2n}$: Ist n ungerade, so sind alle Spiegelungen S_i zueinander konjugiert. Ist n gerade, so zerfallen die Spiegelungen in zwei Konjugationsklassen.

(3) In einer abelschen Gruppe ist jedes Element nur zu sich selbst konjugiert; ebenso für Untergruppen.

(4) In jeder Gruppe ist das neutrale Element 1 nur zu sich selbst konjugiert; ebenso die Untergruppen $H = 1, H = G$.

5.8 Quotientengruppen

Es sei G eine Gruppe und $H \leq G$ eine Untergruppe.

Definition 5.8.1. Für ein $g \in G$ ist seine *Linksnebenklasse* die Teilmenge

$$gH = \{x = gh \mid h \in H\} \subset G,$$

und seine *Rechtsnebenklasse* die Teilmenge

$$Hg = \{y = hg \mid h \in H\} \subset G.$$

Bemerkung 5.8.2. (1) Für $g \in H$ ist $gH = Hg = H$.

(2) Im Allgemeinen ist $gH \neq Hg$, und keine der beiden Nebenklassen ist eine Untergruppe (außer wenn $g \in H$).

(3) G zerfällt sowohl in Links- als auch Rechtsnebenklassen, was man wie folgt sieht:

(4) Wir können zwei Äquivalenzrelationen $\equiv_H$, ${}_H\!\equiv$ auf G definieren:

$$\begin{aligned} x_1 \equiv_H x_2 &\Leftrightarrow \exists h \in H : x_1 h = x_2 \\ &\Leftrightarrow x_1^{-1} x_2 \in H \end{aligned}$$

$$\begin{aligned} x_1 \, {}_H\!\equiv x_2 &\Leftrightarrow \exists h \in H : x_1 = h x_2 \\ &\Leftrightarrow x_1 x_2^{-1} \in H \end{aligned}$$

Für die Äquivalenzklassen $[x]_H$ bzw. ${}_H[H]$ gilt:

$$[x]_H = xH \qquad \text{und} \qquad {}_H[x] = Hx.$$

Lemma 5.8.3. $|gH| = |H|$.

Beweis. $x = gh \mapsto g^{-1}x$ ist eine Bijektion. □

Da alle Äquivalenzklassen somit gleichmächtig sind, ist die Mächtigkeit von G gleich der Mächtigkeit von H mal der Anzahl der Äquivalenzklassen; die letzte Anzahl nennt man den *Index* von H in G, geschrieben $[G : H]$.

Korollar 5.8.4. *Für eine endliche Gruppe gilt: die Ordnung einer Untergruppe H ist immer ein Teiler der Gruppenordnung:*

$$|G| = |H| \cdot [G : H]$$

Lemma 5.8.5. *$H \leq G$ ist genau dann normal, wenn $gH = Hg$ für alle $g \in G$ gilt.*

Beweis. klar (Übung). □

Beispiel 5.8.6. Wir betrachten in $G = \mathfrak{S}_3$ die Untergruppen $H = \{1, \langle 1,2 \rangle\}$, $N = \{1, \langle 1,2,3 \rangle = z, \langle 1,3,2 \rangle = z^2\}$ und listen sämtliche Links- und Rechtsnebenklassen auf, indem wir in den Spalten der folgenden Tabelle für jedes Element σ von $\mathfrak{S}_3$ mit $\times, \bigcirc$ bzw. $\otimes$ markieren, ob σ zur jeweiligen Linksnebenklasse, Rechtsnebenklasse bzw. zu beiden gehört.

$\times$	1	$\langle 1,2 \rangle$	$\langle 1,3 \rangle$	$\langle 2,3 \rangle$	$\langle 1,2,3 \rangle$	$\langle 1,3,2 \rangle$	$\bigcirc$
$1H$	$\otimes$	$\otimes$					$H1$
$\langle 1,2 \rangle H$	$\otimes$	$\otimes$					$H \langle 1,2 \rangle$
$\langle 1,3 \rangle H$			$\otimes$		$\times$	$\bigcirc$	$H \langle 1,3 \rangle$
$\langle 2,3 \rangle H$				$\otimes$	$\bigcirc$	$\times$	$H \langle 2,3 \rangle$
$\langle 1,2,3 \rangle H$			$\times$	$\bigcirc$	$\otimes$		$H \langle 1,2,3 \rangle$
$\langle 1,3,2 \rangle H$			$\bigcirc$	$\times$		$\otimes$	$H \langle 1,3,2 \rangle$
$1N$	$\otimes$				$\otimes$	$\otimes$	$N1$
$\langle 1,2 \rangle N$		$\otimes$	$\otimes$	$\otimes$			$N \langle 1,2 \rangle$
$\langle 1,2 \rangle N$		$\otimes$	$\otimes$	$\otimes$			$N \langle 1,3 \rangle$
$\langle 2,3 \rangle N$		$\otimes$	$\otimes$	$\otimes$			$N \langle 2,3 \rangle$
$\langle 1,2,3 \rangle N$	$\otimes$						$N \langle 1,2,3 \rangle$
$\langle 1,3,2 \rangle N$	$\otimes$				$\otimes$	$\otimes$	$N \langle 1,3,2 \rangle$

Man erkennt insbesondere, dass N ein Normalteiler ist, H aber nicht.

Es sei nun $H \trianglelefteq G$ eine normale Untergruppe. Dann konstruieren wir eine neue Gruppe $\Gamma := G/H = H\backslash G$ genannt *Quotientengruppe*, wie folgt:

- als Menge besteht Γ aus allen Linksnebenklassen (oder Rechtsnebenklassen) gH für $g \in G$. (Ist G endlich, so sind dies $[G : H] = \frac{|G|}{|H|}$ Elemente
- als neutrales Element $1 = 1_\Gamma$ nehmen wir $1H = H$.
- als Multiplikation setzen wir
$$g_1 \cdot g_2 H := (g_1 g_2)H.$$
- als Inverse setzen wir
$$(gH)^{-1} := g^{-1}H$$

Zu zeigen ist nur die Wohldefiniertheit der Multiplikation; alle Gruppenaxiome folgen dann sofort.

Seien $g_1' = g_1 h_1$ und $g_2' = g_2 h_2$ (für $h_1, h_2 \in H$) andere Repräsentanten der gleichen Linksnebenklassen, so haben wir:

$$\begin{aligned} g_1' g_2' &= g_1 h_1 g_2 h_2 \\ &= g_1 g_2 \underbrace{g_2^{-1} h_1 g_2}_{\tilde{h} \in H,\text{ weil H normal}} \\ &= g_1 g_2 \tilde{h} h_2. \end{aligned}$$

Also repräsentieren $g_1' g_2'$ und $g_1 g_2$ die gleiche Nebenklasse.

Proposition 5.8.7. *Sei $H \trianglelefteq G$. Dann gilt:*

i) $\Gamma = G/H$ *ist eine Gruppe.*

ii)
$$\pi = \pi_H^G : \quad G \quad \longrightarrow \quad G/H, \quad g \quad \longmapsto \quad gH$$
ist ein Epimorphismus (genannt kanonische Projektion*).*

iii) $\ker(\pi) = H$.

Beweis. klar nach vorausgegangener Konstruktion. □

Satz 5.8.8 (Universelle Eigenschaft)**.** *Es sei $H \trianglelefteq G$ eine normale Untergruppe und $\varphi : G \to G'$ ein Homomorphismus. Dann sind äquivalent:*

(i) $\ker(\varphi) \geq H$.

(ii) Es gibt einen Homomorphismus $\bar{\varphi} : \Gamma = G/H \to G'$ mit $\varphi = \bar{\varphi} \circ \pi$.

$$\begin{tikzcd} K \arrow[r, hook] & G \arrow[r, "\varphi"] & G' \\ H \arrow[u, dashed] \arrow[r, hook] & G \arrow[u, equal] \arrow[r, "\pi"] & \Gamma \arrow[u, dashed, "\bar{\varphi}"'] \end{tikzcd}$$

Beweis. $(i) \Rightarrow (ii)$: Wir definieren $\bar{\varphi}(gH) := \varphi(g)$.
$\bar{\varphi}$ ist wohldefiniert, denn aus $g' = gh$ $(h \in H)$ folgt mit (i):

$$\bar{\varphi}(g'H) = \varphi(g') = \varphi(gh) = \varphi(g)\underbrace{\varphi(h)}_{=1} = \varphi(g) = \bar{\varphi}(gH).$$

Offenbar ist $\varphi = \bar{\varphi} \circ \pi$.

$(i) \Leftarrow (ii)$: Gibt es ein $\bar{\varphi} : G/H \to G'$ mit $\varphi = \bar{\varphi} \circ \pi$, so folgt sofort für jedes $h \in H$:

$$\varphi(h) = \bar{\varphi}(\pi(h)) = \bar{\varphi}(1) = 1,$$

also $H \leq \ker(\varphi)$. □

Satz 5.8.9. *Für jeden Homomorphismus $\varphi : G \to \tilde{G}$ gilt:*

$$G/\ker(\varphi) \cong \operatorname{im}(\varphi).$$

Beweis. Mit Satz 5.8.8 $(G' = \operatorname{im}(\varphi), H = K = \ker(\varphi), \varphi' = \varphi|^{G'} : G \to G')$ wissen wir, dass es ein $\bar{\varphi}$ gibt mit $\varphi = \bar{\varphi} \circ \pi$.

$$\begin{tikzcd} K = \ker(\varphi) \arrow[r, hook] & G \arrow[r, two heads, "\varphi'"] \arrow[d, "\pi"'] & \operatorname{im}(\varphi) = G' \leq G \\ & G/K \arrow[ur, "\bar{\varphi}"'] & \end{tikzcd}$$

$\bar{\varphi}$ ist also surjektiv. $\bar{\varphi}$ ist auch injektiv, denn aus $\bar{\varphi}(g_1K) = \bar{\varphi}(g_2K)$, also $\varphi(g_1) = \varphi(g_2)$, folgt $g_2^{-1}g_1 \in K$, also $g_1K = g_2K$. □

Beispiel 5.8.10. 1) Die Homomorphismen $\mathrm{id} : G \to G$ und $1 \to G$ zeigen $G/G \cong 1$ (triviale Gruppe) und $G/1 \cong G$.

2) Wegen

$$\mathfrak{A}_n = \ker(\mathrm{sign}) \hookrightarrow \mathfrak{S}_n \xrightarrow{\mathrm{sign}} \mathbb{S}^0 = \{\pm 1\},$$

ist $\mathfrak{S}_n/\mathfrak{A}_n \cong \mathbb{S}^0$.

3) Sei $G = \mathbb{Z}/n = \{\underline{0}, \underline{1}, \ldots, \underline{n-1}\}$ eine zyklische Gruppe und $k \cdot l = n$. Dann haben wir Isomorphismen

$$H = \left\{\underline{0}, \underline{k}, \underline{2k}, \ldots, \underline{(l-1)k}\right\} \xrightarrow{\cong} \mathbb{Z}/l, \underline{xk} \longmapsto \underline{x}, (x = 0, 1, 2, \ldots l-1)$$

und

$$\psi : G/H \xrightarrow{\cong} \mathbb{Z}/k, \underline{g}H \longmapsto \underline{g}, (g = 0, 1, \ldots, k-1).$$

4) Ist $G = G_1 \times G_2$ ein direktes Produkt, so ist $H = G_1 \times 1 = \ker(\pi_{G_2})$ normal und da die Projektion $\pi_{G_2} : G \to G_2$ surjektiv ist, erhalten wir einen Isomorphismus

$$\psi : \begin{array}{ccc} G/H & \xrightarrow{\cong} & G_2 \\ \underbrace{(g_1, g_2)H}_{(1,g_2)H} & \longmapsto & g_2. \end{array}$$

5) Wir haben gesehen, dass in der Diedergruppe $G = D_{2n}$ die Untergruppe $H = \mathrm{Rot}_n \cong Z/n$ der Rotationen um Winkel $\frac{360°}{n} \cdot k, k = 0, 1 \ldots, n-1$ normal ist. Wir erhalten einen Isomorphismus

$$\begin{array}{ccl} G/H & \xrightarrow{\cong} & \mathbb{S}^0 \\ gH & \longmapsto & \begin{cases} +1, & \text{falls } g \in \mathrm{Rot}_n \\ -1, & \text{falls } g \notin \mathrm{Rot}_n \text{ (also Spiegelung oder Produkt aus Spiegelung und Rotation).} \end{cases} \end{array}$$

6) Für die allgemeine lineare Gruppe $G = \mathrm{GL}_n(\mathbb{K})$ ist die Untergruppe der Diagonalmatrizen $H = \mathrm{Zen}(G) = \{D[\lambda, \ldots, \lambda] \mid \lambda \neq 0\}$ ein Normalteiler und

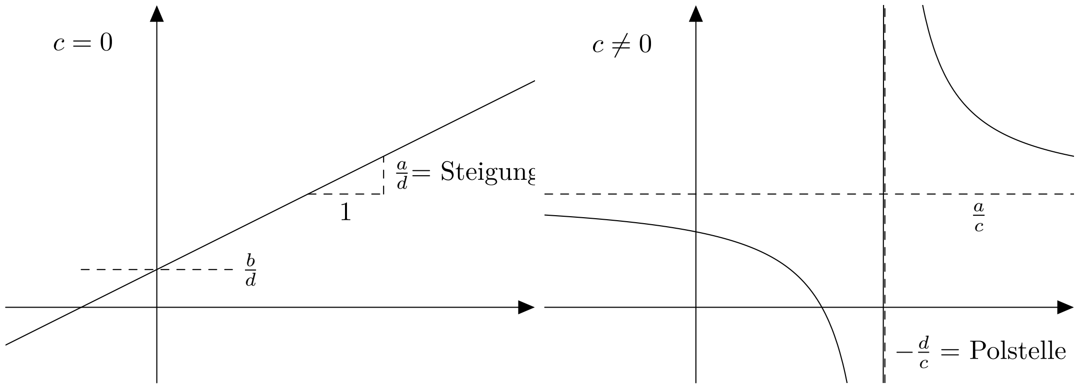
$$\mathrm{PGL}_n(\mathbb{K}) := \mathrm{GL}_n(\mathbb{K})/\mathrm{Zen}(\mathrm{GL}_n(\mathbb{K}))$$

ist die *allgemeine projektive Gruppe*.

7) *Möbius-Gruppe*

$$\mathfrak{M}(\mathbb{R}) := \left\{ f : \mathbb{R} \cup \infty \to \mathbb{R} \cup \infty \mid f(x) = \frac{ax+b}{cx+d}, \text{ für } a,b,c,d \in \mathbb{R} \text{ mit } ad - bc \neq 0 \right\}$$

Dabei setzen wir $f(\infty) := \frac{a}{c}$, insbesondere $f(\infty) = \infty$, falls $c = 0$ und $f(-\frac{d}{c}) = \infty$, falls $c \neq 0$.



$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \overset{\pi}{\longmapsto} f_M(x) = \frac{ax+b}{cx+d}$$

Es gilt:

$$f_{\mathbb{1}} = \mathrm{id}$$
$$f_{M_1 \cdot M_2} = f_{M_1} \circ f_{M_2}$$

$$\begin{tikzcd} \mathrm{Zen}(\mathrm{GL}_2(\mathbb{R})) \arrow[r] \arrow[d, "\cong"'] & \mathrm{GL}_2(\mathbb{R}) \arrow[r, two heads, "\pi"] & \mathfrak{M}(\mathbb{R}) \arrow[d, "\cong"] \\ \mathbb{R}^\times & & \mathrm{PGL}_2(\mathbb{R}) \end{tikzcd}$$

Analoge Betrachtungen können wir für $\mathfrak{M}(\mathbb{C})$ durchführen.

8) *Affine Gruppen*

$$G = \mathrm{Aff}(\mathbb{R}) = \{\text{affine Abbildungen } f : \mathbb{R} \to \mathbb{R}\} = \{f : \mathbb{R} \to \mathbb{R} \mid f(x) = ax + b, a \neq 0\}$$

Diejenigen $f \in \mathrm{Aff}(\mathbb{R})$ mit $b = 0$ bilden die Untergruppe der invertierbaren linearen Abbildungen, welche isomorph ist zu $\mathrm{GL}_1(\mathbb{R}) \cong \mathbb{R}^\times$ (multiplikativ).
Für $a = 1$ erhalten wir die Untergruppe der Translationen $\mathrm{Trans}(\mathbb{R}) \cong \mathbb{R}^\times$ (additiv).
Es gibt eine Bijektion

$$\mathrm{Aff}(\mathbb{R}) \longrightarrow \mathbb{R} \times \mathbb{R}^\times, f \longmapsto (b, a),$$

aber diese ist *kein* Isomorphismus von Gruppen, wenn man rechts die Gruppenstruktur des direkte Produkts $\mathbb{R} \times \mathbb{R}^\times$ nimmt. Stattdessen muss man rechts wie folgt multiplizieren:

$$(b', a') \cdot (b, a) := (a'b + b', a'a).$$

Diese Gruppenstruktur nennt man ein *semidirektes Produkt*.

- $\mathrm{Trans}(\mathbb{R}) \trianglelefteq \mathrm{Aff}(\mathbb{R})$ ist eine normale Untergruppe.
- $\mathrm{GL}_1(\mathbb{R}) \leq \mathrm{Aff}(\mathbb{R})$ ist eine nicht-normale Untergruppe.
- $\mathrm{GL}_1(\mathbb{R})$ ist auch eine Quotientengruppe:

$$\begin{tikzcd} \mathrm{Trans}(\mathbb{R}) \arrow[r, tail] & \mathrm{Aff}(\mathbb{R}) \arrow[r, two heads, "\pi"] & \mathrm{Aff}(\mathbb{R}) / \mathrm{Trans}(\mathbb{R}) \arrow[l, bend left, "s"] & \cong & \mathrm{GL}_1(\mathbb{R}), \end{tikzcd}$$

  Dabei ist $\pi(f) := a$ und $s(a) := (x \mapsto ax)$ und damit $\pi \circ s = \mathrm{id}$.
- Analog: $\mathrm{Aff}(\mathbb{R}^n)$.

6 Determinanten

6.1 Einleitung

Betrachten wir ein LGS $Ax = w$ der Größe 2×2:

$$\begin{bmatrix} ax_1 + bx_2 & = w_1 \\ cx_1 + dx_2 & = w_2 \end{bmatrix}, \qquad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$$

Der Gauß-Algorithmus erlaubt uns, erstens die *Lösbarkeit* (und sogar den Rang) zu bestimmen und zweitens, die *Lösungsmenge* zu parametrisieren. Unser nächstes Ziel ist eine *Lösungsformel*, falls es eine eindeutige Lösung gibt:

$$x_1 = F_1(a, b, c, d, w_1, w_2) = F_1(A, w)$$
$$x_2 = F_2(a, b, c, d, w_1, w_2) = F_2(A, w)$$

Im obigen Beispiel kann man die Formeln rasch herleiten:

(1) $(A \mid w)$ ist genau dann eindeutig lösbar, wenn $D = ad - bc \neq 0$,
(2) und in diesem Fall gilt:

$$x_1 = \frac{dw_1 - bw_2}{D}, \quad x_2 = \frac{aw_2 - cw_1}{D}.$$

Die Größe $D = ad - bc =: \begin{vmatrix} a & b \\ c & d \end{vmatrix}$ nennt man *Determinante* und wir sehen

$$x_1 = \frac{\begin{vmatrix} w_1 & b \\ w_2 & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}, \quad x_2 = \frac{\begin{vmatrix} a & b \\ w_1 & w_2 \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}$$

6.2 Axiomatische Definition

Wir wollen den Begriff der Determinante einer Matrix durch wenige Eigenschaten charakterisieren; erst danach beweisen wir die Existenz und kommen zu einer Formel.

Ist $A = (a_{ij}) \in \mathrm{Mat}_{n,n}(\mathbb{K})$ eine $(n \times n)$-Matrix, so fassen wir sie auch als das n-Tupel $(a_1, \ldots, a_n)$ ihrer Spalten $a_j = a_{\bullet j} = (a_{1j}, a_{2j}, \ldots, a_{nj})^\top$ auf.

Definition 6.2.1. Eine Funktion $D : \mathrm{Mat}_{n,n}(\mathbb{K}) \to \mathbb{K}$ heißt *Determinantenfunktion*, wenn sie folgende Eigenschaften hat:

(D1) *Scherungsinvarianz*: Für alle $1 \leq i < j \leq n$ gilt

$$D(a_1, \ldots, \underbrace{a_i + a_j}_{i\text{-te Stelle}}, \ldots, \underbrace{a_j}_{j\text{-te Stelle}}, \ldots, a_n) = D(a_1, \ldots, \underbrace{a_i}_{i\text{-te Stelle}}, \ldots, \underbrace{a_j}_{j\text{-te Stelle}}, \ldots a_n).$$

(D2) *Spaltenmultiplikativität*: Für alle $1 \leq i \leq n, \lambda \in \mathbb{K}$ gilt

$$D(a_1, \ldots, \underbrace{\lambda a_i}_{i\text{-te Stelle}}, \ldots, a_n) = \lambda D(a_1, \ldots, \underbrace{a_i}_{i\text{-te Stelle}}, \ldots, a_n).$$

(D3) *Alterniertheit*: Für alle $1 \leq i < j \leq n$ gilt

$$D(a_1, \dots, \underbrace{a_i}_{i\text{-te Stelle}}, \dots, \underbrace{a_j}_{j\text{-te Stelle}}, \dots, a_n) = -D(a_1, \dots, \underbrace{a_j}_{i\text{-te Stelle}}, \dots, \underbrace{a_i}_{j\text{-te Stelle}}, \dots, a_n).$$

Bemerkung 6.2.2. Diese Axiome besagen *nicht*, dass D eine lineare Funktion von dem Vektorraum $\mathbb{M} = \mathrm{Mat}_{n,n}(\mathbb{K})$ in den Vektorraum $\mathbb{K}$ ist. Wir werden das noch genauer untersuchen.

Beispiel 6.2.3. 1) $n = 1 : D(a) = a$ (1 Term).
2) $n = 2$: $D\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$ (2 Terme).
3) $n = 3$, bekannt als *Regel von Sarrus*:

$$D\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{32}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}$$

(6 Terme).
4)
- $D(a_1, \dots, a_n) = 0$ ist die triviale Determinantenfunktion.
- Ist D eine Determinantenfunktion, so auch δD für $\delta \in \mathbb{K}$.
- Sind D_1 und D_2 Determinantenfunktionen, so auch $D_1 + D_2$.

Fazit: Die Menge der Determinantenfunktionen $D : \mathrm{Mat}_{n,n}(\mathbb{K}) \to \mathbb{K}$ bilden einen $\mathbb{K}$-Vektorraum, den wir mit $\mathrm{Alt}_n(\mathbb{K})$ bezeichnen. (Aber $\mathrm{Alt}_n(\mathbb{K})$ ist *kein* Untervektorraum von $\mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n \times \dots \times \mathbb{K}^n, \mathbb{K})$.)
Wir werden sehen, dass dieser Vektorraum die Dimension 1 hat und es somit bis auf Normierung nur eine Determinantenfunktion gibt.

Lemma 6.2.4. *Für eine Determinantenfunktion D gilt:*

(i) $D(A) = 0$*, falls A eine Nullspalte besitzt.*
(ii) $D(A) = 0$*, falls zwei Spalten von A gleich sind.*
(iii) $D(A) = 0$*, falls die Spalten von A linear abhängig sind (also* $\mathrm{rg}(A) < n$ *ist).*

Beweis. (i) folgt aus dem Axiom (D2) mit $\lambda = 0$.
(ii) Sei etwa $a_i = a_j$ $(i < j)$, dann rechnet man nach:

$$\begin{aligned} D(a_1, \dots, a_n) &= -D(a_1, \dots, a_i, \dots, -a_j, \dots, a_n) && \text{wegen (D2)} \\ &= -D(a_1, \dots, \underbrace{a_i - a_j}_{=0}, \dots, -a_j, \dots, a_n) && \text{wegen (D1)} \\ &= 0 && \text{wegen (i).} \end{aligned}$$

(iii) Sei etwa $\lambda_1 a_{i_1} + \dots + \lambda_s a_{i_s} = 0$ für $1 \leq i_1 < \dots < i_s \leq n, \lambda_1, \dots, \lambda_s \in \mathbb{K}^\times$, dann rechnen wir nach:

$$\begin{aligned} D(a_1, \dots, a_n) =& \frac{1}{\lambda_1 \cdots \lambda_s} D(a_1, \dots, \lambda_1 a_{i_1}, \dots, \lambda_s a_{i_s}, \dots, a_n) \\ & \quad \text{wegen (D2), angewandt auf die Spalten } i_1, \dots, i_s \\ =& \frac{1}{\lambda_1 \cdots \lambda_s} D(a_1, \dots, \underbrace{\lambda_1 a_{i_1} + \lambda_2 a_{i_2} \dots + \lambda_s a_{i_s}}_{=0}, \dots, \lambda_2 a_{i_2}, \dots, \lambda_s a_{i_s}, \dots, a_n) \\ & \quad \text{die Spalten } i_2, \dots, i_s \text{ zur Spalte } i_1 \text{ addiert} \\ =& \, 0 \qquad \text{wegen (1).} \end{aligned}$$

□

Wir erinnern an die Elementarmatrizen $E_{ij}(\lambda) = \mathbb{1} + \lambda \tilde{E}_{ij}$ (für $i \neq j$) und $E_{ii}(\lambda) = D[1, \dots, \lambda, \dots, 1], \lambda \in \mathbb{K}^\times$ und an die Permutationsmatrizen P_{ij} (für $i \neq j$). Folgende Formeln hatten wir hergeleitet:

1. $E_{ij}(\lambda) = E_{jj}(\lambda^{-1}) E_{ij}(1) E_{jj}(\lambda)$,
2. $P_{ij} = E_{ij}(1) E_{ij}(-1) E_{ji}(1) E_{jj}(-1)$.

Linksmultiplikation mit diesen Matrizen bewirkt folgendens:

$$\begin{aligned} AE_{ij}(\lambda) &= \text{das } \lambda\text{-fache der Spalte } j \text{ wird zur Spalte } i \text{ addiert,} \\ AP_{ij} &= \text{Spalte } i \text{ und Spalte } j \text{ werden vertauscht.} \end{aligned}$$

Lemma 6.2.5. *Für eine Determinantenfunktion D gilt:*

(i) $D(AE_{ij}(\lambda)) = D(A)$, *falls* $i \neq j$.
(ii) $D(AE_{ii}(\lambda)) = \lambda D(A)$.
(iii) $D(AP_{ij}) = -D(A)$, *falls* $i \neq j$.

Beweis. Wir beginnen mit (ii): dies ist eine Umformulierung des Axioms (D2).
(i): Aus Gleichung 1) und (ii) folgt, dass D durch Hinzumultiplizieren der drei Elementarmatrizen erst durch λ dividiert, dann nicht verändert und dann wieder mit λ multipliziert wird.
(iii): Aus Gleichung 2) und (i) und (ii) folgt, dass nur die vierte Elementarmatrix $E_{jj}(-1)$ den Faktor -1 beiträgt. □

Satz 6.2.6. *Es sei $D : \mathrm{Mat}_{n,n}(\mathbb{K}) \to \mathbb{K}$ eine Determinantenfunktion und es gelte $D(\mathbb{1}) = D(e_1, \ldots, e_n) = 0$. Dann ist D die triviale Determinantenfunktion, d.h. $D(A) = 0$ für alle $A \in \mathrm{Mat}_{n,n}(\mathbb{K})$.*

Beweis. Wir bringen A durch Zeilentransformationen Z und Spaltentransformationen S auf Normalform, d.h.

$$Z \cdot A \cdot S = \left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right) \qquad \text{mit } r = \mathrm{rg}(A).$$

Ist $\mathrm{rg}(A) = r < n$, so ist $D(A) = 0$ nach Lemma 6.2.4(iii).
Ist $\mathrm{rg}(A) = n$, so haben wir

$$A = Z^{-1} \cdot S^{-1}.$$

$Z^{-1}S^{-1}$ ist ein Produkt aus Elementarmatrizen $E_{ij}(\lambda)$; nach Lemma 6.2.5 kann man sie von rechts nach links abbauen, wobei sich entweder die Determinante durch Weglassen eines $E_{ij}(\lambda)$ nicht ändert (Fall $i \neq j$, Lemma 6.2.5(i)) oder die Determinante sich um einen Faktor λ ändert (Fall $i \neq j$, Lemma 6.2.5(ii)).
Also haben wir insgesamt

$$D(A) = \delta D(\mathbb{1}) = \delta \cdot 0 = 0,$$

mit $0 \neq \delta =$ Produkt der erwähnten Faktoren. □

Aus diesem Satz folgt nun ein erstaunliches Resultat:

Satz 6.2.7. *Je zwei nicht-triviale Determinantenfunktionen sind proportional.*

Damit folgt: $\mathrm{Alt}_n(\mathbb{K})$ ist höchstens 1-dimensional. Wir wissen aber noch nicht, ob es überhaupt eine nicht-triviale Determinantenfunktion gibt (außer durch Beispiele in den Fällen $n = 1, 2, 3$).

Beweis. Es seien D_1, D_2 zwei nicht-triviale Determinantenfunktionen. Nach Satz 6.2.6 gilt dann $D_1(\mathbb{1}), D_2(\mathbb{1}) \neq 0$.
Wir betrachten die neue Determinantenfunktion

$$D : \mathrm{Mat}_{n,n}(\mathbb{K}) \longrightarrow \mathbb{K}, \quad D(A) := D_2(\mathbb{1})D_1(A) - D_1(\mathbb{1})D_2(A).$$

Wegen $D(\mathbb{1}) = D_2(\mathbb{1})D_1(\mathbb{1}) - D_1(\mathbb{1})D_2(\mathbb{1}) = 0$, muss D trivial sein nach Satz 6.2.6. Also

$$D_2(\mathbb{1})D_1(A) = D_1(\mathbb{1})D_2(A) \quad \text{für alle } A,$$

und somit

$$D_2(A) = \frac{D_2(\mathbb{1})}{D_1(\mathbb{1})} \cdot D_1(A).$$

□

Definition 6.2.8. Eine Determinantenfunktion D heißt *normiert*, wenn $D(\mathbb{1}) = 1$ gilt.

Es gibt also höchstens eine normiertes Element in $\mathrm{Alt}_n(\mathbb{K})$. Unsere Beispiele in den Dimensionen $n = 1, 2, 3$ sind normiert.

Satz 6.2.9 (Produktsatz für Determinanten)**.** *Für eine normierte Determinantenfunktion D gilt:*

$$D(A \cdot B) = D(A) \cdot D(B) \quad \textit{für alle } A, B \in \mathrm{Mat}_{n,n}(\mathbb{K}).$$

Beweis. Für ein festes $A \in \mathbb{M} = \mathrm{Mat}_{n,n}(\mathbb{K})$ betrachten wir zwei Funktionen

$$\begin{aligned} D_1 &: \mathbb{M} \longrightarrow \mathbb{K}, \quad D_1(X) := D(A)D(X), \\ D_2 &: \mathbb{M} \longrightarrow \mathbb{K}, \quad D_2(X) := D(AX). \end{aligned}$$

(1) Offenbar ist D_1 eine Determinantenfunktion.

(2) Um zu zeigen, dass auch D_2 eine ist, schreiben wir $X = (x_1, \dots, x_n)$ als Spaltentupel, genau wie $A^\top = (a_1^\top, \dots, a_n^\top)$ mit $a_i^\top = i$-te Zeile von A.
Dann hat die Matrix $C = AX$ die Einträge $c_{ij} = \langle a_i^\top, x_j \rangle$ mit dem Standardskalarprodukt $\langle \cdot, \cdot \rangle$ auf $\mathbb{K}^n \times \mathbb{K}^n$, d.h.

$$c_j = \begin{pmatrix} \langle a_1^\top, x_1 \rangle \\ \vdots \\ \langle a_n^\top, x_j \rangle \end{pmatrix}.$$

Vergleichen wir also

$$A \cdot (x_1, \dots, x_i + x_j, \dots, x_j, \dots, x_n) = (c_1, \dots, c_i + c_j, \dots, c_j, \dots, c_n)$$

mit

$$A(x_1, \dots, x_i, \dots, x_j, \dots, x_n) = (c_1, \dots, c_i, \dots, c_j, \dots, c_n)$$

unter D, so ergibt sich die Scherungsinvarianz von D_2.
Genauso ergibt sich aus

$$A(x_2, \dots, \lambda x_i, \dots, x_n) = (c_1, \dots, \lambda c_i, \dots, c_n)$$

die Linearität in der i-ten Spalte von D_2.

(3) Demnach ist

$$\begin{aligned} D_3(X) :=& D_1(X) - D_2(X) \\ =& D(A)D(X) - D(AX) \end{aligned}$$

auch eine Determinantenfunktion.
Wegen

$$\begin{aligned} D_3(\mathbb{1}) &= D(A)D(\mathbb{1}) - D(A\mathbb{1}) \\ &= D(A) \cdot 1 - D(A) = 0 \end{aligned}$$

folgt $D_3(X) = 0$ für alle $X \in \mathbb{M}$ nach Satz 6.2.6. Das ist gerade der Produktsatz.

□

Für normierte Determinantenfunktionen haben wir folgende Korollare:

Korollar 6.2.10. *$A \in \mathrm{Mat}_{n,n}(\mathbb{K})$ ist genau dann invertierbar, wenn $D(A) \neq 0$ ist; in diesem Fall ist $D(A^{-1}) = D(A)^{-}$*

Beweis. Ist A nicht invertierbar, so ist $\mathrm{rg}(A) < n$, und deshalb $D(A) = 0$ nach Lemma 6.2.4.
Ist A invertierbar, so folgt einerseits $D(A \cdot A^{-1}) = D(\mathbb{1}) = 1$ und andererseits $D(A \cdot A^{-1}) = D(A) \cdot D(A^{-1})$. □

Korollar 6.2.11. *(i) $D(AB) = D(BA)$ für alle $A, B \in \mathrm{Mat}_{n,n}(\mathbb{K})$.*

(ii) $D(\Omega A \Omega^{-1}) = D(A)$ für alle $A \in \mathrm{Mat}_{n,n}(\mathbb{K})$ und $\Omega \in \mathrm{GL}_n(\mathbb{K})$.

Beweis. (i) ist klar, weil $\mathbb{K}^\times$ kummutativ ist, also

$$D(AB) = D(A)D(B) = D(B)D(A) = D(BA).$$

(ii) folgt aus (i).

□

Korollar 6.2.12. *$D(A^\top) = D(A)$ für alle $A \in \mathrm{Mat}_{n,n}(\mathbb{K})$.*

Beweis. Dies folgt aus einer Zerlegung in Elementarmatrizen E_{ij} mit $E_{ij}(\lambda)^\top = E_{ji}(\lambda)$. □

6.2.1 Einige Berechnungen für normierte Determinantenfunktionen

Sei D eine normierte Determinantenfunktion. Wir berechnen $D(A)$ für spezielle Matrizen $A \in \mathrm{Mat}_{n,n}$:

a) *Diagonalmatrizen*:
$$D(D[\lambda_1, \ldots, \lambda_n]) = \lambda_1 \cdots \lambda_n$$
Um das einzusehen wendet man das Axiom (D2) auf jede Spalte an:
$$D(\lambda_1 e_1, \ldots, \lambda_n e_n) = \lambda_1 \cdots \lambda_n D(e_1, \ldots, e_n) = \lambda_1 \cdots \lambda_n.$$

b) *Zentrale Matrizen*:
$$D(D[\lambda, \ldots, \lambda]) = \lambda^n$$

c) *Permutationsmatrizen*:
$$D(P_\sigma) = \mathrm{sign}(\sigma)$$
Für eine Transposition $\sigma = (ij)$ mit $i \neq j$ gilt $D(P_{ij}) = -1$ nach dem Axiom $(D3)$ bzw. Lemma 6.2.5(iii). Schreibt man σ als Produkt von s Transpositionen, so folgt $D(P_\sigma) = (-1)^s = \mathrm{sign}(\sigma))$.

d) *Elementarmatrizen*:
$$D\left(E_{ij}(\lambda)\right) = \begin{cases} 1, & i \neq j \\ \lambda, & i = j \end{cases}$$
Dies folgt aus Lemma 6.2.5(i) und (ii) mit dem Produktsatz.

e) *Dreiecksmatrizen* (obere und untere):
$$D\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = D\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ * & & \lambda_n \end{pmatrix} = \lambda_1 \cdots \lambda_n$$
Wir können $\lambda_1, \ldots, \lambda_n = 0$ annehmen. Wir ziehen λ_1 aus der ersten Spalte, subtrahieren dann das a_{12}-fache der ersten Spalte von der zweiten, ziehen den Faktor λ_2 aus der zweiten Spalte, u.s.w.:
$$\begin{aligned} D(\lambda_1 e_1, a_2, \ldots, a_n) &= \lambda_1 D(e_1, \lambda_2 e_2 + a_{12} e_1, a_3, \ldots, a_n) \\ &= \lambda_1 D(e_1, \lambda_2 e_2, a_3, \ldots, a_n) \\ &= \lambda_1 \lambda_2 D(e_1, e_2, \lambda_3 e_3 + a_{13} e_1 + a_{23} e_2, a_4, \ldots, a_n) \\ &= \lambda_1 \lambda_2 D(e_1, e_2, e_3, \lambda_4 e_4 + \ldots, \ldots, a_n) \\ &= \lambda_1 \lambda_2 \cdots \lambda_n D(e_1, e_2, \ldots, e_n). \end{aligned}$$

f) *Vielfache von Matrizen*:
$$D(\lambda A) = \lambda^n D(A)$$
Man wendet das Axiom (D2) auf jede Spalte an.

6.2.2 Multilineare Funktionen

Definition 6.2.13. Seien V und W $\mathbb{K}$-Vektorräume. Eine Funktion

$$F : V \times \ldots \times V = V^k \longrightarrow W$$

- heißt (k-fach) *multilinear*, falls für alle $v_1, \ldots, v_k, v_i' \in V, \alpha, \alpha' \in \mathbb{K}, 1 \leq i < j \leq k$ gilt:

$$F(v_1, \ldots, \alpha v_i + \alpha' v_i', \ldots, v_k) = \alpha F(v_1, \ldots, v_i, \ldots, v_k) + \alpha' F(v_1, \ldots, v_i', \ldots, v_k)$$

- heißt *alternierend*, falls für alle $1 \leq i < j \leq k$ gilt:

$$F(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = -F(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k).$$

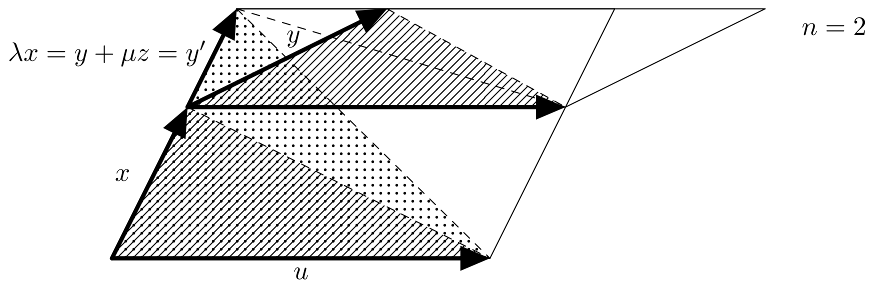
Beispiel 6.2.14. 1) $k = 1$: 1-fach multilinear ist unser altes linear.

2) $k = 2$: Das Skalarprodukt $\langle \cdot, \cdot \rangle : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$,

$$\langle x, y \rangle = x_1 y_1 + \ldots + x_n y_n$$

ist bilinear. Es ist nicht alternierend (sondern es ist symmetrisch: $\langle x, y \rangle = \langle y, x \rangle$).

Geometrisches Argument (Scherungsinvarianz und Multilinearität)



Das Volumen im $\mathbb{R}^n$ ist eine scherungsinvariante Größe:

$$\text{vol}(x, z) = \text{Flächeninhalt des von } x \text{ und } z \text{ aufgespannten Dreiecks},$$

$$\text{vol}(x + y, z) = vol(x, z) + vol(y, z),$$

d.h. die Flächeninhalte der beiden schraffierten Dreiecke summieren sich zum Flächeninhalt des gepunkteten Dreiecks.

Satz 6.2.15. *Eine Determinantenfunktion*

$$D : \text{Mat}_{n,n}(\mathbb{K}) = \mathbb{K}^n \times \ldots \times \mathbb{K}^n \longrightarrow \mathbb{K}$$

ist n-fach multilinear und alternierend.

Beweis. Lemma 6.2.5(iii) besagt, dass D alternierend ist.
Die Multilinearität muss man nur für eine Spalte (sagen wir der ersten) zeigen; für die anderen folgt sie wegen der Alterniertheit. Die Multiplikativität der ersten Spalte folgt aus Axiom (D2); also geht es nur noch um die Additivität.
Um nun

$$D(a_1 + b_1, a_2, \ldots, a_n) = D(a_1, a_2, \ldots, a_n) + D(b_1, a_2, \ldots, a_n)$$

zu zeigen, unterscheiden wir drei Fälle.

<u>1. Fall</u> $\text{rg}(a_2, \ldots, a_{n-1}) < n - 1$.
Dann sind die Ränge aller drei Matrizen $(a_1 + b_1, a_2, \ldots, a_n), (a_1, a_2, \ldots, a_n)$ und $(b_1, a_2, \ldots, a_n)$ kleiner als n. Nach Lemma 6.2.4 sind alle drei Werte von D null.

2. Fall $\mathrm{rg}(a_2, \ldots, a_n) = n-1$ und $\mathrm{rg}(a_1, a_2, \ldots, a_n) = n-1$.
Dann ist $D(a_1, a_2, \ldots, a_n) = 0$ und $a_1 \in \mathrm{Span}(a_2, \ldots, a_n)$, etwa $a_1 = \alpha_2 a_2 + \ldots + \alpha_n a_n$. Man subtrahiert nacheinander das α_i-fache der Spalte i von der Spalte 1 $(i = 2, \ldots, n)$ in der Matrix $(a_1 + b_1, a_2, \ldots, a_n)$ und erhält:

$$D(a_1 + b_1, a_2, \ldots, a_n) = D(b_1, a_2, \ldots, a_n).$$

3. Fall $\mathrm{rg}(a_2, \ldots, a_n) = n-1$ und $\mathrm{rg}(a_1, a_2, \ldots, a_n) = n$.
Dann ist $b_1 \in \mathrm{Span}(a_1, a_2, \ldots, a_n)$, etwa $b_1 = \beta_1 a_1 + \beta_2 a_2 + \ldots + \beta_n a_n$. Man subtrahiert nacheinander das β_i-fache der Spalte i von Spalte 1 $(i = 2, \ldots, n)$ in der Matrix $(a_1 + b_1, a_2, \ldots, a_n)$ und erhält:

$$\begin{aligned} D(a_1 + b_1, a_2, \ldots, a_n) &= D(a_1 + \beta_1 a_1 + \beta_2 a_2 + \ldots + \beta_n a_n, a_2, \ldots, a_n) \\ &= D(a_1 + \beta_1 a_1, a_2, \ldots, a_n) \\ &= (1 + \beta_1) D(a_1, a_2, \ldots, a_n). \end{aligned}$$

Andererseits ist ebenso nach sukzessiver Subtraktion

$$\begin{aligned} D(b_1, a_2, \ldots, a_n) &= D(\beta_1 a_1 + \beta_2 a_2 + \ldots + \beta_n a_n, a_2, \ldots, a_n) \\ &= D(\beta_1 a_1, a_2, \ldots, a_n) \\ &= \beta_1 D(a_1, a_2, \ldots, a_n). \end{aligned}$$

Also ist

$$D(a_1, a_2, \ldots, a_n) + D(b_1, a_2, \ldots, a_n) = (1 + \beta_1) D(a_1, a_2, \ldots, a_n).$$

□

Bemerkung 6.2.16. 1) Man beachte, dass wir die Alterniertheit aus den Axiomen (D1) und (D2) gefolgert haben, siehe Beweis von Lemma 6.2.5(iii).

2) Dann folgte der Produktsatz und die Multilinearität.

6.3 Leibniz-Formel

Wir entwickeln nun endlich eine Formel für eine Determinantenfunktion. Damit erst erhalten wir einen Existenzbeweis, dass es für jedes n eine nicht-triviale Determinantenfunktion gibt.

Satz 6.3.1 (Leibniz-Formel)**.** *Für eine normierte Determinantenfunktion D gilt für jede Matrix $A = (a_{ij}) \in \mathrm{Mat}_{n,n}(\mathbb{K})$:*

$$D(A) = \sum_{\pi \in \mathfrak{S}_n} \mathrm{sign}(\pi) a_{\pi(1)1} \cdot \ldots \cdot a_{\pi(n)n}.$$

Beweis. Wir benutzen die Standardbasis $e_1, \ldots, e_n$ in $\mathbb{K}^n$ und schreiben jeden Spaltenvektor

$$a_j = \sum_{i=1}^{n} a_{ij} e_i \qquad (j = 1, \ldots, n).$$

Wir benutzen die Linearität in jeder Spalte und entwickeln:

$$\begin{aligned} D(a_1, a_2, \ldots, a_n) &= D\left(\sum_{i_1=1}^{n} a_{i_1 1} e_{i_1}, a_2, \ldots, a_n\right) \\ &= \sum_{i_1=1}^{n} a_{i_1,1} D\left(e_{i_1}, \sum_{i_2=1}^{n} a_{i_2 2} e_{i_2}, a_3, \ldots, a_n\right) \\ &= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} a_{i_1 1} a_{i_2 2} D\left(e_{i_1}, e_{i_2}, \sum_{i_3=1}^{n} a_{i_3 3} e_3, a_4, \ldots, a_n\right) \\ &\vdots \\ &= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_n=1}^{n} a_{i_1 1} a_{i_2 2} \cdots a_{i_n n} D\left(e_{i_1}, \ldots, e_{i_n}\right) \end{aligned}$$

In dieser Summe stehen n^n Terme:

- Der Term für den Multiindex $I = (i_1, \ldots, i_n)$ besteht aus dem Produkt der folgenden Einträge

$$a_I := a_{i_1 1} a_{i_2 2} \cdots a_{i_n n}$$

multipliziert mit dem Wert $D(e_{i_1}, \ldots, e_{i_n}$ der Matrix $E_I := (e_{i_1}, \ldots, e_{i_n})$.

Kommt in I ein Index zweimal vor ($i_r = i_s$ für $r \neq s$), so ist $D(E_I) = 0$ nach Lemma 6.2.4(ii). Also brauchen wir nur die wiederholungsfreien I zu berücksichtigen; diese I entsprechen genau den Permutationen $\pi \in \mathfrak{S}_n$ durch die Bijektion $I = (i_1, \ldots, i_n) \leftrightarrow \pi(k) = i_k \quad (k = 1, \ldots, n)$.
Dann ist $E_I = P_\pi$, also

$$D(E_I) = D(P_\pi) = \operatorname{sign}(\pi).$$

Damit haben wir die Leibniz-Formel

$$D(A) = \sum_{\pi \in \mathfrak{S}_n} \operatorname{sign}(\pi) a_{\pi(1)1} \cdot \ldots \cdot a_{\pi(n)n}.$$

□

Damit haben wir endlich auch die Existenz einer nicht-trivialen Determinantenfunktion bewiesen, denn die rechte Seite der Leibniz-Formel definiert offenbar für jedes $n \geq 1$ eine solche.

Satz 6.3.2. *Für jedes $n \geq 1$ ist der Vektorraum $\operatorname{Alt}_n(\mathbb{K})$ 1-dimensional; es gibt genau eine normierte Determinantenfunktion.*

Definition 6.3.3. Die eindeutige normierte Determinantenfunktion nennen wir *die* Determinante

$$\operatorname{Det} : \operatorname{Mat}_{n,n}(\mathbb{K}) \to \mathbb{K}.$$

6.4 Determinante für Endomorphismen

Die Eigenschaft $\operatorname{Det}(\Omega A \Omega^{-1}) = \operatorname{Det}(A)$ erlaubt uns nun, auch für Endomorphismen eine Determinante zu erklären:

Definition 6.4.1. Sei $f : V \to$ ein Endomorphismus, $\dim V = n$. Wir wählen eine Basis $\mathfrak{B}$ von V und setzen:

$$\operatorname{Det}(f) := \operatorname{Det}\left(M_{\mathfrak{B}\mathfrak{B}}(f)\right).$$

Dies ist unabhängig von der Wahl der Basis $\mathfrak{B}$, denn ist $\mathfrak{A}$ eine andere Basis und $\Omega = \Omega_{\mathfrak{A}\mathfrak{B}} = \operatorname{Mat}_{\mathfrak{A}\mathfrak{B}}(\operatorname{id}_V)$ die Basiswechselmatrix, so ist

$$M_{\mathfrak{A}\mathfrak{A}}(\operatorname{id}_V) = \Omega_{\mathfrak{A}\mathfrak{B}} M_{\mathfrak{B}\mathfrak{B}}(\operatorname{id}_V) \Omega_{\mathfrak{A}\mathfrak{B}}^{-1}.$$

Also haben beide darstellenden Matrizen die gleiche Determinante.

Satz 6.4.2. *Die Determinante* $\operatorname{Det} : \operatorname{End}_{\mathbb{K}}(V) \to \mathbb{K}$ *hat folgende Eigenschaften:*

1) $\operatorname{Det}(\operatorname{id}_V) = 1$.

2) $\operatorname{Det}(f \circ g) = \operatorname{Det}(f) \cdot \operatorname{Det}(g)$.

3) $\operatorname{Det}(f) \neq 0$ *genau dann, wenn f ein Isomorphismus ist. In diesem Fall ist* $\operatorname{Det}(f^{-1}) = \operatorname{Det}(f)^{-1}$.

Beispiel 6.4.3. 1) $\operatorname{Det}(0) = 0$.
2) $\operatorname{Det}(\lambda \operatorname{id}_V) = \lambda^n$, insbesondere $\operatorname{Det}(-\operatorname{id}_V) = (-1)^k$.

6.5 Spur

Die Spur einer Matrix ist eine Größe, die mit der Determinante in gewisser Beziehung steht.

Definition 6.5.1. Die *Spur* $A = (a_{ij}) \in \mathrm{Mat}_{n,n}(\mathbb{K})$ ist

$$\mathrm{Spur}(A) = \sum_{i=0}^{n} a_{ii}.$$

Satz 6.5.2. *Die Spur ist eine lineare Abbildung*

$$\mathrm{Spur} : \mathrm{Mat}_{n,n}(\mathbb{K}) \longrightarrow \mathbb{K}.$$

Also gilt für alle $A, B \in \mathrm{Mat}_{n,n}(\mathbb{K})$ *und* $\lambda \in \mathbb{K}$*:*

(1) $\mathrm{Spur}(\lambda A) = \lambda \,\mathrm{Spur}(A)$*,*

(2) $\mathrm{Spur}(A + B) = \mathrm{Spur}(A) + \mathrm{Spur}(B)$*.*

Weiter gilt:

(4) $\mathrm{Spur}(AB) = \mathrm{Spur}(BA)$*, und insbesondere*

(5) $\mathrm{Spur}(\Omega A \Omega^{-1}) = \mathrm{Spur}(A)$ *für invertierbares* Ω*.*

Beweis. Übung. □

Bemerkung 6.5.3. Spur-Bedingungen haben oft etwas mit Fixpunkten zu tun:

$$\mathrm{Fix}(f) = \emptyset \implies \mathrm{Spur}(f_*) = 0,$$

wobei wir offen lassen, was die „induzierte“ Abbildung f_* ist.
Zum Beispiel gilt für eine Permutationsmatrix

$$\mathrm{Spur}(P_\sigma) = |\mathrm{Fix}(\sigma)|\,.$$

Definition 6.5.4. Für einen Endomorphismus $f : V \to V$ definieren wir

$$\mathrm{Spur}(f) = \mathrm{Spur}(M_{\mathfrak{B}\mathfrak{B}}(f)) \in \mathbb{K}$$

für irgendeine Basis $\mathfrak{B}$ von V. Dieser Skalar ist von der Wahl der Basis unabhängig: ist $\mathfrak{A}$ eine andere Basis und $\Omega_{\mathfrak{A}\mathfrak{B}}$ die Basiswechselmatrix von $\mathfrak{B}$ nach $\mathfrak{A}$, so gilt

$$M_{\mathfrak{A}\mathfrak{A}}(f) = \Omega_{\mathfrak{A}\mathfrak{B}} M_{\mathfrak{B}\mathfrak{B}}(f) \Omega_{\mathfrak{A}\mathfrak{B}}^{-1},$$

also haben beide Matrizen $M_{\mathfrak{A}\mathfrak{A}}(f)$ und $M_{\mathfrak{B}\mathfrak{B}}(f)$ die gleiche Spur.

Satz 6.5.5. *Die Spur* $\mathrm{Spur} : \mathrm{End}_{\mathbb{K}}(V) \to \mathbb{K}$ *ist eine lineare Abbildung mit*

(1) $\mathrm{Spur}(f \circ g) = \mathrm{Spur}(g \circ f)$*, insbesondere*
(2) $\mathrm{Spur}(\varphi \circ f \circ \varphi^{-1}) = \mathrm{Spur}(f)$*.*

Beispiel 6.5.6.

-
$\mathrm{Spur} : \mathrm{Mat}_{n,n}(\mathbb{K}) \to \mathbb{K}$	$\mathrm{Spur} : \mathrm{End}_{\mathbb{K}}(V) \to \mathbb{K}$
$\mathrm{Spur}(\mathbb{1}) = n$	$\mathrm{Spur}(\mathrm{id}_V) = \dim V$
$\mathrm{Spur}(0) = 0$	$\mathrm{Spur}(0) = 0$
$\mathrm{Spur}(\lambda \mathbb{1}) = n\lambda$	$\mathrm{Spur}(\lambda\, \mathrm{id}_V) = \lambda \dim V$
$\mathrm{Spur}(-\mathbb{1}) = -n$	$\mathrm{Spur}(-\mathrm{id}_V) = -\dim V$

- $\mathrm{Spur}(A^\top) = \mathrm{Spur}(A)$
- obere/untere Dreiecksmatrizen

$$\mathrm{Spur}\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \mathrm{Spur}\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ * & & \lambda_n \end{pmatrix} = \lambda_1 + \ldots + \lambda_n$$

Beispiel 6.5.7 (Heisenberg-Gleichung). Die Gleichung

$$\Psi \circ \Phi - \Phi \circ \Psi = \mathrm{id}$$

für zwei lineare Abbildungen $\Phi, \Psi : V \to V$ spielt in der Quantenmechanik eine große Rolle.

- Ist V endlich-dimensional, so können wir die Spur anwenden und erhalten

$$\begin{aligned} n = \dim_{\mathbb{K}}(V) = \operatorname{Spur}(\mathrm{id}_V) &= \operatorname{Spur}(\Psi \circ \Phi - \Phi \circ \Psi) \\ &= \operatorname{Spur}(\Psi \circ \Phi) - \operatorname{Spur}(\Phi \circ \Psi) = 0. \end{aligned}$$

  Ist also $V \neq 0$ und $\operatorname{char}(\mathbb{K}) \neq n$, so gibt es keine Lösungen dieser Gleichung.

- Für den unendlich-dimensionalen reellen Vektorraum $V = \mathcal{C}^1(\mathbb{R})$ der einmal stetig-differenzierbaren Funktionen $f : \mathbb{R} \to \mathbb{R}$ seien

$$\begin{aligned} \Phi &: V \to V, \quad \Phi(f)(x) := xf(x), \\ \Psi &: V \to V, \quad \Psi(f) := f' \end{aligned}$$

  und man sieht sofort mit der Produktregel für die Ableitung:

$$\Psi \circ \Phi - \Phi \circ \Psi = \mathrm{id}\,.$$

  In der Quantenmechanik benutzt man den Hilbert-Raum $\mathrm{L}^2(\mathbb{R})$ der quadrat-integrierbaren Funktionen; sie beschreiben die Aufenthaltswahrscheinlichkeit eines Teilchens. Φ ist der Ortsoperator, Ψ dier Impulsoperator. Aus der Gleichung folgt die Heisenbergsche Unschärferelation: man kann Ort und Impuls eines Teilchens nicht gleichzeitig exakt messen.

6.6 Cramersche Regel

Wir können nun die Formeln für die Lösungen eines eindeutig lösbaren LGS herleiten. Es sei $A = (a_1, \ldots, a_n)$, $b \in \mathbb{K}^n$ und $\operatorname{Det}(A) \neq 0$. Dann ist das LGS

$$Ax = b$$

für jedes b eindeutig lösbar. Angenommen, $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ ist die Lösung, dann ist

$$x_1 a_1 + x_2 a_2 + \ldots + x_n a_n = b, \tag{6.1}$$

als Linearkombination der Spalten von A. Setzen wir (6.1) als erste Spalte in die Determinante, so erhalten wir

$$\begin{aligned} \operatorname{Det}(b, a_2, \ldots, a_n) &= \operatorname{Det}\left(\sum_{i=1}^n x_i a_i, a_2, \ldots, a_n\right) \\ &= \sum_{i=1}^n x_i \underbrace{\operatorname{Det}(a_i, a_2, \ldots, a_n)}_{=0,\text{ falls } i=2,3,\ldots,n} \\ &= x_1 \operatorname{Det}(a_1, a_2, \ldots, a_n), \end{aligned}$$

denn nur für $i = 1$ ist der Summand $\neq 0$. Also haben wir

$$x_1 = \frac{\operatorname{Det}(b, a_2, \ldots, a_n)}{\operatorname{Det}(a_1, a_2, \ldots, a_n)}.$$

Genauso verfahren wir für die anderen Variablen x_k.

Satz 6.6.1. *Das LGS $Ax = b$ mit n Unbekannten $x_1, \ldots, x_n$ und n Gleichungen ist genau dann für jedes b eindeutig lösbar, wenn* $\operatorname{Det}(A) \neq 0$ *gilt; und in diesem Fall ist*

$$x_k = \frac{\operatorname{Det}(a_1, \ldots, a_{k-1}, b, a_{k+1}, \ldots, a_n)}{\operatorname{Det}(a_1, \ldots, a_{k-1}, a_k, a_{k+1}, \ldots, a_n)} \qquad (k = 1, \ldots, n).$$

Bemerkung 6.6.2. Wir sehen, dass jedes x_k eine rationale Funktion in den Einträgen a_{ij} von A und den Koordinaten $b_1, \dots, b_n$ von b ist.
Liegen die Einträge a_{ij} und b_k in einem Unterkörper $\mathbb{L}$ von $\mathbb{K}$, so auch die Lösungen $x_1, \dots, x_n$.

Beispiel 6.6.3.

$$A = \begin{pmatrix} \lambda_1 & \alpha & \beta \\ 0 & \lambda_2 & \gamma \\ 0 & 0 & \lambda_3 \end{pmatrix} \qquad b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \qquad \operatorname{Det}(A) = \lambda_1\lambda_2\lambda_3 \neq 0$$

$$x_1 = \frac{\begin{vmatrix} b_1 & \alpha & \beta \\ b_2 & \lambda_2 & \gamma \\ b_3 & 0 & \lambda_3 \end{vmatrix}}{\lambda_1\lambda_2\lambda_3} = \frac{\lambda_2\lambda_3 b_1 - \alpha\lambda_3 b_2 + (\alpha\gamma - \beta\lambda_2) b_3}{\lambda_1\lambda_2\lambda_3}$$

$$x_2 = \frac{\begin{vmatrix} \lambda_1 & b_1 & \beta \\ 0 & b_2 & \gamma \\ 0 & b_3 & \lambda_3 \end{vmatrix}}{\lambda_1\lambda_2\lambda_3} = \frac{\lambda_1\lambda_3 b_2 - \lambda_1\gamma b_3}{\lambda_1\lambda_2\lambda_3} = \frac{\lambda_3 b_2 - \gamma b_3}{\lambda_2\lambda_3}$$

$$x_2 = \frac{\begin{vmatrix} \lambda_1 & \alpha & b_1 \\ 0 & \lambda_2 & b_2 \\ 0 & 0 & b_3 \end{vmatrix}}{\lambda_1\lambda_2\lambda_3} = \frac{\lambda_1\lambda_2 b_3}{\lambda_1\lambda_2\lambda_3} = \frac{b_3}{\lambda_3}.$$

6.7 Laplacescher Entwicklungssatz

Der Laplacesche Entwicklungssatz ist vielleicht die wichtigste Methode, eine Determinante auszurechnen: man ersetzt eine $n \times n$-Determinante durch n $(n-1) \times (n-1)$-Determinanten.

Es sei $A = (a_{ij}) = (a_1, \dots, a_n) \in \operatorname{Mat}_{n,n}(\mathbb{K}), n \geq 2$. Für jedes $1 \leq i, j \leq n$ betrachtet man die *Streichungsmatrix*

$$A'_{ij} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix} \in \operatorname{Mat}_{n-1,n-1}(\mathbb{K}),$$

welche aus A durch Streichung der i-ten Zeile und der j-ten Spalte entsteht.

Satz 6.7.1 (Laplacescher Entwicklungssatz)**.** *(1) Entwicklung nach der j-ten Spalte ($j = 1, 2, \dots, n$ fest):*

$$\operatorname{Det}(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \operatorname{Det}(A'_{ij}).$$

(2) Entwicklung nach der i-ten Zeile ($i = 1, \dots, n$ fest):

$$\operatorname{Det}(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \operatorname{Det}(A'_{ij}).$$

Beweis. Wegen der Invarianz unter Transposition, also $\operatorname{Det}(A^\top) = \operatorname{Det}(A)$, folgt (2) aus (1). Weil das Vertauschen von Spalten nur ein Vorzeichen einträgt, also $\operatorname{Det}(AP_{ij}) = -\operatorname{Det}(A)$, brauchen wir (1) nur für irgendeine Spalte (sagen wir die n-te) zu zeigen.
Wir sortieren in der Leibniz-Formel

$$\operatorname{Det}(A) = \sum_{\pi \in \mathfrak{S}_n} \operatorname{sign}(\pi) \underbrace{a_{\pi(1),1} \cdots a_{\pi(n-1),n-1} a_{\pi(n),n}}_{=:A[\pi]}$$

die Permutationen π in $\mathfrak{S}_n$ nach ihrem Wert auf n, also $\pi(n) = i$; das ergibt n Teil*mengen*

$$\mathfrak{S}_n(i) := \{\pi \in \mathfrak{S}_n \mid \pi(n) = i\} \qquad (i = 1, \dots, n)$$

und damit n Teilsummen $(i = 1, \dots, n)$

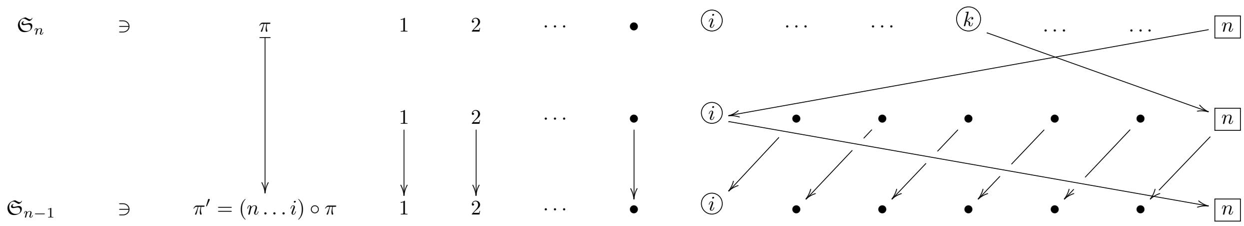
$$\Sigma(i) : = \sum_{\pi \in \mathfrak{S}_n(i)} \operatorname{sign}(\pi) a_{\pi(1),1} \cdots a_{\pi(n-1),n-1} \tag{6.2}$$

$$= a_{i,n} \sum_{\pi \in \mathfrak{S}_n(i)} \operatorname{sign}(\pi) a_{\pi(1),1} \cdots a_{\pi(n-1),n-1}, \tag{6.3}$$

worin der letzte Faktor $a_{\pi(n),n} = a_{i,n}$ nicht mehr von π abhängt und deshalb vor die Summe gezogen werden kann.
Durch die Bijektion

$$\begin{array}{ccl} \mathfrak{S}_n(i) & \longrightarrow & \mathfrak{S}_n(n) \leq \mathfrak{S}_n \\ \pi & \longmapsto & \pi' = \begin{pmatrix} n & n-1 & \dots & i+1 & i \end{pmatrix} \circ \pi \end{array}$$

auf die Unter*gruppe* $\mathfrak{S}_n(n)$:



Wir können $\mathfrak{S}_n(n)$ mit $\mathfrak{S}_{n-1} \leq \mathfrak{S}_n$ identifizieren.
Das Signum hat sich aber geändert:

$$\operatorname{sign}(\pi') = (-1)^{n-i} \operatorname{sign}(\pi) = (-1)^{i+n} \operatorname{sign}(\pi).$$

Die Indizes der Zeilen haben sich auch geändert, aber die Permutation $\begin{pmatrix} n & \dots & i \end{pmatrix}$ entspricht genau der neuen Nummerierung der Zeilen in A'_{in}:

$$a_{in} A'_{ij}[\pi'] = A[\pi] \text{ für } \pi \in \mathfrak{S}_n(i).$$

Also haben wir als Teilsumme:

$$\Sigma(i) = a_{in} \operatorname{Det}(A'_{ij}).$$

□

Beispiel 6.7.2.

$$A = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 1 & 2 \\ 4 & -2 & -3 \end{pmatrix}, \quad \operatorname{Det}(A) = 15$$

Entwicklung nach der ersten Spalte:

$$\operatorname{Det}(A) = 3 \operatorname{Det} \begin{pmatrix} 1 & 2 \\ -2 & -3 \end{pmatrix} - 0 \operatorname{Det} \begin{pmatrix} 2 & 1 \\ -2 & -3 \end{pmatrix} + 4 \operatorname{Det} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$
$$= 3 \cdot (+1) - 0 \cdot (-4) + 4 \cdot 3 = 15.$$

6.8 Adjunkte einer Matrix

Für ein $A = (a_{ij}) = (a_1, \dots, a_n) \in \operatorname{Mat}_{n,n}(\mathbb{K})$ betrachten wir nun die Determinanten der Streichungsmatrizen, also

$$a^{\#}_{ij} := (-1)^{i+j} \operatorname{Det}(A'_{ji}),$$

wobei wir die Indizes vertauscht haben und ein Vorzeichen wie im Laplaceschen Entwicklungssatz gesetzt haben.

Definition 6.8.1. Die Zahl $a_{ij}^{\#}$ heißt *Kofaktor* von A an der Stelle (i, j). Die Matrix der Kofaktoren

$$A^{\#} := (a_{ij}^{\#} = ((-1)^{i+j} \operatorname{Det}(A'_{ij}))$$

heißt *Adjunkte*[1] von A. Im Falle $n = 1$ setzt man für $A = (a_{11})$ dann $a_{11}^{\#} = 1, A^{\#} = (1) = \mathbb{1}$.

Lemma 6.8.2.

$$a_{ij}^{\#} = \operatorname{Det}(a_1, \ldots, a_{i-1}, e_j, a_{i+1}, \ldots, a_j, \ldots, a_n).$$

Beweis. Man entwickle die Determinante $\operatorname{Det}(a_1, \ldots, a_{i-1}, ej, a_{i+1}, \ldots, a_j, \ldots, a_n)$ mit Laplace nach der i-ten Spalte. □

Wir haben eine Funktion

$$()^{\#} : \operatorname{Mat}_{n,n}(\mathbb{K}) \to \operatorname{Mat}_{n,n}(\mathbb{K}).$$

Sie ist nicht linear.

Lemma 6.8.3. *Für die Adjunkte gelten folgende Rechenregeln:*

1) $\mathbb{1}^{\#} = \mathbb{1}$, $0^{\#} = 0$.

2) $(\lambda A)^{\#} = \lambda^{n-1} A^{\#}$.

3) $(A \cdot B)^{\#} = B^{\#} \cdot A^{\#}$.

4) $(A^{-1})^{\#} = (A^{\#})^{-1}$ *für invertierbares* A.

5) $(A^{\top})^{\#} = (A^{\#})^{\top}$.

Lemma 6.8.4. *(i)* $(A^{\#})^{\#} = \operatorname{Det}(A)^{n-2} A$, *für* $n \geq 2$.

(ii) $\operatorname{Det}(A^{\#}) = \operatorname{Det}(A)^{n-1}$, *für* $n \geq 1$.

Beweis. Übung. □

Wir kommen nun zu folgenden Entwicklungssätzen.

Satz 6.8.5.

$$A^{\#} \cdot A = A \cdot A^{\#} = \operatorname{Det}(A) \cdot \mathbb{1}.$$

Beweis. Wir schreiben die Matrizengleichung

$$B := A^{\#} \cdot A = (b_{ij})$$

für jede Stelle (i, j) aus:

$$\begin{aligned} b_{ij} = \sum_{k=1}^{n} a_{ik}^{\#} a_{kj} &= \sum_{k=1}^{n} a_{kj} \operatorname{Det}(a_1, \ldots, a_{i-1}, e_j, a_i, \ldots, a_j, \ldots, a_n) \\ &= \operatorname{Det}(a_1, \ldots, a_{i-1}, \underbrace{\sum_{k=1}^{n} a_{kj} e_j}_{=a_j}, a_{i+1}, \ldots, a_j, \ldots, a_n) \\ &= \operatorname{Det}(a_1, \ldots, a_{i-1}, \underline{a_j}, a_{i+1}, \ldots, \underline{a_j}, \ldots, a_n) \end{aligned}$$

wegen der Linearität der Determinante. Aber diese letzte Determinante ist null für $i \neq j$, und sie ist $= \operatorname{Det}(A)$ für $i = j$. Also ist $B = \operatorname{Det}(A)\mathbb{1}$.
Die zweite Gleichung beweist man entsprechend. □

Bemerkung 6.8.6. Die Gleichung $A \cdot A^{\#} = \operatorname{Det}(A) \cdot \mathbb{1}$ an der Stelle $i = j$ bzw. aus $A^{\#} \cdot A = \operatorname{Det}(A) \cdot \mathbb{1}$ an der Stelle $j = i$ ist der Lapacesche Entwicklungssatz.

[1] Nicht zu verwechseln mit der später noch kommenden *Adjungierten*.

Berechnung der Inversen einer Matrix

Satz 6.8.7. *Für ein invertierbares A gilt*

$$A^{-1} = \operatorname{Det}(A)^{-1} \cdot A^{\#}.$$

Beweis. Folgt sofort aus der Gleichung $A^{\#} \cdot A = \operatorname{Det}(A) \cdot \mathbb{1}$. □

Beispiel 6.8.8.

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad D = ad - bc \neq 0$$

$$\begin{array}{lll} A'_{11} = d & a^{\#}_{11} = +d & \\ A'_{12} = c & a^{\#}_{12} = -b & A^{-1} = \dfrac{1}{D}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \\ A'_{21} = b & a^{\#}_{21} = -c & \\ A_{22} = a & a^{\#}_{22} = +a & \end{array}$$

6.9 Spezielle lineare Gruppe

Nach dem Produktsatz für Determinanten ist

$$\begin{array}{rccc} \operatorname{Det}: & \operatorname{GL}_n(\mathbb{K}) & \longrightarrow & \mathbb{K}^\times = \mathbb{K} \setminus \{0\} \\ & \cap\!\!\mid & & \cap\!\!\mid \\ & \operatorname{Mat}_{n,n}(\mathbb{K}) & \longrightarrow & \mathbb{K} \end{array}$$

nach Einschränkung auf die allgemeine lineare Gruppe $\operatorname{GL}_n(\mathbb{K})$ bzw. die multiplikative Untergruppe $\mathbb{K}^\times$ ein Homomorphismus; beide Abbildungen sind surjektiv.

Definition 6.9.1. Die Untergruppe

$$\operatorname{SL}_n(\mathbb{K}) = \ker(\operatorname{Det}) = \{A \in \operatorname{GL}_n(\mathbb{K}) \mid \operatorname{Det}(A) = 1\}$$

heißt (n-te) *spezielle lineare Gruppe* über $\mathbb{K}$. Und

$$\operatorname{SL}_{\mathbb{K}}(V) = \{f \in \operatorname{Aut}_{\mathbb{K}}(V) \mid \operatorname{Det}(f) = 1\}$$

heißt *spezielle lineare Gruppe von V* für endlich-dimensionale V.

Bemerkung 6.9.2. 1) $\operatorname{SL}_n(\mathbb{K})$ ist normal in $\operatorname{GL}_n(\mathbb{K})$ und $\operatorname{GL}_n(\mathbb{K})/\operatorname{SL}_n(\mathbb{K}) \cong \mathbb{K}^\times$. Hier ist $\mathbb{K}^\times$ über den Schnitt

$$s: \mathbb{K}^\times \longrightarrow \operatorname{GL}_n(\mathbb{K}), \quad s(\lambda) = \begin{pmatrix} \lambda & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} = D[\lambda, 1, \ldots, 1]$$

nicht nur Quotientengruppe, sondern auch Untergruppe von $\operatorname{GL}_n(\mathbb{K})$.

2) Isomorphismen:

$$\begin{array}{ccc} \operatorname{GL}(V) & \underset{L_{\mathfrak{B}\mathfrak{B}}}{\overset{M_{\mathfrak{B}\mathfrak{B}}}{\rightleftarrows}} & \operatorname{GL}_n(\mathbb{K}) \\ \cup\!\!\mid & \cong & \cup\!\!\mid \\ \operatorname{SL}(V) & \underset{\cong}{\rightleftarrows} & \operatorname{SL}_n(\mathbb{K}) \end{array}$$

Beispiel 6.9.3. 1) *Drehungen*

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}, \quad 0 \leq \alpha \leq 2\pi.$$

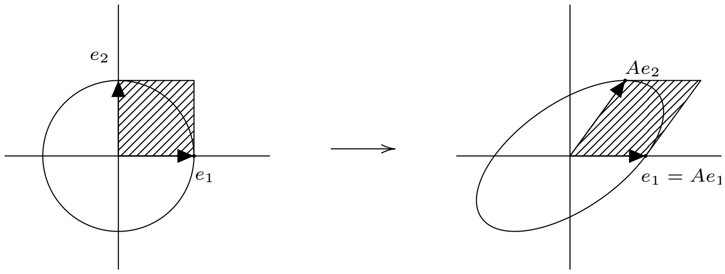
$\operatorname{Det}(A) = \cos^2\alpha + \sin^2\alpha = 1$, also $A \in \operatorname{SL}_2(\mathbb{R})$.

2) *Elementarmatrizen* $E_{ij}(\lambda)$

$$\begin{aligned} \operatorname{Det}(E_{ij}(\lambda)) &= 1 && \text{für } i \neq j, \\ \operatorname{Det}(E_{ij}(\lambda)) &= 1 && \text{für } \lambda = 1. \end{aligned}$$

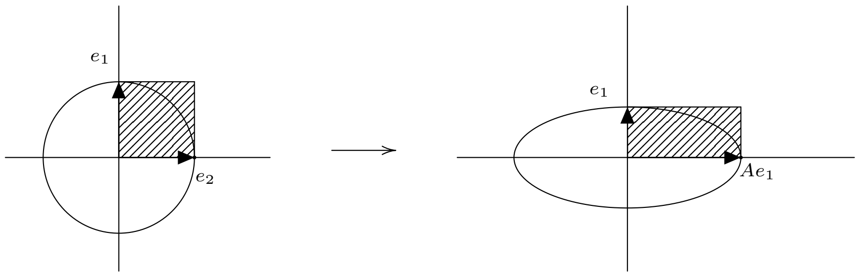
3) *Scherungen*(Transvektionen)

$$A = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}, \qquad \operatorname{Det}(A) = 1$$



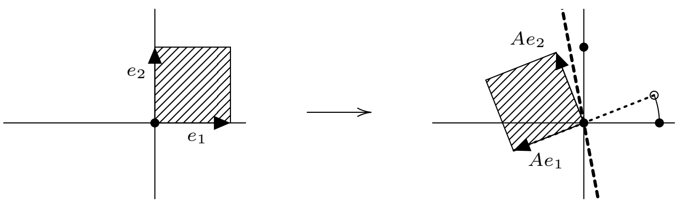
4) *Streckungen* $\lambda \neq 0$

$$A = \begin{pmatrix} \lambda & 0 \\ 0 & \frac{1}{\lambda} \end{pmatrix} = E_{11}(\lambda) E_{22} \left(\frac{1}{\lambda} \right) \qquad \operatorname{Det}(A) = 1$$



5) *Nicht-Beispiel*: Spiegelung

$$A = \begin{pmatrix} -\cos\alpha & -\sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \operatorname{Det}(A) = -1$$



6.10 Orientierung reeller Vektorräume

Für den Körper $\mathbb{K} = \mathbb{R}$ haben wir die besondere Situation, dass $\mathbb{R}^\times = \mathbb{R} \setminus \{0\}$ aus zwei Zusammenhangskomponenten besteht (im Gegensatz zu $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$):

$$\mathrm{GL}_{(}\,\mathbb{R}) \xrightarrow{\operatorname{Det}} \mathbb{R} \setminus 0 = \mathbb{R}_{>0} \sqcup \mathbb{R}_{<0}.$$

Hier ist $\mathbb{R}_{>0}$ eine Untergruppe von $\mathbb{R}^\times$, während $\mathbb{R}_{<0}$ eine Nebenklasse $(-1)\mathbb{R}_{>0} = \mathbb{R}_{<0}$ ist.

Definition 6.10.1. Die Untergruppe

$$\begin{array}{rlll} & \mathrm{GL}_n^+(\mathbb{R}) & = \mathrm{Det}^{-1}(\mathbb{R}_{>0}) & = \{A \in \mathrm{GL}_n(\mathbb{R}) \mid \mathrm{Det}(A) > 0\} \\ \text{bzw.} & \mathrm{GL}(V) & = \mathrm{Det}^{-1}(\mathbb{R}_{>0}) & = \{f \in \mathrm{Aut}(V) \mid \mathrm{Det}(f) > 0\} \end{array}$$

heißen *positive lineare Gruppe* über $\mathbb{R}$ bzw. von V.

Bemerkung 6.10.2. (1) Da $\mathbb{M} = \mathrm{Mat}_{n,n}(\mathbb{R}) = \mathbb{R}^{n^2}$ ist, wissen wir, was eine offene/abgeschlossene Teilmenge von $\mathrm{Mat}_{n,n}(\mathbb{R})$ ist, was eine stetige Kurve $A(t)$ von Matrizen ist und was eine stetige Funktion auf $\mathrm{Mat}_{n,n}(\mathbb{R})$ ist.

(2) $\mathrm{Det} : \mathrm{Mat}_{n,n}(\mathbb{R}) \to \mathbb{R}$ ist stetig, also ist $\mathrm{GL}_n(\mathbb{R}) = \mathrm{Det}^{-1}(\mathbb{R}_{>0} \sqcup \mathbb{R}_{<0})$ offen in $\mathrm{Mat}_{n,n}(\mathbb{R})$.

(3) $\mathrm{GL}_n(\mathbb{R})$ besteht aus zwei disjunkten, offenen Teilen:

- $\mathrm{GL}_n^+(\mathbb{R}) = \mathrm{Det}^{-1}(]0, \infty[)$ ist Untergruppe
- $\mathrm{GL}_n^-(\mathbb{R}) = \mathrm{Det}^{-1}(]-\infty, 0[)$ ist eine Nebenklasse, aber keine Untergruppe.

$$\mathrm{GL}_n^-(\mathbb{R}) = S \cdot \mathrm{GL}_n^+(\mathbb{R}) = \{S \cdot A \mid A \in \mathrm{GL}_n^+(\mathbb{R})\},$$

  $S = D(-1, 1, \ldots, 1)$ eine Spiegelung.

(4) Sei $A^+ \in \mathrm{GL}_n^+(\mathbb{R}), A^- \in \mathrm{GL}_n^-(\mathbb{R})$. Es gibt keine stetige Kurve (Schar, Weg)

$$A : [0, 1] \longrightarrow \mathrm{GL}_n(\mathbb{R}) \text{ mit } A(0) = A^+, A(1) = A^-.$$

Weil die Determinante stetig ist (denn sie ist ein Polynome vom Grad n^2 Einträgen der Matrix), hätten wir eine stetige Funktion

$$\varphi = \mathrm{Det} \circ A : [0, 1] \longrightarrow \mathrm{GL}_n \longrightarrow \mathbb{R} - 0$$

mit $\varphi(0) > 0$ und $\varphi(1) < 0$; diese müsste nach dem Zwischenwertsatz eine Nullstelle haben. (Man sagt: $\mathrm{GL}_n(\mathbb{R})$ ist nicht wegzusammenhängend.)

(5) Beide Teile $\mathrm{GL}_n^+(\mathbb{R})$ und $\mathrm{GL}_n^-(\mathbb{R})$ sind wegzusammenhängend: wir zeigen, dass jedes $A \in \mathrm{GL}_n^+(\mathbb{R})$ bzw. $A \in \mathrm{GL}_n^-(\mathbb{R})$ stetig nach $S = \mathbb{1} = D(1, \ldots, 1)$ bzw. nach $S = D(-1, 1, \ldots, 1)$ bewegt werden kann. Dazu sei $A = E_{i_1 j_1}(\lambda_1) \cdot \ldots \cdot E_{i_s j_s}(\lambda_s)$, $\lambda_1, \ldots, \lambda_s \neq 0$, eine Produktdarstellung von A durch Elementarmatrizen.

- Für ein $E_{ij}(\lambda)$ mit $i \neq j$ nehmen wir den Weg

$$\varepsilon_{ij}^{\lambda}(t) := E_{ij}((1-t)\lambda);$$

  er verläuft ganz in $\mathrm{GL}_n^+ \mathbb{R}$ wegen $\mathrm{Det}(\varepsilon_{ij}^{\lambda}(t)) = +1$ und von $\varepsilon_{ij}^{\lambda}(0) = E_{ij}(\lambda)$ nach $\varepsilon_{ij}^{\lambda}(1) = E_{ij}(0) = \mathbb{1}$.
- Für ein $E_{ii}(\lambda)$ (also $i = j$) müssen wir vorsichtiger wählen: ist $\lambda > 0$, so setzen wir

$$\tilde{\varepsilon}_{ii}^{\lambda}(t) := E_{ii}(t + (1-t)\lambda),$$

  und ist $\lambda < 0$, setzen wir

$$\tilde{\varepsilon}_{ii}^{\lambda}(t) := E_{ii}(-t + (1-t)\lambda).$$

  Der Weg $\tilde{\varepsilon}_{ii}^{\lambda}(t)$ verläuft ganz in $\mathrm{GL}_n^+(\mathbb{R})$ für $\lambda > 0$, und ganz in $\mathrm{GL}_n^-(\mathbb{R})$ für $\lambda < 0$; und er verläuft von $\tilde{\varepsilon}_{ii}^{\lambda}(0) = E_{ii}(\lambda)$ nach $\tilde{\varepsilon}_{ii}^{\lambda}(1) = E_{ii}(+1) = D(1, 1, \ldots, 1) = \mathbb{1}$ für $\lambda > 0$ und für $\lambda < 0$ von $\tilde{\varepsilon}_{ii}^{\lambda}(0) = E_{ii}(\lambda)$ nach $\tilde{\varepsilon}_{ii}^{\lambda}(1) = D(1, \ldots, 1, -1, 1, \ldots, 1)$.
  Nun müssen wir - nur im Falle $i = j$ und $\lambda < 0$ - noch Drehungen $\delta_{12}(t) \cdot \delta_{23}(t) \cdot \ldots \cdot \delta_{i-1,i}(t)$ nachschalten (um die -1 an die erste Stelle wie in S zu schaffen): Wir setzen also

$$\delta_{k-1,k}(t) := \begin{array}{rc} & \begin{array}{cccccccc} & & & k-1 & k & & & \end{array} \\ \begin{array}{r} \\ \\ \\ k-1 \\ k \\ \\ \\ \\ \end{array} & \left(\begin{array}{ccc|cc|ccc} 1 & & & & & & & \\ & \ddots & & & 0 & & 0 & \\ & & 1 & & & & & \\ \hline & 0 & & \cos(\pi t) & -\sin(\pi t) & & 0 & \\ & & & \sin(\pi t) & \cos(\pi t) & & & \\ \hline & & & & & 1 & & \\ & 0 & & & 0 & & \ddots & \\ & & & & & & & 1 \end{array}\right) \end{array}$$

und dann

$$\varepsilon_{ii}^{\lambda}(t) := \tilde{\varepsilon}_{ii}^{\lambda}(t) \quad \text{für } \lambda > 0$$
$$\varepsilon_{ii}^{\lambda}(t) := \delta_{1,2}(t) \cdots \delta_{i-1,i}(t)\tilde{\varepsilon}_{ii}^{\lambda}(t) \quad \text{für } \lambda < 0$$

Weil die Drehmatrizen alle $\mathrm{Det}(\delta_{k-1,k}(t)) = +1$ haben, ist $\varepsilon_{ii}^{\lambda}(t)$ ein Weg in $\mathrm{GL}_n^-(\mathbb{R})$ und geht von $\varepsilon_{ii}^{\lambda}(0) = E_{ii}(\lambda)$ nach $\varepsilon_{ii}^{\lambda}(t) = D(-1, 1, \ldots, 1) = S$.
Nach dieser Vorbereitung definieren wir unseren Weg als

$$A(t) = \varepsilon_{i_1 j_1}^{\lambda_1}(t) \cdot \varepsilon_{i_2 j_2}^{\lambda_2}(t) \cdot \ldots \cdot \varepsilon_{i_s j_s}^{\lambda_s}(t), \quad 0 \leq t \leq 1.$$

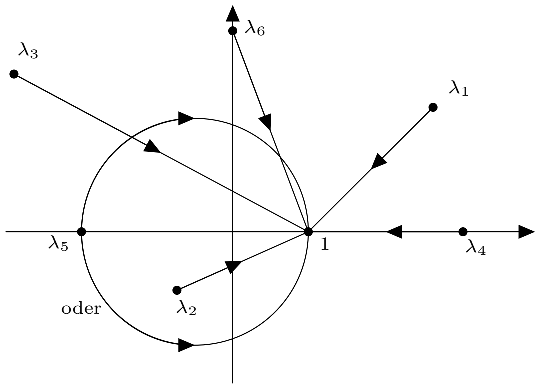
Das ist ein stetiger Weg von $A(0) = A$ nach $A(1)$. Um dieses zu bestimmen, erinnern wir uns:

$$\mathrm{Det}(A(t)) = \prod_{\substack{i_k = j_k \\ \lambda_k > 0}} (t + (1-t)\lambda_k) \prod_{\substack{i_k = j_k \\ \lambda_k < 0}} (-t + (1-t)\lambda_k)\,.$$

$\mathrm{Det}(A(t))$ ist also (für alle t) positiv (negativ), wenn die Anzahl der $\lambda_k < 0$ mit $i_k = j_k$ gerade (ungerade) ist. Also ist

$$D(A(1)) = S^l = \begin{cases} \mathbb{1}, & l \text{ gerade} \\ S, & l \text{ ungerade.} \end{cases}$$

(6) Weil $\mathbb{C}^\times = \mathbb{C} \setminus 0$ wegzusammenhängend ist, kannn man wie in (5) zeigen, dass ganz $\mathrm{GL}_n(\mathbb{C})$ wegzusammenhängend ist: man muss die Wege $t + (1-t)\lambda$ von λ (für $t = 0$) nach $+1$ (für $t = 1$) nur geschickter wählen, wenn λ (wie z.B. für λ_5 im Bild) auf der negativen Halbachse liegt:



Orientierung

Definition 6.10.3. Es sei V ein reeller Vektorraum der Dimension n. Auf der Menge $\mathbb{B}(V)$ aller geordneter Basen $\mathfrak{B} = (b_1, \ldots, b_n)$ von V führen wir eine Äquivalenzrelation ein:

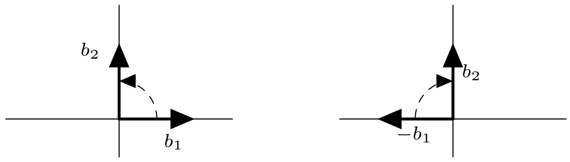
$$\mathfrak{B} \sim \mathfrak{B}' \iff \mathrm{Det}(\Omega_{\mathfrak{B}'\mathfrak{B}}) = \mathrm{Det}(M_{\mathfrak{B}'\mathfrak{B}}(\mathrm{id}_V)) = \mathrm{Det}(\Phi_{\mathfrak{B}'\mathfrak{B}}) > 0.$$

Hier ist $\Omega_{\mathfrak{B}'\mathfrak{B}}$ also die Basiswechselmatrix; und $\Phi = \Phi_{\mathfrak{B}'\mathfrak{B}} : V \to V$ ist definiert durch $\Phi(b_i) = b_i'$ $(i = 1, \ldots, n)$.
Wir nennen äquivalente Basen *gleichorientiert*; eine Äquivalenzklasse nennen wir eine *Orientierung* von V.

Es gibt für V genau zwei Orientierungen: ist $\mathfrak{B} = (b_1, \ldots, b_n)$ ein Repräsentant der einen Orientierung, so ist $\mathfrak{B}' = (-b_1, b_2, \ldots, b_n)$ ein Repräsentant der anderen.

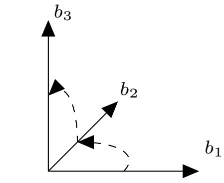
Beispiel 6.10.4. 1) $V = \mathbb{R}$

2) $V = \mathbb{R}^2$



„positiv“ „negativ“

3) $V = \mathbb{R}^3$ (*Rechte-Hand-Regel*)



„positiv“

6.11 Volumenmessung im $\mathbb{R}^n$

Eine wichtige Anwendung der Determinante ist die Volumenmessung in reellen Vektorräumen.

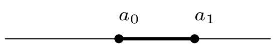
Definition 6.11.1. Es seien $n+1$ Punkte (als Ortsvektoren) $a_0, a_1, \ldots, a_n \in \mathbb{R}^n$ gegeben. Als ihre *konvexe Hülle* oder *Spat* bezeichnet man die Menge

$$\operatorname{Spat}(a_0, a_1, \ldots, a_n) = \left\{ x = \sum_{i=1}^{n} \lambda_i a_i \mid 0 \leq \lambda_i \leq 1, \lambda_0 + \ldots + \lambda_n = 1 \right\}.$$

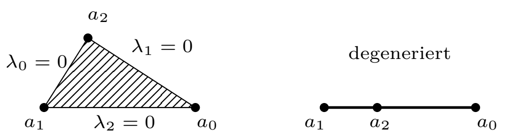
Die $\lambda_0, \ldots, \lambda_n$ nennt man die *baryzentrischen Koordinaten* von x in Bezug auf $a_0, \ldots, a_n$.

Beispiel 6.11.2.

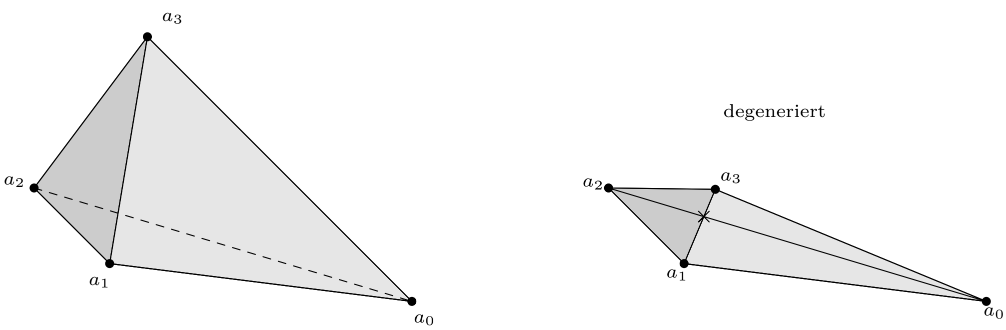
- $n = 1$



- $n = 2$



- $n = 3$

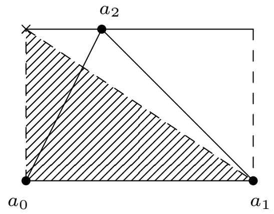


Definition 6.11.3. Das *Volumen* eines von $a_0, a_1, \ldots, a_n \in \mathbb{R}^n$ aufgespannten Spats ist

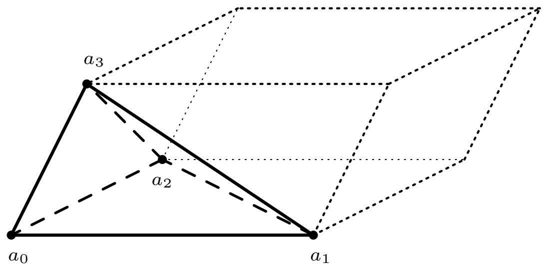
$$\operatorname{vol}(a_0, a_1, \ldots, a_n) := \frac{1}{n!} \left| \operatorname{Det}(a_1 - a_0, \ldots, a_n - a_0) \right|$$

Beispiel 6.11.4.

- $n = 2$



- $n = 3$



- Scherungsinvarianz: *Prinzip des Cavalieri.*
- $\mathrm{vol}(0, e_1, \dots, e_n) = \frac{1}{n!}$ für die Standardeinheitsvektoren e_i.

Lemma 6.11.5. *Ist* $\mathrm{Spat}(a_0, a_1, \dots, a_n)$ *in einem affinen Unterraum* $\mathcal{A}$ *der Dimension* $d < n$ *enthalten, so ist* $\mathrm{vol}(a_0, a_1, \dots, a_n) = 0$.

Beweis. In diesem „degenerierten“ Fall sind die Vektoren $a_1 - a_0, \dots, a_n - a_0$ nicht linear unabhängig; also ist die Determinante $= 0$. □

Lemma 6.11.6 (Translationsinvarianz)**.** *Für alle* $b \in \mathbb{R}^n$ *gilt*

$$\mathrm{vol}(b + a_0, b + a_1, \dots, b + a_n) = \mathrm{vol}(a_0, a_1, \dots, a_n).$$

Lemma 6.11.7 (Streckungsmultiplikativität)**.**

$$\mathrm{vol}(a_0, \dots, \lambda a_i, \dots, a_n) = |\lambda| \,\mathrm{vol}(a_0, \dots, a_i, \dots, a_n),$$
$$\mathrm{vol}(\lambda a_0, \dots, \lambda a_n) = |\lambda|^n \,\mathrm{vol}(a_0, \dots, a_n).$$

Lemma 6.11.8 (Permutationsinvarianz)**.** *Für jedes* $\sigma \in \mathfrak{S}_{n+1} = \mathrm{Sym}(\{0, 1, \dots, n\})$*:*

$$\mathrm{vol}(a_{\sigma(0)}, a_{\sigma(1)}, \dots, a_{\sigma(n)}) = \mathrm{vol}(a_0, a_1, \dots, a_n).$$

Satz 6.11.9. *Für eine affine Abbildung* $\varphi = f + b : \mathbb{R}^n \to \mathbb{R}^n$ *mit linearem Anteil* $f \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R}^n)$ *und Translationsanteil* $b \in \mathbb{R}^n$ *gilt:*

$$\mathrm{vol}(\varphi(a_0), \dots, \varphi(a_n)) = |\operatorname{Det}(f)| \,\mathrm{vol}(a_0, \dots, a_n)$$

Beweis. Nach Lemma 6.11.6 brauchen wir nur $\varphi = f$ zu betrachten. Dann ist

$$\begin{aligned}
\mathrm{vol}(f(a_0), f(a_1), \dots, f(a_n)) &= \frac{1}{n!} |\operatorname{Det}(f(a_1) - f(a_0), \dots, f(a_n) - f(a_0))| \\
&= \frac{1}{n!} |\operatorname{Det}(f(a_1 - a_0), f(a_2 - a_0), \dots, f(a_n - a_0))| \\
&= \frac{1}{n!} |\operatorname{Det}(f) \operatorname{Det}(a_1 - a_0, a_2 - a_0, \dots, a_n - a_0)| \\
&= |\operatorname{Det}(f)| \,\mathrm{vol}(a_0, a_1, \dots, a_n).
\end{aligned}$$

□

Volumenmessung in einem reellen Vektorraum

Man kann in einem reellen Vektorraum V keine absolute Volumenmessung durchführen, sondern nur relativ zu einer Basis $\mathfrak{B} = (b_1, \dots, b_n)$:

$$\mathrm{vol}_{\mathfrak{B}}(a_0, a_1, \dots, a_n) := \frac{1}{n!} |\operatorname{Det}(f)|$$

für $a_0, a_1, \dots, a_n \in V$ und $f : V \to V$ ist bestimmt durch $f(b_i) := a_i - a_0$ für $i = 1, \dots, n$.

Beispiel 6.11.10. 1) $V = \mathbb{R}^n, \mathcal{B} = (e_1, \dots, e_n)$ Standardbasis: $\mathrm{vol}_{\mathcal{B}} = \mathrm{vol}$.

2) V unnd $\mathcal{B}$ beliebig: $\mathrm{vol}_{\mathcal{B}}(0, b_1, \dots, b_n) = \frac{1}{n!}$.

3) V unnd $\mathcal{B}$ beliebig:

$$\begin{aligned}
&\mathcal{B}' = (b_1, \dots, \lambda b_i, \dots, b_n): && \mathrm{vol}_{\mathcal{B}'} = \lambda \, \mathrm{vol}_{\mathcal{B}} \\
&\mathcal{B}' = (b_1, \dots, b_i + b_j, \dots, b_j, \dots, b_n): && \mathrm{vol}_{\mathcal{B}'} = \mathrm{vol}_{\mathcal{B}} \\
&\mathcal{B}' = (b_1, \dots, b_{i-1}, b_j, \dots, b_{j-1}, b_i, \dots, b_n): && \mathrm{vol}_{\mathcal{B}'} = \mathrm{vol}_{\mathcal{B}}
\end{aligned}$$

Satz 6.11.11. *Sei V ein reeller Vektorraum und $\mathcal{B} = (b_1, \dots, b_n)$ eine Basis.*

(i) $\mathrm{vol}_{\mathcal{B}}(a_0, a_1, \dots, a_n) = 0$ *genau dann, wenn* $\dim(\mathrm{Span}(a_1 - a_0, \dots, a_n - a_0)) < n$.

(ii) $\mathrm{vol}_{\mathcal{B}}(\varphi(a_0), \varphi(a_1), \dots, \varphi(a_n)) = |\operatorname{Det}(f)| \, \mathrm{vol}_{\mathcal{B}}(a_0, \dots, a_n)$ *für jede affine Abbidung* $\varphi = f + b$, $f : V \to V$ *linear und* $b \in V$.

(iii) $\mathrm{vol}_{\mathcal{B}'} = |\operatorname{Det}(\Omega_{\mathcal{B}'\mathcal{B}})| \, \mathrm{vol}_{\mathcal{B}}$ *für zwei Basen $\mathcal{B}$ und $\mathcal{B}'$, wobei $\Omega_{\mathcal{B}'\mathcal{B}} = M_{\mathcal{B}'\mathcal{B}}(\mathrm{id}_V)$ die Basiswechselmatrix ist.*

Beweis. Folgt aus Satz 6.11.9 □

6.12 Unterdeterminanten

Die Determinante einer $n \times n$-Matrix $A = (a_{ij})$ gibt uns nur Auskunft, ob der Rang maximal ist oder nicht:

$$\mathrm{rg}(A) = n \iff \operatorname{Det}(A) \neq 0.$$

Den Rang selbst können wir aber durch sogenannte Unterdeterminanten ermitteln. Die Streichungsmatrizen A'_{ij} im Laplaceschen Entwicklungssatz führten zu den Kofaktoren

$$a^{\#}_{ij} = (-1)^{i+j} \operatorname{Det}(A'_{ji}).$$

Diese Kofaktoren sind Beispiele für Unterdeterminanten. Allgemeiner sei $I \subseteq \underline{n} = \{1, \dots, n\}$ eine Teilmenge mit $|I| = p$ und $I' = \underline{n} \setminus I$. Sei $J \subseteq \underline{m}$ und ebenfalls $|J| = p$. Hier kann $0 \leq p \leq n, m$ sein. Dann definieren wir für eine $n \times m$-Matrix A die *Untermatrix*

$$A_{I,J} = \begin{array}{l} \text{nur die Zeilen } i \text{ mit } i \in I, \\ \text{und nur die Spalten} j \text{ mit } j \in J. \end{array}$$

Ihre Determinante $\operatorname{Det}(A_{I,J})$ nennt man eine *p-reihige Unterdeterminante.* Als *Streichungsmatrizen* bezeichnen wir Untermatrizen

$$A'_{I,J} = A_{I',J'}.$$

Für ein Paar (I, J) definiert man als Signum

$$\mathrm{sign}(I, J) := (-1)^s \text{ mit } s = \sum_{k=1}^{p} (i_k - 1) + \sum_{k=1}^{p} (j_k - 1) = \sum_{k=1}^{p} (i_k + j_k),$$

wenn $I = \{i_1, \dots, i_p\}, J = \{j_1, \dots, j_p\}$; denn man braucht s Zeilen- bzw. Spaltenvertauschungen, um die Matrix $(a_1, \dots, a_{i_1 - 1}, e_{j_1}, \dots, a_{i_2 - 1}, e_{j_2}, \dots, a_n)$ in die Blockmatrix

$$\left(\begin{array}{c|c} A'_{I,J} & 0 \\ \hline 0 & * \end{array} \right)$$

zu verwandeln.
Man hat nun einen verallgemeinerten

Satz 6.12.1 (Allgemeiner Satz von Laplace)**.** *Es sei A eine quadratische $n \times n$-Matrix.*

(i) Für festes $I \subseteq \underline{n}$ mit $|I| = p$ gilt:

$$\operatorname{Det}(A) = \sum_{\substack{K \subseteq \underline{n} \\ |K| = p}} \operatorname{sign}(I, K) \operatorname{Det}(A_{I,K}) \operatorname{Det}(A'_{I,K}).$$

(ii) Für festes $J \subseteq \underline{n}$ mit $|J| = p$ gilt:

$$\operatorname{Det}(A) = \sum_{\substack{K \subseteq \underline{n} \\ |K| = p}} \operatorname{sign}(K, J) \operatorname{Det}(A_{K,J}) \operatorname{Det}(A'_{K,J}).$$

Bemerkung 6.12.2. Für $p = 1$ erhalten wir den einfachen Satz von Laplace.

Für eine quadratische Matrix A folgt natürlich sofort: für jedes $p = 1, 2, \ldots, n = \operatorname{rg}(A)$ gibt es mindestens eine p-reihige Unterdeterminante $\operatorname{Det}(A_{I,J})$, die nicht verschwindet. Es gilt allgemein:

Satz 6.12.3. *Für eine $m \times n$-Matrix A sind äquivalent:*

(i) $p \leq \operatorname{rg}(A)$

(ii) Es gibt eine p-reihige Unterdeterminante, die nicht null ist.

Beweis. Hat $A = (a_1, \ldots, a_n)$ den Rang $\operatorname{rg}(A) = r$, so kann man aus den Spalten r linear unabhängige $a_{j_1}, \ldots, a_{j_r}$ auswählen; jeweils p dieser Spalten sind auch linear unabhängig. Wir wählen $J \subseteq \{j_1, \ldots, j_r\}$ mit $|J| = p$.
Nun ist der Zeilenrang der $(m \times p)$-Matrix $\tilde{A} = A_{\underline{m},J}$ natürlich genau p. Also können wir p linear unabhängige Zeilen $I \subseteq \underline{m}$, $|I| = p$, auswählen. Dann ist $A_{I,J}$ eine $p \times p$-Untermatrix mit Rang p, also $\operatorname{Det}(A_{I,J}) \neq 0$. $\square$

Man überträgt diesen Satz auf den Rang einer linearen Abbildung:

Satz 6.12.4. *Für eine lineare Abbildung $f : V \to W$ sind äquivalent:*

(i) $p \leq \operatorname{rg}(f) = \dim(\operatorname{im}(f))$

(ii) Es gibt einen p-dimensionalen Unterraum $V' \subseteq V$ und einen p-dimensionalen Quotientenraum $\overline{W}$ von W, so dass $\tilde{F} = \pi \circ f \circ i$ ein Isomorphismus ist.

$$\begin{tikzcd} V \arrow[r, "f"] & W \arrow[d, two heads, "\pi"] \\ V' \arrow[u, hook, "i"] \arrow[r, "\cong"] & \overline{W} \end{tikzcd}$$

(Hierbei ist $i : V' \hookrightarrow V$ die Inklusion, und $\pi : W \twoheadrightarrow \overline{W}$ die Projektion.)

7 Eigenwerte und Eigenvektoren

7.1 Invariante Unterräume

Bei einer linearen Abbildung $f : V \to V$ interessiert uns, ob sie (außer der Null) Fixpunkte hat. Etwas allgemeiner will man wissen, ob sie eine Gerade invariant lässt:
Gibt es $v \neq 0$ und $\lambda \in \mathbb{K}$ mit

$$f(v) = \lambda v \quad ?$$

Diese Gleichung nennt man *Eigenwertgleichung*, wie sie uns z.B. bei der Google-Matrix (siehe Einleitung) begegnet ist; man nennt v einen *Eigenvektor* und λ einen *Eigenwert*.

Wählt man nun eine Basis $\mathcal{B} = (b_1, b_2, \dots, b_n)$, welche mit $b_1 = v$ beginnt, so ist

$$M = M_{\mathcal{B}\mathcal{B}}(f) = \left(\begin{array}{c|ccc} \lambda & * & \dots & * \\ \hline 0 & & & \\ \vdots & & M' & \\ 0 & & & \end{array}\right).$$

Ist nicht nur $U_1 = \operatorname{Span}(v) = \operatorname{Span}(b_1)$ invariant, sondern auch der komplementäre Unterraum $U_1' = \operatorname{Span}(b_2, \dots, b_n)$, so hat M sogar die Gestalt

$$M = \left(\begin{array}{c|c} \lambda & 0 \\ \hline 0 & M' \end{array}\right).$$

Nun könnte man einen Eigenvektor in U' suchen, usw. Könnte man so n linear unabhängige Eigenvektoren finden? Dann hätte M sogar Diagonalgestalt

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

und alles wäre sehr einfach.

7.2 Eigenwerte und Eigenvektoren

Wir betrachten die Eigenwertgleichung genauer.

Definition 7.2.1. Es sei $f : V \to V$ ein Endomorphismus.

1) Ein Skalar $\lambda \in \mathbb{K}$ heißt *Eigenwert* von f, falls es ein $v \in V, v \neq 0$ gibt mit $f(v) = \lambda v$.

2) Ein Vektor $v \in V, v \neq 0$ heißt *Eigenvektor* von f, wenn es ein $\lambda \in \mathbb{K}$ gibt mit $f(v) = \lambda v$.

Die Menge der Eigenwerte heißt *Spektrum* von f und wird mit $\operatorname{Spek}(f)$ bezeichnet; für einen Eigenwert λ ist

$$\operatorname{Eig}_\lambda(f) = \{v \in V \mid f(v) = \lambda v\}$$

der *Eigenraum* von f zu λ und $\dim \operatorname{Eig}_\lambda(f)$ heißt seine *Vielfachheit*.
Für eine Matrix $A \in \operatorname{Mat}_{n,n}(\mathbb{K})$ seien die Begriffe Eigenwert, Eigenvektor und Eigenraum analog definiert, also für $V = \mathbb{K}^n, f = T_A : \mathbb{K}^n \to \mathbb{K}^n, f(x) = Ax$.

Bemerkung 7.2.2. (1) Falls $\dim_K V < \infty$, sehen wir sofort:

$$\mathrm{Eig}_\lambda(f) = \{\lambda \in \mathbb{K} \mid f - \lambda\,\mathrm{id}_V \text{ nicht invertierbar}\}.$$

Falls $\dim_{\mathbb{K}} V < \infty$ ist, ist „nicht invertierbar" äquivalent zu „nicht injektiv"; aber für $\dim_{\mathbb{K}} V = \infty$ it dies anders. Deshalb nennt man für $\dim_{\mathbb{K}} V = \infty$ die Menge $\{\lambda \in \mathbb{K} \mid f - \lambda\,\mathrm{id}_V \text{ nicht invertierbar}\}$ das Spektrum von f und bezeichnet die kleinere Menge mit „nicht injektiv" anstelle von „nicht-invertierbar" dann als die Eigenwerte.

(2) $\mathrm{Eig}_\lambda(f) = \ker(f - \lambda\,\mathrm{id}_V) \neq 0$ für ein $\lambda \in \mathrm{Spek}(f)$ und die Eigenräume sind Untervektorräume.

(3)

$$\begin{aligned} 0 \in \mathrm{Spek}(f) &\iff f \text{ nicht injektiv} \\ 1 \in \mathrm{Spek}(f) &\iff \mathrm{Fix}(f) \neq 0 \end{aligned}$$

(4) $\mathrm{Eig}_\lambda(f) \cap \mathrm{Eig}_\mu(f) = 0$, falls $\lambda, \mu \in \mathrm{Spek}(f), \lambda \neq \mu$; dies ist ein Spezialfall des folgenden Satzes.

Satz 7.2.3. *Sind $v_1, \ldots, v_k$ Eigenvektoren von $f : V \to V$ zu jeweils verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_k$, so sind sie linear-unabhängig.*

Beweis. Für $k = 1$ ist die Aussage richtig, weil ein Eigenvektor immer $\neq 0$ ist.
Sei $k \geq 2$ und es sei

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_k v_k = 0 \tag{7.1}$$

eine Relation. Wir multiplizieren (7.1) mit λ_k und erhalten

$$\lambda_k \alpha_1 v_1 + \lambda_k \alpha_2 v_2 + \ldots + \lambda_k \alpha_k v_k = 0, \tag{7.2}$$

und wir wenden f auf (7.1) an und erhalten

$$\alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \ldots + \alpha_k \lambda_k v_k = 0. \tag{7.3}$$

Die Differenz (7.2) – (7.3) ist eine Relation zwischen den $v_2, \ldots, v_k$:

$$(\lambda_k - \lambda_1)\alpha_1 v_1 + (\lambda_k - \lambda_2)\alpha_2 v - 2 + \ldots + (\lambda_k - \lambda_{k-1})\alpha_{k-1} v_k = 0.$$

Nach Induktionsvoraussetzung folgt: $(\lambda_k - \lambda_i)\alpha_i = 0$ für $i = 1, \ldots, k-1$. Wegen $\lambda_k \neq \lambda_i$ folgt $\alpha_i = 0$ für $i = 1, \ldots, k-1$; und dies in (7.1) eingesetzt ergibt auch $\alpha_k = 0$. □

Korollar 7.2.4. *Eine lineare Abbildung $f : V \to V$ besitzt höchstens $n = \dim_{\mathbb{K}} V$ verschiedene Eigenwerte. Ist $|\mathrm{Spek}(f)| = k$, so gibt es einen f-invarianten Unterraum der Dimension k*

Beispiel 7.2.5. 1)

$$\begin{array}{lll} f = \mathrm{id}_V & \mathrm{Spek}(f) = \{1\} & \mathrm{Eig}_1(f) = V \\ f = 0 & \mathrm{Spek}(f) = \{0\} & \mathrm{Eig}_0(f) = V \end{array}$$

2)

$$\begin{array}{llll} f = -\mathrm{id}_V & \mathrm{Spek}(f) = \{-1\} & \mathrm{Eig}_{-1}(f) = V & \text{(Spiegelung am Nullpunkt)} \\ f = \lambda\,\mathrm{id}_V & \mathrm{Spek}(f) = \{\lambda\} & \mathrm{Eig}_\lambda(f) = V & \end{array}$$

3)

$$f = \text{Spiegelung an einer Geraden im } \mathbb{R}^n$$
$$f(x_1, \ldots, x_n) = (-x_1, x_2, \ldots, x_n)$$
$$\mathrm{Spek}(f) = \{\pm 1\} \qquad \mathrm{Eig}_{+1}(f) = \mathrm{Span}(x_1) \qquad \mathrm{Eig}_{-1}(f) = \mathrm{Span}(x_2, \ldots, x_n)$$

4) Eine lineare Abbildung kann auch keine Eigenwerte haben:

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}, \qquad \operatorname{Spek}(A) = \emptyset.$$

5)

$$A = D[\lambda_1, \dots, \lambda_n] = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \quad \text{Diagonalmatrix}$$

$$\operatorname{Spek}(A) = \{\lambda_1, \dots, \lambda_n\} \qquad \operatorname{Eig}_{\lambda_i}(A) = \operatorname{Span}(\{e_j \mid \lambda_j = \lambda_i\}).$$

6)

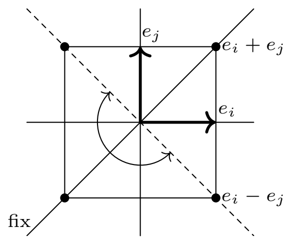
$$A = E_{ij}(\lambda),\ \text{Elementarmatrix } (i \neq j)$$

$$\operatorname{Spek}(A) = \{1\} \qquad \operatorname{Eig}_1(A) = \operatorname{Span}(\{e_k \mid k \neq j\})$$

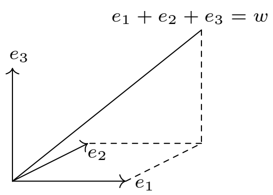
7)

$$A = P_{ij} \text{ Permutationsmatrix} \qquad \operatorname{Spek}(A) = \{\pm 1\}$$

$$\operatorname{Eig}_{+1}(A) = \operatorname{Span}(\{e_k \mid k \neq i, j\} \cup \{e_i + e_j\}) \qquad \operatorname{Eig}_{-1}(A) = \operatorname{Span}(\{e_i - e_j\})$$



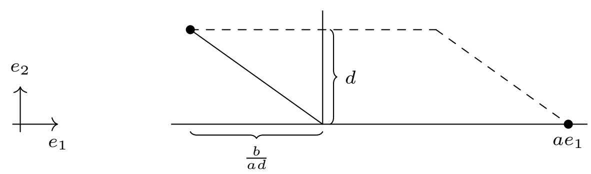
$A = P_{(123)}$ Permutationsmatrix für den 3-Zykel $(123) \in \mathfrak{S}_3$



- $\operatorname{Span}(w) = \operatorname{Fix}(A) = \operatorname{Eig}_1(A)$
- auf $\operatorname{Span}(w)^\perp$: Drehung um 120°

8)

$$A = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = \underbrace{\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & d \end{pmatrix}}_{\text{Streckungen}}\underbrace{\begin{pmatrix} 1 & \frac{b}{a} \\ 0 & 1 \end{pmatrix}}_{\text{Scherung}} \qquad a \neq 0$$



$$\operatorname{Spek}(A) = \{a, d\}$$

$a \neq d:$	$\operatorname{Eig}_a(A) = \operatorname{Span}(e_1),$	$\dim \operatorname{Eig}_d(A) = 1;$
$a = d$ und $b \neq 0:$	$\operatorname{Eig}_a(A) = \operatorname{Span}(e_1);$	
$a = d$ und $b = 0:$	$\operatorname{Eig}_a(A) = \mathbb{K}^2.$	

Lemma 7.2.6. *(a) Ist v Eigenvektor von f zum Eigenwert λ, so ist v auch Eigenvektor von αf zum Eigenwert $\alpha\lambda$ ($\alpha \in \mathbb{K}$):*

$$\mathrm{Eig}_\lambda(f) \subseteq \mathrm{Eig}_{\alpha\lambda}(\alpha f).$$

(b) Ist v Eigenvektor von f zum Eigenwert λ, so ist v auch Eigenvektor von f^q zum Eigenwert λ^q ($q \in \mathbb{N}$):

$$\mathrm{Eig}_\lambda(f) \subseteq \mathrm{Eig}_{\lambda^q}(f^q).$$

Für invertierbares f gilt $\mathrm{Eig}_\lambda(f) = \mathrm{Eig}_{\lambda^{-1}(f^{-1})}$.

(c) Ist v Eigenvektor von f_1 zum Eigenwert λ_1 und ebenso Eigenvektor von f_2 zum Eigenwert λ_2, so ist v auch Eigenvektor von $f_1 + f_2$ zum Eigenwert $\lambda_1 + \lambda_2$:

$$\mathrm{Eig}_{\lambda_1}(f_1) \cap \mathrm{Eig}_{\lambda_2}(f_2) \subseteq \mathrm{Eig}_{\lambda_1+\lambda_2}(f_1 + f_2).$$

(d) Ist v Eigenvektor von f zum Eigenwert λ, so ist v auch Eigenvektor von $p(f) := \alpha_0\,\mathrm{id}_V + \alpha_1 f + \ldots + \alpha_q f^q$ zum Eigenwert $p(\lambda) := \alpha_0 + \alpha_1\lambda + \ldots + \alpha_q\lambda^q$ ($\alpha_0, \ldots, \alpha_q \in \mathbb{K}$):

$$\mathrm{Eig}_\lambda(f) \subseteq \mathrm{Eig}_{p(\lambda)}(p(f))$$

Beweis. klar. □

Lemma 7.2.7. *Sei $f \in \mathrm{End}_{\mathbb{K}}(V), \varphi \in \mathrm{Aut}_{\mathbb{K}}(V)$.*

i) $f' := \varphi \circ f \circ \varphi^{-1}$ und f haben die gleichen Eigenwerte:

$$\mathrm{Spek}(\varphi \circ f \circ \varphi^{-1}) = \mathrm{Spek}(f).$$

ii) Ist v ein Eigenvektor von f zum Eigenwert λ, so ist $v' := \varphi(v)$ Eigenvektor von f' zum Eigenwert λ:

$$\varphi(\mathrm{Eig}_\lambda(f)) = \mathrm{Eig}_\lambda(\varphi \circ f \circ \varphi^{-1}).$$

Beweis. Beide Behauptungen folgen aus der Rechnung

$$(\varphi \circ f \circ \varphi^{-1})(\varphi(v)) = \varphi(f(v)) = \varphi(\lambda v) = \lambda\varphi(v)$$

für $v \in \mathrm{Eig}_\lambda(f)$. □

Lemma 7.2.8. *Für zwei $f, g \in \mathrm{End}_{\mathbb{K}}(V)$ gilt:*

i) $f \circ g$ und $g \circ f$ haben dieselben Eigenwerte:

$$\mathrm{Spek}(f \circ g) = \mathrm{Spek}(g \circ f).$$

ii) Ist v Eigenvektor von $f \circ g$ zum Eigenwert λ, so ist $v' = g(v)$ Eigenvektor von $g \circ f$ zum Eigenwert λ (falls $g(v) \neq 0$):

$$g(\mathrm{Eig}_\lambda(f \circ g)) \subseteq \mathrm{Eig}_\lambda(g \circ f), \qquad f(\mathrm{Eig}_\lambda(g \circ f)) \subseteq \mathrm{Eig}_\lambda(f \circ g)$$

Beweis. Übungsaufgabe. □

7.3 Charakteristisches Polynom

Wie findet man alle Eigenwerte eines Endomorphismus $f : V \to V$?

$$\begin{aligned}\lambda \text{ ist Eigenwert} &\iff \text{es gibt } v \in V \setminus 0 \text{ mit } f(v) = \lambda v\\ &\iff \text{es gibt } v \in V \setminus 0 \text{ mit } (f - \lambda \operatorname{id}_V)(v) = 0\\ &\iff \ker(f - \lambda \operatorname{id}_V) \neq 0\\ &\iff f - \lambda \operatorname{id}_V \text{ nicht invertierbar}\\ &\iff \operatorname{Det}(f - \lambda \operatorname{id}_V) = 0.\end{aligned}$$

Wir haben für festes f eine Funktion

$$\mathbb{K} \longrightarrow \mathbb{K}, \quad \lambda \longmapsto \operatorname{Det}(f - \lambda \operatorname{id}_V)$$

mit Nullstellenmenge $\mathrm{CP}_f^{-1}(0) = \operatorname{Spek}(f)$.

Definition 7.3.1. $\mathrm{CP}_f(t) := \operatorname{Det}(f - t \operatorname{id}_V)$ heißt *charakteristisches Polynom* von f. Für eine Matrix $A \in \mathrm{Mat}_{n,n}(\mathbb{K})$ ist entsprechend

$$\mathrm{CP}_A(t) = \operatorname{Det}(A - t\mathbb{1}) = \operatorname{Det}\begin{pmatrix} a_{11} - t & a_{12} & \dots & a_{1n}\\ a_{21} & a_{22} - t & \dots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{n1} & a_{n2} & \dots & a_{nn} - t\end{pmatrix}.$$

Nach Leibniz-Formel ist dies (bei festem f bzw. A bzw. allen a_{ij}) ein Polynom in t vom Grad n:

$$\mathrm{CP}_f(t) = \prod_{\pi \in \mathfrak{S}_n} \operatorname{sign}(\pi) \tilde{a}_{1,\pi(1)} \cdots \tilde{a}_{n,\pi(n)},$$

mit $\tilde{a}_{ij} = a_{ij}$, falls $i \neq j$, und $\tilde{a}_{ii} = a_{ii} - t$.

Beispiel 7.3.2.

(1) Für die Nullfunktion $f = 0$ ist

$$\mathrm{CP}_0(t) = (-1)^n t^n, \qquad \operatorname{Spek}(0) = \{0\}.$$

(2) Für die Identität $f = \operatorname{id}_V$ ist

$$\mathrm{CP}_{\operatorname{id}_V}(t) = (1 - t)^n, \qquad \operatorname{Spek}(\operatorname{id}_V) = \{1\}.$$

(3) Für eine Diagonalmatrix $D = D[\lambda_1, \dots, \lambda_n]$ ist

$$\mathrm{CP}_D(t) = \prod_{i=1}^{n} (\lambda_i - t), \qquad \operatorname{Spek}(A) = \{\lambda_1, \dots, \lambda_n\}.$$

(4) Für eine Elementarmatrix $A = E_{ij}(\lambda), i \neq j$ ist

$$\mathrm{CP}_A(t) = (1 - t)^n, , \qquad \operatorname{Spek}(A) = \{1\}.$$

(5) Für eine Permutationsmatrix $A = P_{ij}$ für eine Transposition $\sigma = (ij)$ ist

$$\mathrm{CP}_A(t) = \mathrm{Det}\begin{pmatrix} 1-t &&&&&&&&& \\ & \ddots &&&&&&&& \\ && -t & 0 & \cdots & 0 & 1 &&& \\ && 0 & 1-t &&& 0 &&& \\ && \vdots && \ddots && \vdots &&& \\ && 0 &&& 1-t & 0 &&& \\ && 1 & 0 & \cdots & 0 & -t &&& \\ &&&&&&& 1-t && \\ &&&&&&&& \ddots & \\ &&&&&&&&& 1-t \end{pmatrix}$$

$$= \mathrm{Det}\left(\begin{array}{cc|ccc} -t & 1 & & & \\ 1 & -t & & 0 & \\ \hline & & 1-t & & \\ & 0 & & \ddots & \\ & & & & 1-t \end{array}\right) = (t^2-1)(1-t)^{n-2},$$

$$\mathrm{Spek}(A) = \{\pm 1\}.$$

(6) Für eine Drehung $A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$ um Winkel α ist

$$\mathrm{CP}_A(t) = 1 - 2\cos(\alpha)t + t^2,$$
$$\mathrm{Spek}(A) = \begin{cases} \emptyset & \alpha \neq 0, \pi \\ \{1\} & \alpha = 0 \\ \{-1\} & \alpha = \pi \end{cases}$$

(7) Für eine obere (bzw. untere) Dreiecksmatrix

$$A = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \quad \text{bzw.} \quad A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ * & & \lambda_n \end{pmatrix} \quad \text{ist}$$

$$\mathrm{CP}_A(t) = \prod_{i=1}^{n}(\lambda_i - t), \qquad \mathrm{Spek}(A) = \{\lambda_1, \ldots, \lambda_n\}$$

(8) $A \in \mathrm{GL}_n(\mathbb{R})$, $n = 1, 3, 5, \ldots$ ungerade. Dann hat $\mathrm{CP}_A(t)$ mindestens eine Nullstelle und somit ist $\mathrm{Spek}(A) \neq \emptyset$.

(9) $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

$$\mathrm{CP}_A(t) = (a-t)(d-t) - bc = t^2 - \underbrace{(a+d)}_{-\,\mathrm{Spur}(A)}\, t + \underbrace{(ad-bc)}_{\mathrm{Det}(A)}$$

Wir sehen, dass die Berechnung des charakteristischen Polynomes nicht ganz einfach ist; es gibt weniger Rechenregel als für die Determinante (man versuche einmal, den Laplaceschen Entwicklungssatz anzuwenden). Insbesondere ist $\mathrm{CP}_A(t)$ nicht „trivial“, d.h. nicht das Nullpolynom oder ein konstantes Polynome, wenn $\mathrm{rg}(A) < n$ ist.

Satz 7.3.3. *Für eine Blockmatrix* $M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ *bzw.* $M' = \begin{pmatrix} A & 0 \\ C & D \end{pmatrix}$ *gilt*

$$\mathrm{CP}_M(t) = \mathrm{CP}_A(t) \cdot \mathrm{CP}_C(t) \qquad \textit{bzw.}$$

Beweis. $M - t\mathbb{1}$ bzw. $M' - t\mathbb{1}$ hat Blockgestalt $\begin{pmatrix} A - t\mathbb{1} & B \\ 0 & C - t\mathbb{1} \end{pmatrix}$ bzw. $\begin{pmatrix} A - t\mathbb{1} & 0 \\ C & D - t\mathbb{1} \end{pmatrix}$, und so folgt die Behauptung aus dem entsprechenden Ergebnis für die Determinante. □

Satz 7.3.4. *Für zwei $n \times n$-Matrizen A, B gilt:*

$$\mathrm{CP}_{AB}(t) = \mathrm{CP}_{BA}(t).$$

Beweis. Wir haben

$$\begin{pmatrix} -t\mathbb{1} & A \\ 0 & -\mathbb{1} \end{pmatrix} \begin{pmatrix} \mathbb{1} & A \\ B & t\mathbb{1} \end{pmatrix} = \begin{pmatrix} -t\mathbb{1} + AB & 0 \\ -B & -t\mathbb{1} \end{pmatrix}$$

und andersherum

$$\begin{pmatrix} \mathbb{1} & A \\ B & t\mathbb{1} \end{pmatrix} \begin{pmatrix} -t\mathbb{1} & A \\ 0 & -\mathbb{1} \end{pmatrix} = \begin{pmatrix} -t\mathbb{1} & 0 \\ -tB & BA - t\mathbb{1} \end{pmatrix}.$$

Nach dem Produktsatz für Determinanten (also $\mathrm{Det}(AB) = \mathrm{Det}(BA)$) und Satz 7.3.3 folgt

$$\mathrm{CP}_{AB}(-t)^n = (-t)^n \,\mathrm{CP}_{BA}(t),$$

also die Behauptung. □

Satz 7.3.5. *Ähnliche Matrizen haben gleiches charakteristisches Polynom: Für $A \in \mathrm{Mat}_{n,n}(\mathbb{K})$ und $\Omega \in \mathrm{GL}_n(\mathbb{K})$ gilt*

$$\mathrm{CP}_{\Omega A \Omega^{-1}}(t) = \mathrm{CP}_A(t).$$

Beweis. Dies folgt sofort aus vorigem Satz. Man kann aber auch direkt nachrechnen:

$$\Omega A \Omega^{-1} - t\mathbb{1} = \Omega A \Omega^{-1} - \Omega \cdot t\mathbb{1} \cdot \Omega^{-1} = \Omega(A - t\mathbb{1})\Omega^{-1},$$

und benutzen, dass ähnliche Matrizen die gleiche Determinante haben. □

Wir notieren noch

Satz 7.3.6. *(i)* $\mathrm{CP}_{A^\top} = \mathrm{CP}_A(t)$.

(ii) $\mathrm{CP}_{\lambda A}(t) = \lambda^n \,\mathrm{CP}_A(\frac{t}{\lambda})$ *für* $\lambda \neq 0$.

Beweis. Für (i) rechnet man nach

$$A^\top - t\mathbb{1} = A^\top - t\mathbb{1}^\top = (A - t\mathbb{1})^\top$$

und benutzt $\mathrm{Det}(M^\top) = \mathrm{Det}(M)$. Für (ii) rechnet man

$$\lambda A - t\mathbb{1} = \lambda(A - \frac{t}{\lambda}\mathbb{1})$$

und benutzt $\mathrm{Det}(\lambda M) = \lambda^n \,\mathrm{Det}(M)$. □

Korollar 7.3.7. *Für eine schiefsymmetrische Matrix $A = -A^\top$ gilt: Ist n gerade bzw. ungerade, so ist $\mathrm{CP}_A(t)$ ein gerades Polynom $\mathrm{CP}_A(-t) = \mathrm{CP}_A(t)$ bzw. ein ungerades Polynom $\mathrm{CP}_A(-t) = -\mathrm{CP}_A(t)$.*

Koeffizienten des charakteristischen Polynoms

Betrachten wir nun das charakteristische Polynom genauer: Es seien $c_k(A) \in \mathbb{K}$ die Koeffizienten des charakteristischen Polynoms:

$$\mathrm{CP}_A(t) = c_0(A) + c_1(A)t + \ldots + c_{n-1}(A)t^{n-1} + c_n(A)t^n.$$

Aus der Definition und Leibniz-Formel folgt sofort

(1) $c_0(A) = \mathrm{Det}(A)$ (setze $t = 0$).

(2) $c_n(A) = (-1)^n$ (nach Ausmultiplizieren) und somit insbesondere $\operatorname{grad}\mathrm{CP}_A(t) = n$.

Wir kennen noch einen weiteren Koeffizienten:

Proposition 7.3.8. $c_{n-1}(A) = (-1)^{n-1}\,\mathrm{Spur}(A)$.

Beweis. In der Leibniz-Formel für $\mathrm{Det}(A - t\mathbb{1})$ erreicht nur der Term für $\pi = 1 \in \mathfrak{S}_n$ den Grad n für die Unbekannte t:

$$\tilde{a}(1) = (a_{11} - t)(a_{22} - t) \cdot \ldots \cdot (a_{nn} - t),$$

denn fehlt in einem $\tilde{a}(\pi)$ ein Diagonalterm $a_{jj} - t$ mit $i \neq j$ (z.B. $j = \pi(i)$). Also ist

$$\mathrm{CP}_A(t) = \tilde{a}(\pi) + \underbrace{\sum_{\pi \neq 1} \mathrm{sign}(\pi)\tilde{a}(\pi)}_{P(t)}$$

mit $\operatorname{grad}(P(t)) \leq n - 2$. Also trägt $P(t)$ nichts zum Koeffizienten $c_{n-1}(A)$ und $c_n(A) = (-1)^n$ bei, und $c_{n-1}(A)$ ist der $(n-1)$-te Koeffizient von $\tilde{a}(1)$. Durch Ausmultiplizieren erhält man

$$\tilde{a}(1) = (-1)^n t^n (-1)^{n-1}(a_{11} + a_{22} + \ldots + a_{nn})t^{n-1} + \text{niedere Terme}.$$

$\square$

Satz 7.3.9. *Für $k = 0, 1, \ldots, n$ gilt:*

(i) $c_k(\Omega A \Omega^{-1}) = c_k(A)$.

(ii) $c_k(AB) = c_k(BA)$.

(iii) $c_k(A^\top) = c_k(A)$

(iv) $c_k(\lambda A) = \lambda^{n-k} c_k(A)$

(v) $c_k \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \sum_{i+j=k} c_i(A)c_j(D)$.

Gibt es Formeln für die anderen Koeffizienten $c_k(A)$, wie für $k = 0, n-1, n$?

Bemerkung 7.3.10. (1) Wir haben jetzt schon öfter einer Matrix $A \in \mathrm{Mat}_{n,n}(\mathbb{K})$ oder einem Endomorphismus $f \in \mathrm{End}(V)$ eine „Größe“ zugeordnet, d.h. eine natürliche Zahl, oder ein Polynom. Beispiele hierfür waren

$$\begin{aligned} \mathrm{rg}(A) &\in \mathbb{N}, \\ \mathrm{def}(A) &\in \mathbb{N}, \\ \mathrm{Det}(A) &\in \mathbb{K}, \\ \mathrm{Spur}(A) &\in \mathbb{K}, \\ \mathrm{CP}_A(t) &\in \mathbb{K}[t] = \text{ Ring der Polynome in } t \text{ mit Koeffizienten in } \mathbb{K}. \end{aligned}$$

Man nennt eine solche „Größe“ eine Ähnlichkeitsinvariante, wenn sie sich beim Übergang von A zu einer ähnlichen Matrix $A' = BAB^{-1}$ nicht ändert. Die obigen „Größen“ sind allesamt Beispiele für Ähnlichkeitsinvarianten. Nicht nur Det und Spur, sondern alle Koeffizienten des charakteristischen Polynoms sind Ähnlichkeitsinvarianten.

(2) Wir haben zwar das charakteristische Polynom $\mathrm{CP}_A(t)$ als Funktion $\mathbb{K} \to \mathbb{K}$ eingeführt; aber es ist für viele Zwecke besser, es als *formalen Ausdruck* zu betrachten:

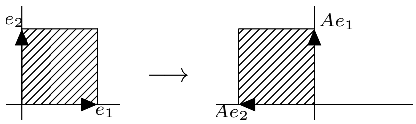
$$p(t) = p_0 + p_1 t + p_2 t^2 + \ldots + p_n t^n$$

Solche Ausdrücke kann man addieren (koeffizientenweise), mit einem Skalar $\alpha \in \mathbb{K}$ skalieren (koeffizientenweise), und multiplizieren (durch distributives Ausmultiplizieren und Zusammenfassen durch $t^k \cdot t^l := t^{k+l}$). In dieser Auffassung, also in $\mathbb{K}[t]$ sind zwei Polynome genau dann gleich, wenn ihre Koeffizienten gleich sind. Für $\mathbb{K} = \mathbb{Q}$, $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ ist zwischen den beiden Auffassungen kein Unterschied.

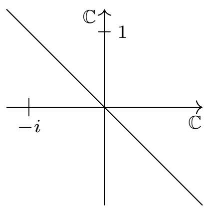
Beispiel 7.3.11. Beispiele für verschiedene Körper

$$A = \begin{pmatrix} 0 & -1//1 & 0 \end{pmatrix}, \qquad \mathrm{CP}_A(t) = t^2 + 1.$$

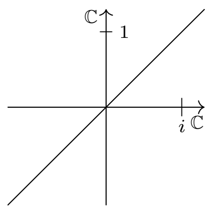
$\boxed{\mathbb{K} = \mathbb{R}}$ $\mathrm{Spek}(A) = \emptyset$



$\boxed{\mathbb{K} = \mathbb{C}}$ $\mathrm{Spek}(A) = \{\pm i\}$

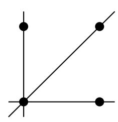


$$\mathrm{Eig}_i(A) = \mathrm{Span}\left(\begin{pmatrix} -i \\ 1 \end{pmatrix}\right) \subseteq \mathbb{C}^2,$$



$$\mathrm{Eig}_{-i}(A) = \mathrm{Span}\left(\begin{pmatrix} -i \\ 1 \end{pmatrix}\right) \subseteq \mathbb{C}^2,$$

$\boxed{\mathbb{K} = \mathbb{F}_2}$ $\mathrm{Spek}(A) = \{1\}$



$$\mathrm{Eig}_1(A) = \mathrm{Span}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right),$$

$\boxed{\mathbb{K} = \mathbb{F}_3}$ $\mathrm{Spek}(A) = \emptyset$

7.4 Diagonalisierbarkeit

Definition 7.4.1. (i) Eine Matrix $A \in \mathrm{Mat}_{n,n}(\mathbb{K})$ heißt *diagonalisierbar*, falls A zu einer Diagonalmatrix ähnlich ist; d.h. es gibt ein $\Omega \in \mathrm{GL}_n(\mathbb{K})$, so dass

$$\Omega A \Omega^{-1} = A' = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

für gewisse $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$.

(ii) Ein Endomorphismus $f \in \mathrm{End}(V)$ heißt *diagonalisierbar*, falls es eine Basis $\mathcal{B}$ von V gibt, so welcher f als Diagonalmatrix dargestellt wird, d.h. $A' = M_{\mathcal{B}\mathcal{B}}(f)$ ist eine Diagonalmatrix $\mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$.

Bemerkung 7.4.2.

1. Offenbar ist f genau dann diagonalisierbar, wenn $A = M_{\mathcal{B}'\mathcal{B}'}(f)$ für irgendeine Basis $\mathcal{B}'$ diagonalisierbar ist; und A ist genau dann diagonalisierbar, wenn $f = L_{\mathcal{B}'\mathcal{B}'}(A)$ für irgendeine Basis $\mathcal{B}'$ diagonalisierbar ist.
2. In (i) sind $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von $\Omega A \Omega^{-1}$, und wegen Lemma 7.2.7(i) auch die von A. Der Standardeinheitsvektor e_i ist Eigenvektor von $\Omega A \Omega^{-1}$ zum Eigenwert λ; also ist nach Lemma 7.2.7(ii) dann $\Omega^{-1} e_i$ Eigenvektor von A zu λ. Also:

   a) $\mathrm{Spek}(A) = \{\lambda_1, \ldots, \lambda_n\}$.

   b) $\mathrm{Eig}_\lambda(A) = \mathrm{Span}(\{\Omega^{-1} e_i \mid \lambda_i = \lambda\})$, $\lambda \in \mathrm{Spek}(A)$.

   In (ii) sind die $\lambda_1, \ldots, \lambda_n$ die Eigenwerte und die Basisvektoren $b_1, \ldots, b_n$ sind Eigenvektoren, genauer ist b_i Eigenvektor zu λ_i und

   a) $\mathrm{Spek}(f) = \{\lambda_1, \ldots, \lambda_n\}$,

   b) $\mathrm{Eig}_\lambda(f) = \mathrm{Span}(\{b_i \mid \lambda_i = \lambda\})$, $\lambda \in \mathrm{Spek}(f)$.

Was sind die Vorteile, wenn man eine Matrix oder einen Endomorphismus diagonalisieren kann?

(I) Alle Ähnlichkeitsinvarianten kann man leicht ermitteln:

- Rang $\mathrm{rg}(A) = |\{\lambda_i \mid \lambda_i \neq 0\}|$,
- Determinante $\mathrm{Det}(A) = \lambda_1 \cdot \ldots \cdot \lambda_n$,
- Spur $\mathrm{Spur}(A) = \lambda_1 + \ldots + \lambda_n$,
- charakterisches Polynom $\mathrm{CP}_A(t) = \prod_{i=1}^n (\lambda_i - t)$,
- Koeffizienten von CP_A $c_k(A) = s_{n-k}(\lambda_1, \ldots, \lambda_n)$, wobei s_{n-k} ein elementar-symmetrisches Polynom ist,
- Spektrum $\mathrm{Spek}(A) = \{\lambda_1, \ldots, \lambda_n\}$,

Für die Eigenräume erhält man $\mathrm{Eig}_\lambda(A) = \mathrm{Span}(\{\Omega^{-1} e_i \mid \lambda_i = \lambda\})$.

(II) Mit einer Basis aus Eigenvektoren hat man Koordinaten $K_\mathcal{B}(v) = (x_1, \ldots, x_n)$ für ein $v \in V$ und f ist in diesen Koordinaten denkbar einfach:

$$K_\mathcal{B}(f(v)) = (\lambda_1 x_1, \ldots, \lambda_n x_n).$$

(III) Man kann die Iterationen von A, also $A, A^2, A^3, \ldots$, leicht ausrechnen durch

$$A^k = \Omega^{-1} \mathrm{Diag}(\lambda_1^k, \ldots, \lambda_n^k) \Omega,$$

und damit das dynamische System $v \mapsto A^k v$.

(IV) Man kann die für die Theorie der Differentialgleichungen wichtige Exponentialfunktion

$$\exp(A) := \sum_{q \geq 0} \frac{A^q}{q!}$$

leicht ausrechnen durch

$$\exp(A) = \Omega^{-1} \mathrm{Diag}(e^{\lambda_1}, \ldots, e^{\lambda_n}) \Omega.$$

(V) Für $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ bilden die diagonalisierbaren Matrizen eine offene und dichte Teilmenge von $\mathbb{K}^{n^2} = \mathrm{Mat}_{n,n}(\mathbb{K})$. Deshalb sind „stetige Formeln“ auf $\mathrm{Mat}_{n,n}(\mathbb{K})$, die für diagonalisierbare Matrizen richtig sind, für alle Matrizen richtig.

Die Aufgabe besteht also darin:

(a) ein gutes Kriterium für die Diagonalisierbarkeit zu finden,

(b) eine Basis aus Eigenvektoren zu bestimmen.

Die Matrix $B = M_{\mathcal{B}\mathcal{B}}(\mathrm{id}_{\mathbb{K}^n}) = (b_1, \ldots, b_n)$ ist im Falle (i) die Basiswechselmatrix von der Standardbasis $\mathcal{E} = (e_1, \ldots, e_n)$ des $\mathbb{K}^n$ in die neue Basis $\mathcal{B} = (b_1, \ldots, b_n)$, die aus den Eigenvektoren besteht. Im Falle (ii) taucht eine Basiswechselmatrix wie folgt auf: hat man f in irgendeiner Basis $\mathcal{A}$ von V durch die Matrix A dargestellt, also $A = M_{\mathcal{A}\mathcal{A}}(f)$, so sei $B = M_{\mathcal{B}\mathcal{A}}(\mathrm{id}_V)$; dann folgt

$$A' = M_{\mathcal{B}\mathcal{B}}(f) = M_{\mathcal{B}\mathcal{A}}(\mathrm{id}_V) \cdot M_{\mathcal{A}\mathcal{A}} \cdot M_{\mathcal{A}\mathcal{B}}(\mathrm{id}_V) = BAB^{-1}.$$

Nun also ein Kriterium zur Diagonalisierbarkeit. Ein Polynom $p(t) = p_0 + p_1 t, p_1 \neq 0$ vom Grad 1 nennt man auch *linear* (obwohl es keine lineare Abbildung $\mathbb{K} \to \mathbb{K}$ im Sinne der Linearen Algebra beschreibt, sondern eine affine; nur für $p_0 = 0$ ist es linear). Tritt es in der speziellen Gestalt $\lambda - t$ als Faktor eines Polynoms $q(t) = (\lambda - t)\tilde{q}(t)$ auf, so ist $t = \lambda$ eine Nullstelle (Wurzel) von $q$und $(\lambda - t)$ heißt *Linearfaktor* von q. Ist $q(t) = (\lambda_1 - t) \cdots (\lambda_n - t)$, so sagt man q *zerfalle* in Linearfaktoren.

Satz 7.4.3 (über die Diagonalisierbarkeit)**.**

(a) (Notwendige Bedingung)
Ist die Matrix A (bzw. der Endomorphismus f) diagonalisierbar, so zerfällt das charakteristische Polynom $\mathrm{CP}_A(t)$ (bzw. $\mathrm{CP}_f(t)$) in Linearfaktoren.

(b) (Hinreichende Bedingung)
Zerfällt das charakteristische Polynom $\mathrm{CP}_A(t)$ einer Matrix A (bzw. $\mathrm{CP}_f(t)$ eines Endomorphismus) in Linear faktoren $(\lambda_i - t), i = 1, \ldots, n$ und(!) *sind diese paarweise verschieden, so ist A (bzw. f) diagonalisierbar.*

Beweis. Es genügt, jeweils eine Fassung (für Matrix oder Endomorphismus) zu beweisen.

(a) Zu A gebe es ein $\Omega \in \mathrm{GL}_n(\mathbb{K})$ mit $\Omega A \Omega^{-1} = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$. Dann folgt

$$\mathrm{CP}_A(t) = \mathrm{CP}_{\Omega A \Omega^{-1}}(t) = (\lambda_1 - t) \cdot \ldots \cdot (\lambda_n - t)$$

nach Satz 7.3.5 und Beispiel 7.3.2.
Gibt es eine Basis $\mathcal{B}$ mit $M_{\mathcal{B}\mathcal{B}}(f) = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n) = A'$ so ist

$$\mathrm{CP}_f(t) = \mathrm{CP}_{A'}(t) = (\lambda_1 - t) \cdot \ldots \cdot (\lambda_n - t)$$

nach Definition 7.3.1 und Beispiel 7.3.2.

(b) Ist $\mathrm{CP}_A(t) = (\lambda_1 - t) \cdots (\lambda_n - t)$ mit $\lambda_i \neq \lambda_j$ für $i \neq j$, so wähle man zu jedem Eigenwert λ_i einen Eigenvektor $b_i \in \mathbb{K}^n$. Nach Satz 7.2.3 sind diese linear unabhängig, also ist $\mathcal{B} = (b_1, \ldots, b_n)$ eine Basis und die Matrix $\Omega = (b_1, \ldots, b_n)$ mit den Spaltenvektoren b_i, so ist wegen $\Omega e_i = b_i$ dieses $\Omega = M_{\mathcal{B}\mathcal{E}}$ die Basiswechselmatrix von der Standardbasis $\mathcal{E}$ nach $\mathcal{B}$. Nun folgt

$$\Omega^{-1} A \Omega e_i = \Omega^{-1} A b_i = \Omega^{-1} \lambda_i b_i = \lambda_i \Omega^{-1} b_i = \lambda_i e_i,$$

also ist $\Omega^{-1} A \Omega = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$ eine Diagonalmatrix.
Zerfällt $\mathrm{CP}_f(t)$ und ist b_i ein Eigenvektor zu λ_i, so ist $\mathcal{B} = (b_1, \ldots, b_n)$ eine Basis von Eigenvektoren. Offenbar ist $M_{\mathcal{B}\mathcal{B}}$ die Diagonalmatrix mit Einträgen $\lambda_1, \ldots, \lambda_n$ auf der Diagonale.

□

7.5 Trigonalisierbarkeit

Wenn wir nur wissen, dass $\mathrm{CP}_A(t)$ in Linearfaktoren $(\lambda_i - t)$ zerfällt, diese aber vielleicht nicht verschieden sind, dann kann man A immerhin noch trigonalisieren.

Definition 7.5.1.

(i) Eine Matrix $A \in \mathrm{Mat}_{n,n}(\mathbb{K})$ heißt *trigonalisierbar*, falls sie zu einer oberen (oder unteren) Dreiecksmatrix ähnlich ist:

$$\Omega A \Omega^{-1} = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \qquad \text{für ein } \Omega \in \mathrm{GL}_n(\mathbb{K}).$$

(ii) Ein Endomorphismus $f \in \operatorname{End}_{\mathbb{K}}(V)$ heißt *trigonalisierbar*, falls es eine Basis $\mathcal{B}$ von V gibt mit

$$M_{\mathcal{B}\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Für eine trigonalisierbare Matrix (bzw. Endomorphismus) gilt damit:

- $\mathrm{CP}_A(t) = (\lambda_1 - t) \cdots (\lambda_n - t)$,
- $\mathrm{Det}(A) = \lambda_1 \cdot \ldots \cdot \lambda_n$,
- $\mathrm{Spur}(A) = \lambda_1 + \ldots + \lambda_n$,
- $c_k(A) = s_{n-k}(\lambda_1, \ldots, \lambda_n$ $(k = 0, \ldots, n)$, s_{n-k} elementarsymmetrisches Polynom,
- $\mathrm{Spek}(A) = \{\lambda_1, \ldots, \lambda_n\}$

Was die Eigenräume angeht, kann man nur

$$\mathrm{Eig}_{\lambda_i}(A) \neq 0 \quad \text{und} \quad \mathrm{Eig}_{\lambda_1}(A) \supseteq \mathrm{Span}(e_1),$$

sagen.

Beispiel 7.5.2.

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

$$\mathrm{CP}_A(t) = (\lambda - t)^2 \qquad\qquad \mathrm{CP}_B(t) = (\lambda - t)^2$$

$$\mathrm{Spek}(A) = \{\lambda\} \qquad\qquad \mathrm{Spek}(B) = \{\lambda\}$$

$$\underbrace{\mathrm{Eig}_\lambda(A)}_{\dim = 2} = \mathbb{K}^2 \qquad\qquad \underbrace{\mathrm{Eig}_\lambda(B)}_{\dim = 1} = \mathbb{K}e_1$$

Satz 7.5.3. CP_A *zerfällt genau dann in (nicht notwendigerweise verschiedene) Linearfaktoren*

$$\mathrm{CP}_A(t) = (\lambda_1 - t) \cdots (\lambda_n - t),$$

wenn A zu einer oberen Dreiecksmatrix ähnlich ist.

Beweis. „$\Longrightarrow$“: Für $n = 1$ ist die Aussage offensichtlich. Es sei b_1 ein Eigenvektor zu λ_1. Ergänzt man dann zu einer Basis $\mathcal{B} = (b_1, \ldots, b_n)$, so hat A (genauer $T_A : \mathbb{K}^n \to \mathbb{K}^n$) in dieser Basis die Darstellung

$$M = M_{\mathcal{B}\mathcal{B}}(T_A) = \left(\begin{array}{c|ccc} \lambda_1 & \alpha_2 & \cdots & \alpha_n \\ \hline 0 & & \tilde{A} & \end{array}\right) = \Omega A \Omega^{-1},$$

mit der Basiswechselmatrix $\Omega = M_{\mathcal{B}\mathcal{E}}(\mathrm{id}_{\mathbb{K}^n}) = (b_1, \ldots, b_n)$, wobei $\mathcal{E} = (e_1, \ldots, e_n)$ die Standarbasis ist.
Nun ist nach Satz 7.3.3 über das charakteristische Polynom von Blockmatrizen

$$\mathrm{CP}_A(t) = \mathrm{CP}_M(t) = (\lambda_1 - t)\,\mathrm{CP}_{\tilde{A}}(t),$$

also ist nach Koeffizientenvergleich $\mathrm{CP}_{\tilde{A}} = (\lambda_2 - t) \cdots (\lambda_n - t)$ zerfallend.
Nach Induktionvoraussetzung können wir annehmen, dass

$$\tilde{A} = \tilde{\Omega}^{-1} \tilde{D} \tilde{\Omega}$$

gilt für eine obere Dreiecksmatrix $\tilde{D}$ und ein $\tilde{\Omega} \in \mathrm{GL}_{n-1}(\mathbb{K})$.
Wir setzen

$$P = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & \tilde{\Omega} \end{array}\right) \in \mathrm{GL}_n(\mathbb{K})$$

und

$$D = \left(\begin{array}{c|c} \lambda_1 & \begin{pmatrix}\alpha_2 & \cdots & \alpha_n\end{pmatrix} \tilde{\Omega}^{-1} \\ \hline 0 & \tilde{D} \end{array}\right),$$

und erhalten

$$P^{-1}DP = \left(\begin{array}{c|ccc} \lambda_1 & \alpha_2 & \cdots & \alpha_n \\ \hline 0 & & \tilde{\Omega}^{-1}\tilde{D}\tilde{\Omega} & \end{array}\right) = \left(\begin{array}{c|ccc} \lambda_1 & \alpha_2 & \cdots & \alpha_n \\ \hline 0 & & \tilde{A} & \end{array}\right) = M,$$

also $A = \Omega^{-1}M\Omega = (\Omega^{-1}P^{-1})D(P\Omega)$.
Die umgekehrte Richtung „$\Longleftarrow$“ ist klar. □

Korollar 7.5.4. *Jede komplexe Matrix ist trigonalisierbar.*

Bemerkung 7.5.5. Ist eine Endomorphismus $f : V \to V$ trigonalisiert durch die Basis $\mathcal{B} = (b_1, \ldots, b_n)$, so bilden die Räume

$$V_k := \mathrm{Span}(\{b_1, \ldots, b_n\})$$

eine aufsteigende Kette

$$0 = V_0 \subseteq V_1 \subseteq \ldots \subseteq V_n = V$$

von f-invarianten Unterräumen $f(V_k) \subseteq V_k$. Dies nennt man eine *f-invariante Fahne*. Die Existenz eienr sochen ist zur Trigonalisierbarkeit äquivalent. (Übung)

Bemerkung 7.5.6. Manchmal ist es wünschenswert, mehrere Matrizen $A, B, \ldots$ *simultan* zu *diagonalisieren* (sofern jede einzelne diagonalisierbar ist), d.h. es soll eine invertierbare Matrix Ω gefunden werden, so dass $\Omega A\Omega^{-1}, \Omega B\Omega^{-1}, \ldots$ alle diagonal sind. Da alle Diagonalmatrizen untereinander vertauschen, müssen dafür alle $A, B, \ldots$ untereinander vertauschen. Dies ist auch hinreichend:

Satz 7.5.7. *Zwei diagonalisierbare Matrizen A und B sind genau dann simultan diagonalisierbar, wenn sie vertauschen (d.h. $AB = BA$ gilt).*

Beweis. Übungsaufgabe. □

Beispiel 7.5.8. Zwei Spiegelungen A und B im $\mathbb{R}^n$ vertauschen genau dann, wenn die „Spiegel“ (d.h. $\mathrm{Fix}(A)$ und $\mathrm{Fix}(B)$) sich senkrect schneiden:

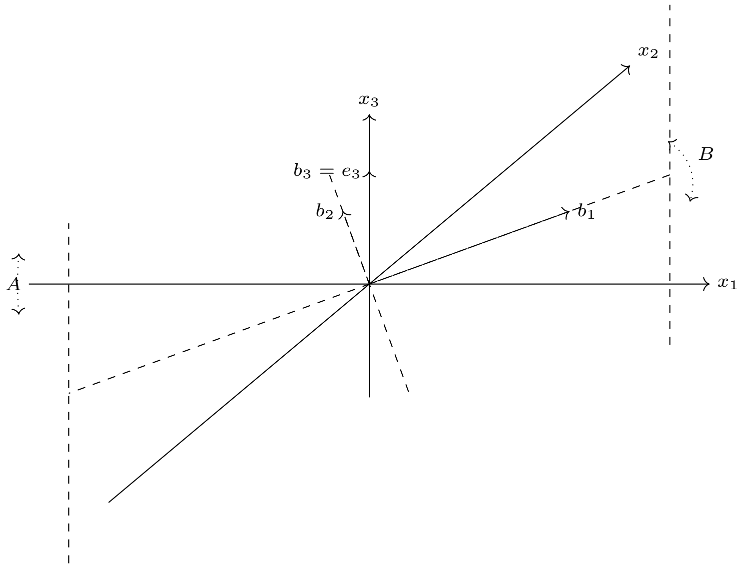
$$v \in \mathrm{Fix}(A), w \in \mathrm{Fix}(B) \implies \langle v, w\rangle = 0.$$

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

Spiegelung an der (x_1, x_2)-Ebene
$\mathrm{Fix}(A) = \mathrm{Span}(e_1, e_2) = \mathrm{Eig}_1(A)$
$\mathrm{Span}(e_3) = \mathrm{Eig}_{-1}(A)$

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

Spiegelung an der Ebene durch die (x_1, x_2)-Diagonale und x_3
$\mathrm{Fix}(B) = \mathrm{Span}(e_1 + e_2, e_3) = \mathrm{Eig}_1(B)$
$\mathrm{Span}(-e_1 + e_2) = \mathrm{Eig}_{-1}(B)$



$$\begin{aligned} b_1 &:= e_1 + e_2 \\ b_2 &:= -e_1 + e_2 \\ b_3 &:= e_3 \end{aligned}$$

1) A ist bereits diagonal, wird also durch $\mathcal{S} = (e_1, e_2, e_3)$ diagonalisiert, aber auch durch $\mathcal{B} = (b_1, b_2, b_3)$ und viele andere Basen.

$$M_{\mathcal{B}\mathcal{B}}(T_A) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = A \qquad \text{(zufälligerweise?)}$$

2) B wird von $\mathcal{S}$ *nicht* diagonalisiert, aber durch $\mathcal{B}$ (und viele andere Basen)

$$M_{\mathcal{B}\mathcal{B}}(T_B) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

8 Ringe

8.1 Ringe

8.2 Unterringe und Ideale

8.3 Polynomringe

Es sei $\mathbb{K}$ ein kommutativer Ring mit 1.

Definition 8.3.1. Der *Polynomring* $\mathbb{K}[x]$ in der Variablen x ist die Menge aller formalen Ausdrücke

$$f = f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$$

für irgendein $n \geq 0$ und $a_0, a_1, \ldots, a_n \in \mathbb{K}$. Man nennt a_i den *Koeffizienten* von x^i in f; das maximale n mit $a_n \neq 0$ heißt *Grad* von f: $n = \mathrm{grad}(f)$. Man nennt a_n den *führenden Koeffizienten*, a_0 den *absoluten Koeffizienten* von f.
Die *Addition* von Polynomen ist definiert durch die Addition der Koeffizienten

$$f(x) = a_0 + a_1 x + \ldots + a_n x^n, \qquad g(x) = b_0 + b_1 x + \ldots + b_m x^m$$
$$f + g := f(x) + g(x) = (a_0 + b_0) + (a_1 + b_1)x + \ldots + (a_i + b_i)x^i + \ldots$$

Die *Multiplikation* von Polynomen ist definiert durch $x^i \cdot x^j = x^{i+j}$ und durch Ausmultiplizieren (Distributivgesetz):

$$f \cdot g = f(x)g(x) = h(x) = c_0 + c_1 x + \ldots + c_{n+m} x^{n+m},$$

mit

$$\begin{aligned} c_0 &= a_0 b_0 \\ c_1 &= a_0 b_1 + a_1 b_0 \\ &\vdots \\ c_k &= \sum_{i+j=k} a_i b_j \\ &\vdots \\ c_{n+m} &= a_n b_m. \end{aligned}$$

Das neutrale Element der Addition ist das *Nullpolynom* $f = 0$ (alle $a_i = 0$); das neutrale Element der Multiplikation ist das *Einspolynom* $f = 1$. Polynome $f = a_0$ vom Grad 0 nennt man *konstant.* Polynome mit führendem Koeffienten 1 heißen *normiert.*

Bemerkung 8.3.2. 1) Die Schreibweise $f = f(x)$ mit einer „Variablen“ oder „Unbekannten“ x ist praktisch und historisch geheiligt.

2) Wer es ganz genau machen möchte, stelle sich den Polynomring $\mathbb{K}[x]$ als abzählbar-unendliche Summe

$$\mathbb{K} \oplus \mathbb{K} \oplus \ldots = \bigoplus_{n \geq 0} \mathbb{K}$$

von Kopien von $\mathbb{K}$ vor; seine Elemente sind die Koeffiziententupel

$$f \leftrightarrow (a_0, a_1, \ldots, a_n, 0, \ldots),$$

die fast überall null sind; die Addition ist komponentenweise erklärt, und die Multiplikation über die obigen „Diagonalformeln“. Diese Identifizierung ist das *Prinzip des Koeffizientenvergleichs.*

3) $R = \mathbb{K}[x]$ ist ein kommutativer Ring; $\mathbb{K}$ selbst ist (aufgefasst als Menge der konstanten Polynome $f(x) = a_0$) ein Unterring.

4) Ist $\mathbb{K}$ ein Körper, so ist $\mathbb{K}[x]$ ein $\mathbb{K}$-Vektorraum abzählbarer Dimension; die Stammpolynome $f_n = x^n$ bilden eine Basis.

5) Man kann das „Adjungieren" einer Unbekannten x wiederholen:

$$(\mathbb{K}[x_1])[x_2] = \mathbb{K}[x_1, x_2]$$

und erhält Polynomringe in mehreren (auch unendlich vielen) Unbekannten:

$$\mathbb{K}[x, y] \ni f(x, y) = a_{00} + a_{10}x + a_{011}y + a_{11}xy + \ldots = \sum_{\substack{i=1,\ldots,n \\ j=1,\ldots,m}} a_{ij}x^i y^j$$

6) Ist $\mathbb{K}$ ein Integritätsbereich, so auch $\mathbb{K}[x]$ und es gilt für $f, g \neq 0$:

$$\operatorname{grad}(f \cdot g) = \operatorname{grad}(f) + \operatorname{grad}(g).$$

7) Die Einheiten in $\mathbb{K}[x]$ sind die Einheiten in $\mathbb{K}$.

8) Ist $\mathbb{K}$ ein Körper, so nennen wir den Quotientenkörper des Rings $R = \mathbb{K}[x]$ den *Körper der rationalen Funktionen*

$$\mathbb{K}(x) := \operatorname{Quot}(\mathbb{K}[x]).$$

Seine Elemente sind die formalen Quotienten $\frac{f}{g} = \frac{f(x)}{g(x)}$, $g \neq 0$ von Polynomen.

8.3.1 Einsetzen in Polynome

Streng zu unterscheiden von einem *(formalen) Polynom*

$$f(x) = a_0 + a_1x + \ldots + a_qx^q$$

ist die *assoziierte Polynomfunktion*

$$f_* = f : \mathbb{K} \longrightarrow \mathbb{K}, \qquad \lambda \longmapsto „f(\lambda)\text{“} := a_0 + a_1\lambda + \ldots + a_q\lambda^q.$$

Diese entsteht durch Einsetzen von $\lambda \in \mathbb{K}$ anstelle von x, die Potenzen sind λ^n, diese werden mit den Koeffizienten a_i skaliert zu $a_i\lambda^n$, und alles aufaddiert.
Man nennt $\lambda \in \mathbb{K}$ eine *Nullstelle* der Polynomfunktion oder auch des Polynoms f, falls $f(\lambda) = 0$ gilt.
Was könnte man noch in ein Polynom einsetzen?

Beispiel 8.3.3. (a) $\mathbb{M}_n = \operatorname{Mat}_{n,n}(\mathbb{K})$

$$f_* : \mathbb{M}_n \longrightarrow \mathbb{M}_n, \qquad f_*(A) = „f(A)\text{“} := a_0\mathbb{1} + a_1A + \ldots + a_nA^n.$$

(b) $\mathbb{K}$ Körper, $\mathbb{E} = \operatorname{End}_{\mathbb{K}}(V)$, $\dim V$ beliebig

$$f_* : \mathbb{E} \longrightarrow \mathbb{E}, \qquad f_*(\varphi) = „f(\varphi)\text{“} := a_0 \operatorname{id}_V + a_1\varphi + \ldots + a_n\varphi^n.$$

Offenbar gilt:

0)	$0_*(A) = 0$	$0_*(\varphi) = 0$
1) $f = 1$	$1_*(A) = \mathbb{1}_n$	$1_*(\varphi) = \operatorname{id}_V$
2) λf	$(\lambda f)_*(A) = \lambda f_*(A)$	$(\lambda f)_*(\varphi) = \lambda f_*(\varphi)$
3) $f \cdot g$	$(f \cdot g)_*(A) = f_*(A)g_*(A)$	$(f \cdot g)_*(\varphi) = f_*(\varphi) \circ g_*(\varphi)$
4) $f + g$	$(f + g)_*(A) = f_*(A) + g_*(A)$	$(f + g)_*(\varphi) = f_*(\varphi) + g_*(\varphi)$

Definition 8.3.4. Sei $\mathbb{K}$ ein Körper. Eine $\mathbb{K}$-*Algebra* $\mathbb{A}$ ist ein Ring mit 1, der auch ein $\mathbb{K}$-Vektorraum ist, so dass für Skalierung und Produkt gilt:

$$(\lambda a)b = \lambda(ab) = a(\lambda b) \qquad (\lambda \in \mathbb{K}, a, b \in \mathbb{A})$$

Beispiel 8.3.5. 1) $\mathbb{A} = \mathbb{K}$

2) $\mathbb{A} = \mathbb{K}[x]$

3) $\mathbb{A} = \mathrm{Mat}_{n,n}(\mathbb{K})$

4) $\mathbb{A} = \mathrm{End}_{\mathbb{K}}(V)$ (wobei V ein $\mathbb{K}$-Vektorraum ist)

1) und 2) sind kommutative, 3) und 4) (i.A.) nicht-kommutative $\mathbb{K}$-Algebren.

Für den Polynomring $\mathbb{A} = \mathbb{K}[x]$ ist die assozierte Polynomfunktion

$$f_* = \mathrm{eval}_f : \mathbb{K}[x] \longrightarrow \mathbb{K}[x]$$
$$f_*(g) = a_0 + a_1 g + a_2 g^2 + \ldots + a_n g^n$$

für ein festes $f \in \mathbb{K}[x]$ und (variables) $g = g(x) \in \mathbb{K}[x]$, also das Einsetzen von g in f:

$$f_*(g) = f(g(x)) = a_0 + a_1 g(x) + a_2 g(x)^2 + \ldots + a_n g(x)^n.$$

Es ist $\mathrm{grad}(f(g)) = \mathrm{grad}(f) \cdot \mathrm{grad}(g)$.

Für eine $\mathbb{K}$-Algebra $\mathbb{A}$ gilt nun:

$$f(g)_* = \mathrm{eval}^{\mathbb{A}}_{f(g)} = (v \mapsto f(g)(v) = f(g(v)))$$

$$\begin{aligned} \mathrm{eval} : \mathbb{K}[x] \times \mathbb{A} &\longrightarrow \mathbb{A} \\ (f, b) &\longmapsto f(b) = a_0 + a_1 b + \ldots + a_n b^n \end{aligned}$$

- f fest: Polynomfunktion $f_* : \mathbb{A} \to \mathbb{A}$ *nicht* linear
- $\mathbb{K}[x] \to \mathrm{End}_{\mathbb{K}}(\mathbb{A}), f \mapsto f_*$ ist ein Ringhomomorphismus.
- Ist $\mathbb{A}$ eine $\mathbb{K}$-Algebra, so gibt es für jedes $f \in \mathbb{K}[x]$ eine *assoziierte Polynomfunktion*

$$f_* = \mathrm{eval}_f : \mathbb{A} \longrightarrow \mathbb{A}, \qquad f_*(v) := a_o 1 + a_1 v + \ldots + a_q v^q.$$

- Weiter gibt es (für festes $v \in \mathbb{A}$) einen *Evaluationshomomorphismus* von Algebren

$$\mathrm{eval}_v : \mathbb{K}[x] \longrightarrow \mathbb{A}, \qquad f = a_0 + a_1 x + \ldots + a_n x^n \longmapsto \mathrm{eval}_v(f) = a_0 1 + a_1 v + \ldots + a_q v^q,$$

  für den gilt:

$$\begin{array}{lll} \mathrm{eval}_v(0) = 0 & \mathrm{eval}_v(f+g) = \mathrm{eval}_v(f) + \mathrm{eval}_v(g) & \mathrm{eval}_v(\lambda f) = \lambda\, \mathrm{eval}_v(f) \\ \mathrm{eval}_v(1) = 1 & \mathrm{eval}_v(f \cdot g) = \mathrm{eval}_v(f) \cdot \mathrm{eval}_v(g) & \end{array}$$

8.3.2 Division mit Rest in Polynomringen

Es gibt für Polynomringe über Körpern wie im Ring $\mathbb{Z}$ der ganzen Zahlen einen *Euklidischen Algorithmus*; darin übernimmt der Grad die Rolle des Betrags. (Kommutative Ringe mit einem Euklidischen Algorithmus zu passender Betragsfunktion heißen *Euklidische Ringe*.)

Satz 8.3.6. *Zu je zwei $f, g \in \mathbb{K}[x]$ mit $g \neq 0$ gibt es eindeutig bestimmte $q, r \in \mathbb{K}[x]$ mit*

$$\underbrace{f}_{\textit{Dividend}} = \underbrace{q}_{\textit{Quotient}} \cdot \underbrace{g}_{\textit{Divisor}} + \underbrace{r}_{\textit{Rest}} \quad \textit{und } \mathrm{grad}(r) < \mathrm{grad}(g).$$

Beweis. Der Beweis ist fast wörtlich der gleiche wie beim Euklidischen Algorithmus für $\mathbb{Z}$, wenn man den Betrag $|a|$ einer Zahl durch den Grad $\mathrm{grad}(f)$ eines Polynoms ersetzt. Wir skizzieren die wesentliche Idee durch eine Liste von Koeffizienten:

$$\begin{array}{ccc|l}
\text{Grad} & f & g & \\
\hline
x^n & a_n \neq 0 & & \uparrow \\
x^{n-1} & a_{n-1} & & \cdot \frac{a_n}{b_m} x^{n-m} \\
\vdots & \vdots & & \\
x^m & a_m & 0 \neq b_m & \\
x^{m-1} & a_{m-1} & b_{m-1} & \\
\vdots & \vdots & \vdots & \\
x & a_1 & b_1 & \\
1 & a_0 & b_0 &
\end{array}$$

Es sei $n = \mathrm{grad}(f), m = \mathrm{grad}(g)$.

Fall $n = m$: Man setzt $q = \frac{a_n}{b_n}$ und $r = 0$.
Fall $n < m$: Man setzt $\varphi = 0$ und $r = f$.
Fall $n \geq m$: Man „hebt“ g durch Multiplikation mit $\frac{a_n}{b_m} x^{n-m}$ auf den gleichen Grad wie f an und beginnt den ersten Divisionsschritt mit dem „Ansatz“

$$\begin{aligned}
r'(x) :&= f(x) - \frac{a_n}{b_m} x^{n-m} g(x) \\
&= (a_n x + \dots) - (\frac{a_n}{b_m} x^{n-m} b_m x^m + \dots) \\
&= \underbrace{a_n - \frac{a_n}{b_m} b_m}_{=0} x^n + \underbrace{c_{n-1} x^{n-1} + \dots + c_0}_{\text{Grad } < n}.
\end{aligned}$$

Also haben wir

$$f(x) = \frac{a_n}{b_m} x^{n-m} \cdot g(x) + r'(x),$$

mit $\mathrm{grad}(r') < n$. Ist $\mathrm{grad}(r') < m$, sind wir fertig. Ist $\mathrm{grad}(r') > m$, so dividieren wir nun r' durch g mit Rest. Anders gesagt, wir können durch Induktion annehmen, dass in diesem Fall

$$r'(x) = q'(x) g(x) + r''(x) \qquad \text{mit } \mathrm{grad}(r'') < m$$

für ein $q', r'' \in \mathbb{K}[x]$ gilt; zusammen also

$$\begin{aligned}
f(x) &= \frac{a_n}{b_m} x^{n-m} \cdot g(x) + q'(x) g(x) + r(x) \\
&= \left(\frac{a_n}{b_m} x^{n-m} + q'(x) \right) g(x) + r(x).
\end{aligned}$$

□

Satz 8.3.7. *Ist $\lambda \in \mathbb{K}$ eine Nullstelle von $f \in \mathbb{K}[x]$, so ist das Polynom $(\lambda - x)$ ein Teiler von f.*

Beweis. Nach Satz 8.3.6 können wir f durch $(\lambda - x) = g$ mit Rest teilen:

$$f(x) = \varphi(x)(\lambda - x) + r(x).$$

Weil $\mathrm{grad}(r) < \mathrm{grad}(g) = 1$ ist, muss $r(x) = c$ konstant sein, also

$$f(x) = \varphi(x)(\lambda - x) + c.$$

Setzt man $x = \lambda$ ein, so folgt $c = 0$. □

Korollar 8.3.8. *Ein Polynom von Grad n kann höchstens n (verschiedene) Nullstellen haben.*

Beweis. Angenommen, $\lambda_1, \ldots, \lambda_k$ sind die Nullstellen. Dann wäre

$$\begin{aligned} f(x) &= \varphi_1(x)(\lambda_1 - x) \\ &= \varphi_2(x)(\lambda_2 - x)(\lambda_1 - x), && \text{weil } \varphi_1(\lambda_2) = 0 \\ &= \varphi_3(x)(\lambda_3 - x)(\lambda_2 - x)(\lambda_1 - x), && \text{weil } \varphi_2(\lambda_3) = 0 \\ &\vdots \\ &= \varphi_k(x)(\lambda_k - x) \cdots (\lambda_1 - x), && \text{weil } \varphi_{k-1}(\lambda_k) = 0 \end{aligned}$$

und $n = \operatorname{grad}(f) = \operatorname{grad}(\varphi_k) + k$. □

Vielfachheit von Nullstellen

Ist $\lambda \in \mathbb{K}$ eine Nullstelle von $f(x)$, so gilt also

$$f(x) = (\lambda - x)\varphi_1(x)$$

Ist λ eine Nullstelle von $\varphi_1(x)$, so gilt sogar

$$f(x) = (\lambda - x)^2\varphi_2(x), \quad \text{usw.}$$

Definition 8.3.9. Das maximale $l \geq 0$, so dass $(\lambda - x)^l$ ein Teiler von $f(x)$ ist, heißt *Vielfachheit* der Nullstelle λ, geschrieben $\mu(f; \lambda)$.

- $\mu(f; \lambda) \geq 1 \iff \lambda$ ist Nullstelle von f.
- $0 \leq \mu(f; \lambda) \leq n = \operatorname{grad}(f)$.
- $\mu(f; \lambda_1) + \ldots + \mu(f; \lambda_k) \leq \operatorname{grad}(f)$, falls $\lambda_1, \ldots, \lambda_k$ verschiedene Nullstellen von f sind.

Ob ein Polynom $p(t) \in \mathbb{K}[t]$ in Linearfaktoren zerfällt, ist dazu äquivalent, dass es $n = \operatorname{grad}(p)$ Nullstellen hat, denn jede Nullstelle $\lambda_1, \ldots, \lambda_k$ erlaubt die Abspaltung eines Linearfaktors $(\lambda_i - t)$, also

$$p(t) = (\lambda_1 - t) \cdot \ldots \cdot (\lambda_k - t)\tilde{p}(t).$$

Hier ist zu betonen, dass ein Linearfaktor durchaus mehrmals abgespalten werden kann.

Die Suche nach den Nullstellen eines Polynoms ist eine der Kernaufgaben der Algebra. Wir müssen uns hier mit den folgenden Überlegungen bescheiden:

$\underline{\mathbb{K} = \mathbb{C}}$:
Der Körper der komplexen Zahlen ist *algebraisch abgeschlossen*:

Satz 8.3.10 (Fundamentalsatz der Algebra)**.** *Ein komplexes Polynom f mit* $\operatorname{grad}(f) \geq 1$ *zerfällt in Linearfaktoren:*

$$f(x) = c(x - \lambda_1) \cdot \ldots \cdot (x - \lambda_n)$$

für die Nullstellen $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ (mit Vielfachheiten) und ein gewisses $c \in \mathbb{C}, c \neq 0$.

Also besitzt jeder Endomorphismus f eines komplesen Vektorraums mindestens einen Eigenvektor. Wir werden später einen Beweis des Fundamentalsatzes geben, der (fast) nur Mittel der Linearen Algebra benutzt.

$\underline{\mathbb{K} = \mathbb{R}}$:
Fasst man ein reelles Polynom $f(x) \in \mathbb{R}[x]$ als ein komplexes Polynom $f(x) \in \mathbb{C}[x]$ auf, zerfällt es also in Linearfaktoren $(\lambda_i - x)$ für insgesamt n komplexe Nullstellen . Man rechnet sofort nach:

- Ist λ eine Nullstelle von f, so auch $\overline{\lambda}$.

Die Nullstellen von f kann man also wie folgt sortieren:

(1) $\lambda'_1, \ldots, \lambda'_n$ reelle Nullstellen $(0 \leq k \leq n)$

(2) $\lambda_1, \overline{\lambda_1}, \ldots, \lambda_l, \overline{\lambda_l}$ nicht-reelle Nullstellen (l Paare)

(3) $k + 2l = n$.

Wir erhalten (für normiertes $f \in \mathbb{R}[x]$):

$$f(x) = (x - \lambda_1') \cdots (x - \lambda_k') g_1(x) \cdots g_l(x),$$

mit

$$\begin{aligned} g_j(x) :&= (x - \lambda_j)(x - \overline{\lambda_j}) \qquad\qquad (j = 1, \ldots, l) \\ &= \underbrace{\lambda_j \overline{\lambda_j}}_{\in \mathbb{R}} - (\underbrace{\lambda_j + \overline{\lambda_j}}_{\in \mathbb{R}}) x + x^2 \end{aligned}$$

Ein reelles Polynom zerfällt also in $\mathbb{R}[x]$ in Linearfaktoren und quadratische Faktoren ohne reelle Nullstellen. Die letzte Behauptung folgt mit der *Diskriminante*

$$D := b^2 - 4ac, \qquad a = 1, b = \lambda_j + \overline{\lambda_j}, c = \lambda_j \overline{\lambda_j}$$

aus der Lösungsformel

$$x_{1,2} = \frac{-b \pm \sqrt{D}}{2a}$$

eines quadratischen Polynoms $ax^2 + bx + c = 0$.)

Ein reelles Polynom $p(t) = p_0 + p_1 t + \cdots + p_n t^n$ ungeraden Grades $n = 2n' + 1$ hat mindestens eine reelle Nullstelle: für sehr großes $t \gg 0$ ist $p(-t)$ negativ und $p(t)$ positiv, falls $p_n > 0$, und umgekehrt, falls $p_n < 0$. Nach dem Zwischenwertsatz für stetige Funktionen (wie z.B. Polynome) liegt dann irgendwo dazwischen (mindestens) eine Nullstelle. Ein reelles Polynom geraden Grades $n = 2n'$ muss nicht unbedingt eine Nullstelle haben (wie z.B. $p(t) = 1 + t^2$). Es folgt dann durch Induktion:

- Hat $p(t) \in \mathbb{R}[t]$ ungeraden Grad $n = 2n'+1$, so hat $p(t)$ entweder $1, 3, 5, \ldots$ oder $2n'+1$ Linearfaktoren.
- Hat $p(t) \in \mathbb{R}[t]$ geraden Grad $n = 2n'$, so hat $p(t)$ entweder $0, 2, 4, \ldots$ oder $2n'$ Linearfaktoren.

Insbesondere haben wir

Satz 8.3.11. *Ein Endomorphismus $f : \mathbb{R}^{2n+1} \to \mathbb{R}^{2n+1}$ hat mindestens eine invariante Gerade.*

8.4 Symmetrische Polynome

Die symmetrische Gruppe $\mathfrak{S}_n$ operiert durch Vertauschung der Variablen auf dem Polynomring $\mathbb{K}[x_1, \ldots, x_n]$:

$$\S_n \times \mathbb{K}[x_1, \ldots, x_n] \longrightarrow \mathbb{K}[x_1, \ldots, x_n], \qquad \sigma.f(x_1, \ldots, x_n) = f(x_{\sigma(1), \ldots, x_{\sigma(n)}}.$$

Definition 8.4.1. Ein Polynom f heißt *symmetrisch*, falls $\sigma.f = f$ für alle $\sigma \in \mathfrak{S}_n$ gilt. Die Menge der symmetrischen Polynome bezeichnet man mit $\mathbb{K}[x_1, \ldots, x_n]^{\mathfrak{S}_n}$.

(1) $\mathbb{K}[x_1, \ldots, x_n]^{\mathfrak{S}_n}$ ist eine Unteralgebra von $\mathbb{K}[x_1, \ldots, x_n]$.

(2) Die Polynome

$$E_k(x_1, \ldots, x_n = \sum_{i_1 < i_2 < \ldots < i_k} x_{i_1} \cdot \ldots \cdot x_{i_k}$$

heißen *elementarsymmetrische Polynome* ($k = 0, 1, \ldots, n$).

$$\begin{aligned} E_0 &= 1, \\ E_1 &= x_1 + x_2 + \ldots + x_k, \\ E_2 &= x_1 x_2 + x_1 x_3 + \ldots + x_{n-1} x_n, \\ &\vdots \\ E_n &= x_1 \cdot \ldots \cdot x_n. \end{aligned}$$

Bemerkung 8.4.2. Die symmetrischen Polynome snid deshalb wichtig, weil viele Größen wie Determinanten etc. symmetrische Polynome in Nullstellen oder Eigenwerten sind.

Der folgende Satz zeigt bereits die Bedeutung der E_k.

Satz 8.4.3. *Für ein zerfallendes Polynom $f(x)$ mit den Nullstellen $\lambda_1, \dots, \lambda_n$ gilt*

$$f(x) = \prod_{i=1}^{n} (\lambda_i - x) = \sum_{k=0}^{n} E_k(\lambda_1, \dots, \lambda_n) x^k.$$

Beweis. Offensichtlich. □

Der nächste Satz besagt, dass die E_k in $\mathbb{K}[x_1, \dots, x_n]^{\mathfrak{S}_n}$ erstens algebraische Erzeuger sind und zweitens algebraisch unabhängig sind.

Satz 8.4.4. $\mathbb{K}[x_1, \dots, x_n]^{\mathfrak{S}_n}$ *ist ein Polynomring in den E_k, d.h. es gibt einen Ringisomorphismus*

$$\begin{aligned} \varphi_E : \mathbb{K}[e_1, \dots, e_n] &\longrightarrow \mathbb{K}[x_1, \dots, x_n]^{\mathfrak{S}_n} \subseteq \mathbb{K}[x_1, \dots, x_n] \\ e_k &\longmapsto E_k \quad (k = 1, \dots, n) \end{aligned}$$

Zum Beweis müsste man zwei Aussagen beweisen:

(1) Für jedes symmetrische Polynom $f(x_1, \dots, x_n)$ gibt es ein Polynom $\Phi(e_1, \dots, e_n)$ mit

$$f(x_1, \dots, x_n) = \Phi\left(E_1(x_1, \dots, x_n), \dots, E_n(x_1, \dots, x_n)\right).$$

(2) Gilt für ein Polynom $\Phi(e_1, \dots, e_n)$ die Gleichung

$$\Phi\left(E_1(x_1, \dots, x_n), \dots, E_n(x_1, \dots, x_n)\right) = 0,$$

so ist $\Phi(e_1, \dots, e_n) = 0$ das Nullpolynom.

8.5 *Anwendung I*: Vielfachheit von Eigenwerten

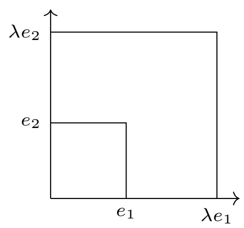
Wo liegt die Diskrepanz zwischen dem notwendigem und dem hinreichenden Kriterium für die Diagonalisierbarkeit?

Beispiel 8.5.1.

1)

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

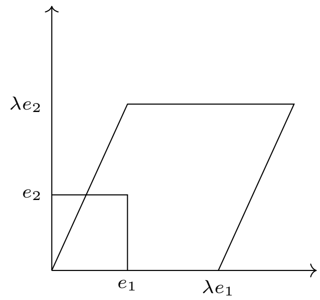
$$\operatorname{Spek}(A) = \{\lambda\}, \quad \mathrm{CP}_A(t) = (\lambda - t)^2, \quad \operatorname{Eig}_\lambda(A) = \mathbb{K}^2, \quad \dim \operatorname{Eig}_\lambda(A) = 2.$$



2)

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

$$\mathrm{Spek}(A) = \{\lambda\}, \quad \mathrm{CP}_A(t) = (\lambda - t)^2, \quad \mathrm{Eig}_\lambda(A) = \mathbb{K}e_1, \quad \dim \mathrm{Eig}_\lambda(A) = 1$$



Ist ein Eigenwert $\lambda \in \mathrm{Spek}(A)$ gegeben, so gibt es

(1) seine Vielfachheit als Nullstelle des charakteristischen Polynoms

(2) die Dimension des Eigenraumes.

Definition 8.5.2. Sei λ ein Eigenwert von $A \in \mathrm{Mat}_{n,n}(\mathbb{K})$ bzw. von $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$, $n = \dim_{\mathbb{K}}(V)$.

(i) Die *algebraische Vielfachheit* $\mathrm{algV}_A(\lambda)$ bzw. $\mathrm{algV}_\varphi(\lambda)$ ist die Vielfachheit von λ im charakteristischen Polynom $\mathrm{CP}_A(t)$ bzw. $\mathrm{CP}_\varphi(t)$.

(ii) Die *geometrische Vielfachheit* $\mathrm{geomV}_A(\lambda)$ bzw. $\mathrm{geomV}_\varphi(\lambda)$ ist die Dimension des Eigenraums von λ, also $\dim \mathrm{Eig}_A(\lambda)$ bzw. $\dim \mathrm{Eig}_\varphi(A)$.

Satz 8.5.3. $1 \leq \mathrm{geomV}_A(\lambda) \leq \mathrm{algV}_A(\lambda) \leq n$.

Beweis. Ist $\mathcal{B}' = (b_1, \ldots, b_k)$ eine Basis des Eigenraums $\mathrm{Eig}_A(\lambda)$ und ist $\mathcal{B} = (b_1, \ldots, b_k, \ldots, b_n)$ eine Ergänzung zu einer Basis von $V)\mathbb{K}^n$, so ist

$$M = M_{\mathcal{B}\mathcal{B}}(\varphi) = \left(\begin{array}{ccc|c} \lambda & & & \\ & \ddots & & A \\ & & \lambda & \\ \hline & 0 & & B \end{array}\right).$$

Also ist $\mathrm{CP}_M(t) = (\lambda_1 - t)^k \, \mathrm{CP}_B(t)$ und folglich $k = \mathrm{geomV}_M(\lambda) \leq \mathrm{algV}_M(\lambda)$. □

Satz 8.5.4. *Eine Matrix* $A \in \mathrm{Mat}_{n,n}(\mathbb{K})$ *ist genau dann diagonalisierbar, wenn das charakteristische Polynom in Linearfaktoren* $(\lambda_i - t)$ *zerfällt* $(i = 1, \ldots, n)$ *und für jedes* λ_i *gilt:*

$$\mathrm{geomV}_A(\lambda_i) = \mathrm{algV}_A(\lambda_i).$$

8.6 *Anwendung II*: Satz von Cayley-Hamilton

Sei A eine feste Matrix.

$$\mathrm{eval}_A : \mathbb{K}[x] \longrightarrow \mathrm{Mat}_{n,n}(\mathbb{K}), \qquad f(x) = a_0 + a_1 x + \ldots + a_r x^r \longmapsto f(A) = a_0 \mathbb{1} + a_1 A + \ldots + a_r A^r$$

- Das Bild ist eine kommutative Unteralgebra

$$\mathrm{im}(\mathrm{eval}_A) =: \mathbb{K}[A].$$

- Der Kern ist ein Ideal genannt „Ideal von A“ bzw. *Auswertungsideal* bzw. *Annihilatorideal* von A

$$\ker(\mathrm{eval}_A) = \{f \in \mathbb{K}[x] \mid f(A) = 0\} =: I_A$$

- Analog für festes φ

$$\mathbb{K}[x] \longrightarrow \mathrm{End}_K(V), \qquad f \longmapsto f(\varphi) = a_0 \,\mathrm{id}_V + a_1\varphi + \ldots + a_r\varphi^r$$

- Gibt es überhaupt nicht-triviale $f \in I_A$?
  Vorüberlegung: Sei $A = D(\lambda_1, \ldots, \lambda_n)$ diagonal.

$$\mathrm{CP}_A(t) = (\lambda_1 - t)(\lambda_2 - t)\cdots(\lambda_n - t) \quad \text{zerfallend}$$

  $\Rightarrow A$ ist trigonalisierbar und wir erhalten somit eine invariante Fahne.

$$\Rightarrow \Phi = (\lambda_1 \mathbb{1} - A)(\lambda_2 \mathbb{1} - A)\cdots(\lambda_n \mathbb{1} - A) = 0$$

  Also ist $f = \mathrm{CP}_A \in I_A$.

Satz 8.6.1 (Satz von Cayley-Hamilton)**.** *(i) Für jede Matrix A gilt* $\mathrm{CP}_A(A) = 0$*, also* $\mathrm{CP}_A \in I_A$.

(i) Für jeden Endomorphismus $\varphi : V \to V$ eines endlich-dimensionalen $\mathbb{K}$-Vektorraums V gilt $\mathrm{CP}_\varphi(\varphi) = 0$*, also* $\mathrm{CP}_\varphi \in I_\varphi$.

CAVEAT: Das folgende „Einsetzen" ist nicht korrekt und deshalb kein Beweis:

$$\mathrm{CP}_A(A)) = \mathrm{Det}(A - A \cdot \mathbb{1}) = \mathrm{Det}(A - A) = \mathrm{Det}(0) = 0.$$

Wenn für t das A eingesetzt wird, dann müsste für jeden Skalar a_{ij} auch $a_{ij}\mathbb{1}$ eingesetzt werden.

Beweis. $n = 1$: $A = (\lambda)$, also $\mathrm{CP}_A(t) = \lambda - t$ und $CP_A(A) = \lambda\mathbb{1} - (\lambda) = 0$.

1) $B(t) := (A - t\mathbb{1})^\top \in \mathrm{Mat}_{n,n}(\mathbb{K}[t])$ ist eine Matrix mit Einträgen in $\mathbb{K}[t]$ und zwar auf der Diagonalen $b_{ii}(t) = a_{ii} - t$ lineare Polynome und weg von der Diagonalen $b_{ij}(t) = a_{ij}$ $(i \neq j)$ konstante Polynome.

2) $\mathrm{Det}(B(t)) = \mathrm{CP}_A(t) \in \mathbb{K}[t]$

$$\mathrm{Det} : \mathrm{Mat}_{n,n}(\underbrace{\mathbb{K}[t]}_{R,\text{kommutativ}}) \longrightarrow \underbrace{\mathbb{K}[t]}_{R}$$

3) Einsetzen von A für t und $a_{ij}\mathbb{1}$ für a_{ij} ergibt

$$\mathrm{Mat}_{n,n}(\mathbb{K}[t]) \longrightarrow \mathrm{Mat}_{n,n}(\mathbb{K}[A]) \subseteq \mathrm{Mat}_{n,n}(\mathbb{M}_n)$$

$$B(A) = \begin{pmatrix} a_{11}\mathbb{1} - A & a_{12}\mathbb{1} & \ldots & a_{1n}\mathbb{1} \\ a_{21}\mathbb{1} & \ddots & & \vdots \\ \vdots & & \ddots & a_{n-1\;n}\mathbb{1} \\ a_{n1}\mathbb{1} & \ldots & a_{n\;n-1}\mathbb{1} & a_{nn}\mathbb{1} - A \end{pmatrix}$$

   Dies ist eine $(n^2 \times n^2)$-Matrix.

4) Man kann $B(A)$ mit einem $(n \times 1)$-Spaltenvektor multiplizieren, dessen Einträge $(n \times 1)$-Spaltenvektoren in $\mathbb{K}^n$ sind:

$$e = \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \qquad \text{mit} \quad e_i = \text{Standardbasisvektor des } \mathbb{K}^n$$

$$B(A)e = \begin{pmatrix} (a_{11}e_1 - Ae_1) + a_{21}e_2 + \ldots + a_{n1}e_n \\ \vdots \\ a_{1n}e_1 + a_{2n}e_2 + \ldots + (a_{nn}e_n - Ae_n) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

5) Es sei $B^\#(t) \in \mathrm{Mat}_{n,n}(\mathbb{K}[t])$ die Adjunkte zu $B(t)$. Ihre Einträge sind Polynome von Grad $\leq n - 1$ in $\mathbb{K}[t]$ und zwar ist

$$b^\#_{ij} = (-1)^{i+j}\,\mathrm{Det}(B'(t)_{ji})$$

   wobei $B'(t)_{ji}$ die (j, i)-Streichungsmatrix von $B(t)$ sei.
   Wir wissen:

$$B^\#(t) \cdot B(t) = \mathrm{Det}(B(t)) \cdot \mathbb{1}_{n^2} = \mathrm{CP}_A(t) \cdot \mathbb{1}_{n^2}$$

6) Setzen wir hier A für t ein, so erhalten wir

$$B^{\#}(A) \cdot B(A) = \mathrm{CP}_A(A) \cdot \mathbb{1},$$

bzw. ausführlicher

$$\begin{pmatrix} \mathrm{CP}_A(A) & & \\ & \ddots & \\ & & \mathrm{CP}_A(A) \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} \mathrm{CP}_A(A)e_1 \\ \vdots \\ \mathrm{CP}_A(A)e_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

Also gilt $\mathrm{CP}_A(A)e_i = 0$ für alle i, und damit $A = 0$.

□

Korollar 8.6.2.

- *Für jedes A und jedes $k \geq 0$ gibt es ein Polynom $p_k(x) \in \mathbb{K}[x]$ mit $p_k(A) = A^k$ und* $\mathrm{grad}(p_k) < n$.
- *Für jedes invertierbare A gibt es ein Polynom $p_{-1}(x) \in \mathbb{K}[x]$ mit $p_{-1}(A) = A^{-1}$ und* $\mathrm{grad}(p_{-1} < n$. *Also gibt es auch für jedes $k \in \mathbb{Z}$ ein $p_k(x)$ mit $p_k(x) = A^k$ und* $\mathrm{grad}(p_k) < n$.

8.7 *Anwendung III*: Minimalpolynom

Wir betrachten für ein festes $A \in \mathrm{Mat}_{n,n}(\mathbb{K}) = \mathbb{M}_n$ mit $A = M_{\mathcal{BB}}(\varphi)$ für $\varphi : V \to V$ mit $\dim V = n$ und $\mathcal{B}$ Basis von V die Auswertung

$$\begin{tikzcd} & f \arrow[r, mapsto] & f(A) \\ I_A \arrow[r, hook] \arrow[d, equal] & \mathbb{K}[t] \arrow[r, "\mathrm{eval}_A"] \arrow[d, equal] & \mathbb{M}_n \arrow[d, shift left, "L_{\mathcal{BB}}"] \\ I_\varphi \arrow[r, hook] & \mathbb{K}[t] \arrow[r, "\mathrm{eval}_A"] & \mathbb{M}_n \arrow[u, shift left, "M_{\mathcal{BB}}"] \\ & f \arrow[r, mapsto] & f(\varphi) \end{tikzcd}$$

Mit dem Auswertungsideal $I_A = I_\varphi$. Da eval_A ein Vektorraumhomomorphismus ist und $\mathbb{M}_n$ die Dimension n^2 hat, können die $n^2 + 1$ Potenzen

$$A^0 = \mathbb{1}, A, A^2, \ldots, A^{n^2}$$

nicht linear unabhängig sein. Eine lineare Relation

$$\alpha_0 \mathbb{1} + \alpha_1 A + \alpha_2 A^2 + \ldots + \alpha_{n^2} A^{n^2} = 0$$

heißt aber, es gibt ein Polynom

$$f(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \ldots + \alpha_{n^2} x^{n^2}$$

vom Grad $\leq n^2$ mit $f(A) = 0$. Dass es sogar ein Polynom vom Grad n mit $f(A) = 0$ gibt, nämlich das charakteristische Polynom $f = \mathrm{CP}_A$, besagt der Satz von Cayley-Hamilton. Und Beispiele zeigen, dass es noch besser geht:

$$A = \begin{pmatrix} \lambda & 1 & 0 \\ & \lambda & 1 \\ & & \lambda \end{pmatrix}, \qquad \mathrm{CP}_A(t) = (\lambda - t)^3$$

Für $f(t) = (\lambda - t)^2$ gilt

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

Satz 8.7.1. *Zu jedem nicht-trivialen Ideal $I \subseteq \mathbb{K}[t]$ gibt es ein eindeutiges normiertes Polynom $M(t)$, das jedes $f(t) \in I$ teilt.*

Beweis. Wir setzen
$$d := \min\{\text{grad}(f) \mid f \in I, f \neq 0\}.$$
Sei M zunächst irgendein Polynom $\neq 0$ in I mit minimalem Grad $= d$. Wir teilen jedes $f \in I$ mit Rest und erhalten
$$f(t) = q(t)M(t) + r(t), \qquad \text{grad}(r) < d.$$
Weil I ein Ideal ist, ist $q \cdot M \in I$, also auch $r = f - q \cdot M$. Damit hätte r kleineren Grad als M; es muss also $r = 0$ sein.
Die Eindeutigkeit von M erhält man durch die Normierung. □

Bemerkung 8.7.2. 1) Wir haben benutzt, dass $\mathbb{K}[t]$ ein Euklidischer Ring ist, d.h. ein kommutativer Ring mit 1, der einen euklidischen Algorithmus besitzt.

2) Und wir haben gezeigt, dass $\mathbb{K}[t]$ ein *Hauptidealring* ist, d.h. ein Ring, in welchem jedes Ideal von einem Element erzeugt werden kann.

Definition 8.7.3. Für eine Matrix $A \in \text{Mat}_{n,n}(\mathbb{K})$ sei das *Minimalpolynom* $\text{MP}_A(t) \in \mathbb{K}[x]$ der normierte Erzeuger des Auswertungsideals $I_A \subseteq \mathbb{K}[t]$ von A.
Analog definieren wir $\text{MP}_\varphi(t)$ für ein $\varphi \in \text{End}_\mathbb{K}(V)$, wobei V ein endlichdimensionaler Vektorraum sei.

Beispiel 8.7.4.
$$A = \begin{pmatrix} \lambda & 1 & 0 \\ & \lambda & 1 \\ & & \lambda \end{pmatrix}$$
Es ist $\text{CP}_A(t) = (\lambda - t)^3$, aber $\text{MP}_A(t) = (\lambda - t)^2$, weil dieses normierte Polynom A annuliert und kein lineares Polynom dies tut.

Satz 8.7.5. *Das charakteristische Polynom* $\text{CP}_A(t)$ *wird vom Minimalpolynom* $\text{MP}_A(t)$ *geteilt.*

Beweis. Folgt sofort aus der Definition un dem Satz von Cayley-Hamilton: $\text{CP}_A(t) \in I_A$. □

Bemerkung 8.7.6.
- Es gilt auch: $\text{CP}_A(t)$ teilt $\text{MP}_A(t)^n$, aber der Beweis benötigt fortgeschrittene Körpertheorie.
- $0 \leq \text{grad}(\text{MP}_A(t)) \leq n$.
- Jede Nullstelle von $\text{MP}_A(t)$ ist ein Eigenwert; jeder Eigenwert ist eine Nullstelle von $\text{MP}_A(t)$.

Beispiel 8.7.7. 1) $A = 0$: $\text{CP}_A(t) = (-1)^n t^n$, $\text{MP}_A(t) = 1$

2) $A = \mathbb{1}$: $\text{CP}_A(t) = (1-t)^n$, $\text{MP}_A(t) = (1-t)$

3) $A = D(\lambda_1, \dots, \lambda_n$, $\lambda_1, \dots, \lambda_n$ alle verschieden: $\text{CP}_A(t) = \prod_{i=1}^n (\lambda_i - t) = \text{MP}_A(t)$.

4) $A = D(\lambda_1, \dots, \lambda_n)$, $\lambda_1, \dots, \lambda_n$ beliebig: $\text{CP}_A(t) = \prod_{i=1}^n (\lambda_i - t) = \prod_{i=1}^k (\lambda'_i - t)^{m_i}$, wobei $\lambda'_1, \dots, \lambda'_k$ die verschiedenen Eigenwerte mit Vielfachheiten m_i seien; $\text{MP}_A(t) = \prod_{i=1}^k (\lambda'_i - t)$

Satz 8.7.8. *Ist A diagonalisierbar und*
$$\text{CP}_A(t) = \prod_{i=1}^k (\lambda_i - t)^{n_i}$$
mit den paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_k$ und $n_1 + \ldots + n_k = n$, so ist
$$\text{MP}_A(t) = \prod_{i=1}^k (t - \lambda_i).$$

Beweis. $\text{MP}_A(t)$ muss die gleichen Nullstellen wie $\text{CP}_A(t)$, nämlich die Eigenwerte, haben; aber es genügt jeweils Vielfachheit 1, weil dies schon A annulliert. □

8.8 Quotientenringe

Es sei R ein Ring mit 1 und $I \subseteq R$ ein (Links-/Rechts-)Ideal.

Definition 8.8.1. Der *Quotientenring* $\overline{R} = R/I$ ist als Menge zunächst die Menge der Äquivalenzklassen unter der Äquivalenzrelation

$$a \equiv b \mod I \iff a - b \in I.$$

Wir schreiben $a + I$ für die Äquivalenzklasse $\underline{a}$ von a.
Die *Addition* ist definiert durch

$$\underline{a} + \underline{b} := \underline{(a+b)}.$$

Die *Null* ist $0 = \underline{0} = 0 + I = I$, das *Negative* von $\underline{a}$ ist $-\underline{a} := \underline{-a}$.
Die *Multiplikation* ist definiert durch

$$\underline{a} \cdot \underline{b} := \underline{ab}.$$

Die *Eins* ist $1 := \underline{1}$.

Bemerkung 8.8.2.

1) Diese Operationen und Elemente sind wohldefiniert und statten $\overline{R}$ mit der Struktur eines Rings mit 1 aus.

2) Ist R kommutativ, so auch $\overline{R}$.

3) Enthält I ein invertierbares Element u, so ist auch $1 = u \cdot u^{-1} \in I$ und somit ist jedes Element von R in I enthalten, d.h. $I = R$. In diesem Fall ist $\overline{R} = 0$.

4) Ist $u \in R^\times$ eine Einheit, so auch $\underline{u}$ in $\overline{R}$ (falls $\overline{R} \neq 0$). Es ist auch möglich, dass für Nicht-Einheiten $u \in R$ trotzdem $\underline{u}$ eine Einheit in $\overline{R}$ ist; siehe Beispiele.

Beispiel 8.8.3.

(1) $R = \mathbb{Z}, I = n\mathbb{Z}$: $\overline{R} = \mathbb{Z}/n\mathbb{Z} = \mathbb{Z}/n$.
Für $n = p$ eine Primzahl ist $\overline{R} = \mathbb{F}_p$ sogar ein Körper.

(2) $R = \mathbb{K}[x]$,

$$I_1 = \langle x \rangle : \qquad \begin{array}{rcl} \overline{R} & \xrightarrow{\cong} & \mathbb{K} \\ f(x) = a_0 + a_1 x + \ldots + a_n x^n & \longmapsto & a_0 \end{array}$$

$$I_2 = \langle x^2 \rangle : \qquad \begin{array}{rcl} \overline{R} & \xrightarrow{\cong} & \mathbb{K} \oplus \mathbb{K} \\ f & \longmapsto & (a_0, a_1) \end{array}$$

mit der Multiplikation $(a_0, a_1) \cdot (b_0, b_1) = (a_0 b_0, a_0 b_1 + a_1 b_0)$ auf $\mathbb{K} \oplus \mathbb{K}$. Die beiden Polynome 1 und x werden zu Erzeugern $1 \mapsto (1, 0)$ und $x \mapsto (0, 1) = \underline{x}$ mit der Multiplikation $\underline{x}^2 = 0$.

(3) $R = \mathbb{R}[x]$,

$$I_3 = \langle x^2 + 1 \rangle : \quad \begin{array}{rcl} \overline{R} & \xrightarrow{\cong} & \mathbb{R} \oplus \mathbb{R} \cong \mathbb{C} \\ f & \longmapsto & (a_0, a_1), \end{array}$$

wie in (2), nun aber mit der Multiplikation

$$(a_0, a_1) \cdot (b_0, b_1) = (a_0 b_0 - a_1 b_1, a_0 b_1 + a_1 b_0).$$

Es gilt also $\underline{x}^2 = -1$, wegen $\underline{0} = \underline{x^2 + 1} = \underline{x}^2 + 1$. Also ist $\mathbb{R}[x]/\langle x^2 + 1 \rangle \cong \mathbb{C}$ der Körper der komplexen Zahlen.

(4) $R = \mathbb{Q}[x]$, $I_4 = \langle x^2 - 2 \rangle$

$$\begin{array}{rl} \overline{R} & \xrightarrow{\cong} \mathbb{Q}(\sqrt{2}) \subseteq \mathbb{R} \\ f & \longmapsto \underbrace{a_0 + a_1\sqrt{2} + 2a_2 + 2a_3\sqrt{2} + 4a_4 + \ldots}_{= \sum_{i \geq 0} 2^i a_{2i} + \sum_{i \geq 0} 2^i a_{2i+1} \sqrt{2}}, \end{array}$$

Insbesondere

$$x \longmapsto \sqrt{2}, \qquad x^2 q \longmapsto 2^q, \qquad x^{2q+1} \longmapsto 2^q \sqrt{2}.$$

(5) $R = \mathbb{K}[x]$, $f(x) = a_0 + a_1 x + \ldots + a_{n-1}x^{n-1} + x^n$ normiert

$$A = \begin{pmatrix} 0 & & & & & -a_0 \\ 1 & 0 & & & & -a_1 \\ & 1 & 0 & & & -a_2 \\ & & \ddots & \ddots & & \vdots \\ & & & 1 & 0 & -a_{n-2} \\ & & & & 1 & -a_{n-1} \end{pmatrix} \quad \text{Begleitmatrix von } f$$

- $\mathrm{CP}_A(x) = f(x)$ charakteristisches Polynom
- $\mathrm{MP}_A(x) = p_0 + p_1 x + \ldots + p_{m-1}x^{m-1} + x^m$ Minimalpolynom
- $I_A =$ Auswertungsideal $= \{\varphi(x) \in \mathbb{K}[x] \mid \varphi(A) = 0\} = \ker(\mathrm{eval}_A)$
- $\mathbb{K}[A] = \mathrm{im}(\mathrm{eval}_A) = \{\varphi(A) \in \mathrm{Mat}_{n,n}(\mathbb{K}) \mid \varphi \in \mathbb{K}[x]\}$

$$\mathbb{K}[x]/I_A \xrightarrow{\cong} \mathbb{K}[A]$$

$$\left\{ \underbrace{\underline{1} \mapsto \mathbb{1},\ \underline{x} \mapsto A}_{\text{Algebraerzeuger}},\ \ldots,\ \underline{x}^{m-1} \mapsto A^{m-1} \right\} \quad \text{Basis für Vektorraum}$$

Übungsaufgabe.

Die *kanonische Quotientenabbildung*

$$\pi : R \longrightarrow \overline{R} = R/I, \qquad \pi(a) = \underline{a} = a + I$$

ist ein surjektiver Ringhomomorphismus mit folgender *universellen Eigenschaft*:

Satz 8.8.4. *Für jeden Ringhomomorphismus $f : R \to R'$ gilt:*
Es gibt genau dann einen Ringhomomorphismus $\overline{f} : \overline{R} \to R'$

$$\begin{tikzcd} R \arrow[r, "f"] \arrow[d, "\pi"'] & R' \\ \overline{R} = R/I \arrow[ur, "\overline{f}"'] & \end{tikzcd}$$

mit $f = \overline{f} \circ \pi$, falls $\ker(f) \supseteq I$.

Beweis. Man setzt $\overline{f}(\underline{a}) = \overline{f}(a + I) = f(a)$; dies ist genau dann wohldefiniert, wenn $\ker(f) \supseteq I$ gilt. □

Satz 8.8.5. *Jeder kommutative Ring R ist eine Quotientenring eines Polynomrings $\mathbb{Z}[\mathcal{X}]$, wobei $\mathcal{X} \ni x$ eine Menge von Variablen sei.*

Beweis. Es sei $\mathcal{E} \subseteq R$ ein Erzeugendensystem von R, d.h. jedes Element $a \in R$ kann als ganzzahlige Linearkombination von Produkten von Potenzen e^n von Elementen $e \in \mathcal{E}$ geschrieben werden (Notfalls muss man $\mathcal{E} = R$ nehmen). Wir indizieren $\mathcal{E} = (e_i \mid I)$ durch die Indexmenge I.
Für eine endliche Menge $J \subseteq I$ und natürliche Zahlen n_j für $j \in J$ schreiben wir $\underline{n}_J = (n_j \mid j \in J)$ und

$$e_J^{\underline{n}} = e_{j_1}^{n_{j_1}} \cdot \ldots \cdot e_{j_k}^{n_{j_k}}, \qquad \text{falls } J = (j_1, \ldots, j_k)$$

Nun nehmen wir für jedes $i \in I$ eine Unbekannte x_i und betrachten den Polynomring $\tilde{R} = \mathbb{Z}[(x_i \mid i \in I)]$ in den Unbekannten x_i. Ein Polynom $f \in \tilde{R}$ hat die Form

$$f = \sum_{J \subseteq I} a_J x_J^{\underline{n}_J}$$

mit $a_J \in \mathbb{Z}$ und $a_J = 0$ für fast alle $J \subseteq I$, wobei

$$x_J^{\underline{n}_J} := x_{j_1}^{n_{j_1}} \cdot \ldots \cdot x_{j_k}^{n_{j_k}}$$

Dann ist $\Phi : f \mapsto \sum_J a_J e_J^{\underline{n}_J}$ ein surjektiver Ringhomomorphismus und

$$R \cong \mathbb{Z}[(x_i \mid i \in I)]/\ker\Phi.$$

□

Neue Nullstellen

Sei $\mathbb{K}$ Körper.

$$\begin{aligned}&\mathbb{K}[x] \ni f(x) = a_0 + a_1 x + \ldots + a_n x^n \quad \text{Polynom}\\ &\mathbb{K}[x] \supseteq I := \langle f \rangle = \{\lambda_0 + \lambda_1 f + \ldots + \lambda_q f^q\} \quad \text{von } f \text{ erzeugtes Ideal.}\end{aligned}$$

Durch Bilden des Quotienten erhalten wir einen neuen Ring R:

$$\begin{array}{ccl}
\mathbb{K}[x] & \xrightarrow{\pi} & R := \mathbb{K}[x]/I \subseteq R[X] \\
\cup\!\vert & & \cup\!\vert \\
\mathbb{K} & & Q := \pi(\mathbb{K}) \\
c & \longmapsto & \underline{c}, \quad c \in \mathbb{K} \subseteq \mathbb{K}[x] \text{ konstant} \\
x & \longmapsto & \underline{x} \\
x^k & \longmapsto & \underline{x}^k = \underline{x^k} \\
f & \longmapsto & \underline{f} = 0
\end{array}$$

Wir erhalten ein neues Polynom $\Phi(X) \in R[X]$

$$\Phi(X) := \pi_{\#}(f(x)) := \underline{a_0} + \underline{a_1} X + \underline{a_2} X^2 + \ldots + \underline{a_n} X^n$$

Satz 8.8.6. $\Phi(X)$ *hat mindestens eine Nullstelle, nämlich* $\lambda = \underline{x} \in R$.

Beweis.

$$\begin{aligned}
\Phi(\underline{x}) &= \underline{a_0} + \underline{a_1}\underline{x} + \underline{a_2}\underline{x}^2 + \ldots + \underline{a_n}\underline{x}^n \\
&= \underline{a_0} + \underline{a_1 x} + \underline{a_2 x^2} + \ldots + \underline{a_n x^n} \\
&= \underline{a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n} \\
&= \underline{f(x)} \\
&= \underline{0} \\
&= 0.
\end{aligned}$$

□

9 Normalformen

9.1 Normalformenprobleme

Was ist ein Normalformenproblem? - Es geht darum, für eine Menge von Matrizen eine Äquivalenzrelation zu definieren und für jede Äquivalenzklasse einen (besonderen) Repräsentanten zu finden.

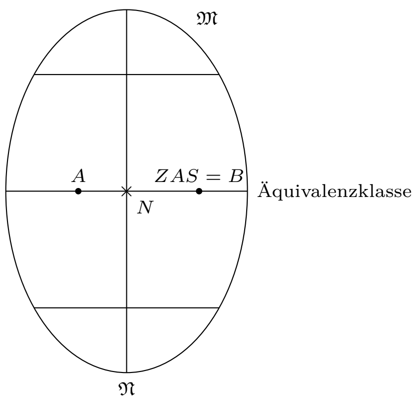
- $\mathfrak{M} \subseteq \mathrm{Mat}_{m,n}(\mathbb{K})$ Matrizen
- $\Gamma \leq \mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})^{\mathrm{op}}$ Gruppe: Für $(Z_1, S_1), (Z_2, S_2) \in \Gamma$ ist die Multiplikation gegeben durch

$$(Z_2, S_2) \cdot (Z_1, S_1) := (Z_2 Z_1, S_1 S_2)$$

  (Für eine Gruppe G definiert man die *oppositionelle Gruppe* G^{op} wie folgt: Als Menge ist $G^{\mathrm{op}} = G$ mit Verknüpfung $A \cdot B := BA$.)
  Definiere eine Äquivalenzrelation durch $A \sim_\Gamma B :\iff \exists \gamma = (Z, S) \in \Gamma : ZAS = B$.

  Fragen

  1. Wie sieht die Menge der Äquivalenzklasse aus? (Wie viele Klassen? Wie groß sind sie? ...)
  2. Gibt es in jeder Klasse einen besonderen Repräsentanten? (Repräsentantensystem: $\mathfrak{N} \subseteq \mathfrak{M}$, so dass jede Äquivalenzklasse $[A]$ die Menge $\mathfrak{N}$ nur einmal schneidet; dieser besondere Repräsentant $N = [A] \cap \mathfrak{N}$ würden wir die Normalform von A nennen.)

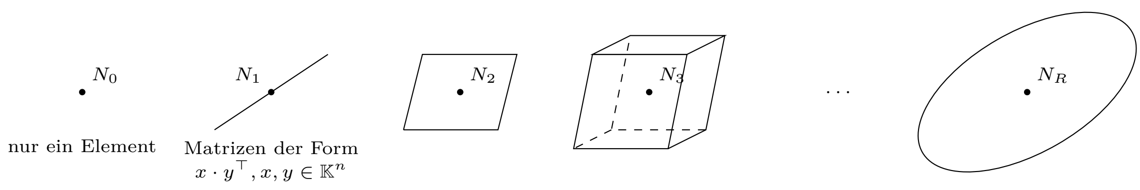


Beispiel 9.1.1. „Gleichrangigkeit“

$$\mathfrak{M} = \mathrm{Mat}_{m,n}(\mathbb{K}) \qquad \Gamma = \mathrm{GL}_m(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})^{\mathrm{op}}$$

$$A \sim_\Gamma B \iff \mathrm{rg}(A) = \mathrm{rg}(B) = r$$

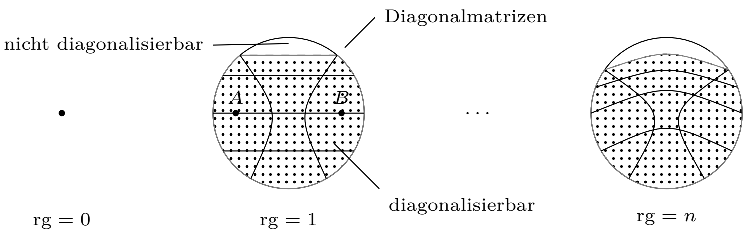
- Es gibt für $r = 0, 1, 2, \ldots, \min\{m, n\} =: R$ je eine Klasse
- $\mathfrak{N} = \{N_0, N_1, \ldots, N_R\}$, $N_r = \left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right)$ sind die Normalformen.
  Klassen:



Beispiel 9.1.2 (Ähnlichkeit)**.**

$$\mathfrak{M} = \mathrm{Mat}_{n,n}(\mathbb{K}) \qquad \Gamma = \{(\Omega, \Omega^{-1}) \mid \Omega \in \mathrm{GL}_n(\mathbb{K})\}$$

$$A \sim_\Gamma B \iff \exists \Omega \in \mathrm{GL}_n(\mathbb{K}) : \Omega A \Omega^{-1} = B \quad \text{Basiswechsel für lineare Abbildungen}$$



Beispiel 9.1.3 (Kongruenz)**.**

$$\mathfrak{M} = \mathrm{Mat}_{n,n}(\mathbb{K}) \qquad \{(\Omega, \Omega^\top) \mid \Omega \in \mathrm{GL}_n(\mathbb{K})\}$$

$$A \sim_\Gamma B \iff \exists \Omega \in \mathrm{GL}_n(\mathbb{K}) : \Omega A \Omega^\top = B \quad \text{Basiswechsel für Bilinearformen}$$

9.2 Nilpotente Endomorphismen

Definition 9.2.1. Ein $\varphi \in \mathrm{End}_\mathbb{K}(V)$ heißt *nilpotent*, falls $\varphi^d = 0$ für ein $d \in \mathbb{N}$ gilt. Das kleinste solche d heißt *(Nilpotenz)index* von φ.
Analog heißt eine Matrix $N \in \mathrm{Mat}_{n,n}(\mathbb{K})$ *nilpotent*, wenn $N^d = 0$ gilt für ein $d \in \mathbb{N}$.

Beispiel 9.2.2. 1) $N = 0$ ist nilpotent mit Index $d = 1$.

2)

$$N = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 \\ 0 & \cdots & \cdots & 0 \end{pmatrix}, \qquad d = n$$

3)

$$N = \begin{pmatrix} 0 & \cdots & 0 & * & \cdots & * \\ \vdots & \ddots & & \ddots & \ddots & \vdots \\ & & & & \ddots & * \\ & & & & & 0 \\ \vdots & & & & \ddots & \vdots \\ 0 & \cdots & & & \cdots & 0 \end{pmatrix}$$

4) Die Summe $(A + B)$ und das Produkt AB zweier nilpotenter Matrizen müssen nicht nilpotent sein.

Lemma 9.2.3. *Falls A und B vertauschen, so gilt*

(1)

$$(A+B)^m = \sum_{k=0}^{m} \binom{m}{k} A^k B^{m-k} \quad \textit{und}$$

(2)

$$(AB)^m = A^m B^m.$$

Damit erhält man:

Proposition 9.2.4. *Seien A und B vertauschbare Matrizen.*

(i) Ist A oder B nilpotent mit Index d_A bzw. d_B, so ist AB nilpotent mit Index höchstens d_A bzw. d_B.

(ii) Sind A und B nilpotent mit Index d_A bzw. d_B, so ist $A+B$ nilpotent mit Index höchstens d_A+d_B.

Das wichtigste Kriterium für Nilpotenz ist folgender Satz.

Satz 9.2.5.

(i) Ein Endomorphismus eine endlichdimensionalen Vektorraums ist genau dann nilpotent, wenn er durch eine echte obere Dreiecksmatrix dargestellt werden kann.

(ii) Eine Matrix ist genau dann nilpotent, wenn sie zu einer echten oberen Dreiecksmatrix ähnlich ist.

Beweis. Übung. □

Korollar 9.2.6. *Für eine nilpotente Matrix N gilt:*

(1) $\mathrm{CP}_N(t) = (-1)^n t^n$.

(2) $\mathrm{MP}_N(t) = t^d$, $d =$ *Nilpotenzindex.*

(3) $\mathrm{Det}(N) = 0$.

(4) $\mathrm{Spur}(N) = 0$.

Die Addition einer vertauschbaren nilpotenten Matrix N zu A ändert das charakteristische Polynome nicht:

Satz 9.2.7. *Sei* $\mathrm{char}(\mathbb{K}) = 0$. *Sind A und N vertauschbare Matrizen und N nilpotent, so gilt:*

(1) $\mathrm{CP}_{A+N}(t) = \mathrm{CP}_A(t)$,

(2) $\mathrm{Det}(A+N) = \mathrm{Det}(A)$,

(3) $\mathrm{Spur}(A+N) = \mathrm{Spur}(A)$.

Beweis. Es ist nur (1) zu zeigen; dann gilt $c_k(A+N) = c_k(A)$ für alle Koeffizienten des charakteristischen Polynoms.

1) Da $\mathbb{K}$ unendlich viele Elemente hat, können nicht alle Nullstellen von $\mathrm{CP}_A(t)$ sein; sei also $\mu \in \mathbb{K}$ so gewählt, dass $\mathrm{CP}_A(\mu) = \mathrm{Det}(A - \mu\mathbb{1}) \neq 0$ gilt und somit $\mu\mathbb{1} - A$ invertierbar ist. Wir setzen

$$B := (\mu\mathbb{1} - A)^{-1}.$$

2) *Behauptung*: BN ist nilpotent.
Diese Behauptung folgt aus der Proposition, wenn B und N vertauschen. Zunächst ist

$$(\mu\mathbb{1} - A)N = \mu N - AN = N\mu - NA = N(\mu\mathbb{1} - A).$$

Aber aus $B^{-1}N = NB^{-1}$ folgt sofort

$$NB = BB^{-1}NB = BNB^{-1}B = BN.$$

3) *Behauptung*: $\mathrm{Det}(\mathbb{1} - BN) = 1$.
Nach Satz 9.2.5 ist das nilpotente BN ähnlich zu einer echten oberen Dreiecksmatrix $T = \Omega BN\Omega^{-1}$. Damit folgt

$$\begin{aligned}\mathrm{Det}(\mathbb{1} - BN) &= \mathrm{Det}(\mathbb{1} - \Omega^{-1}T\Omega)\\ &= \mathrm{Det}(\Omega^{-1}(\mathbb{1} - T)\Omega) = \mathrm{Det}(\mathbb{1} - T) = 1.\end{aligned}$$

4) Für das oben gewählte $\mu \in \mathbb{K}$ rechnen wir nach:

$$\begin{aligned}
\mathrm{CP}_{A+N}(\mu) &= (-1)^n \operatorname{Det}(\mu \mathbb{1} - (A+N)) \\
&= (-1)^n \operatorname{Det}(\underbrace{(\mu \mathbb{1} - A)(\mathbb{1} - BN)}_{(\mu\mathbb{1}-A) - \underbrace{(\mu\mathbb{1}-A)B}_{\mathbb{1}} N}) \\
&= (-1)^n \operatorname{Det}(\mu \mathbb{1} - A) \underbrace{\operatorname{Det}(\mathbb{1} - BN)}_{=1} \\
&= \mathrm{CP}_A(\mu)
\end{aligned}$$

Die Polynome $\mathrm{CP}_{A+N}(t)$ und $\mathrm{CP}_A(t)$ stimmen für die unendlich vielen Nicht-Nullstellen $t = \mu$ überein; also sind sie gleich. □

Bemerkung 9.2.8. Die analoge Aussage gilt nicht für das Minimalpolynom:

$$\begin{aligned}
A &= \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} & N &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\
\mathrm{CP}_A(t) &= (\lambda - t)^2 & \mathrm{CP}_{A+N}(t) &= (\lambda - t)^2 \\
\mathrm{MP}_A(t) &= -(\lambda - t) & \mathrm{MP}_{A+N}(t) &= (\lambda - t)^2
\end{aligned}$$

9.3 Hauptraumzerlegung

Wie wir wissen, ist ein Hindernis für die Diagonalisierbarkeit einer Matrix die Diskrepanz zwischen der geometrischen und der algebraischen Vielfachheit eines Eigenwerts, d.h. die Dimension des Eigenraumes ist zu klein. Man definiert deshalb zur weiteren Untersuchung die verallgemeinerten Eigenräume wie folgt:

Definition 9.3.1. Für einen Endomorphismus $\varphi : V \to V$ mit Eigenwert λ mit Vielfachheit

$$r = \operatorname{algV}(\varphi, \lambda) \geq 1$$

sei

$$\mathrm{Hau}_\lambda(\varphi) := \ker((\varphi - \lambda \operatorname{id}_V)^r)$$

der *Hauptraum*.

Es ist

$$\mathrm{Hau}_\lambda(\varphi) = \ker((\varphi - \lambda \operatorname{id}_V)^r) \supseteq \ker((\varphi - \lambda \operatorname{id}_V)^{r-1}) \supseteq \ldots \supseteq \ker((\varphi - \lambda \operatorname{id}_V)) = \mathrm{Eig}_\lambda(\varphi).$$

Beispiel 9.3.2.

$$A = \left(\begin{array}{ccccccc|c}
\multicolumn{7}{c}{\overbrace{\hphantom{\lambda \quad * \quad \cdots \quad * \quad \lambda \quad \ddots \quad \lambda}}^{r}} & \\
\multicolumn{4}{c}{} & \multicolumn{3}{c}{\overbrace{\hphantom{\lambda \quad \ddots \quad \lambda}}^{d}} & \\
\lambda & * & \cdots & * & & & & \\
& \ddots & \ddots & \vdots & & & & \\
& & \ddots & * & & 0 & & B \\
& & & \lambda & & & & \\
& & & & \lambda & & & \\
& & & & & \ddots & & \\
0 & & & & & & \lambda & \\
\hline
& & & 0 & & & & C
\end{array}\right)$$

$$\operatorname{algV}(A, \lambda) = r, \qquad \operatorname{geomV}(A, \lambda) = d, \qquad \lambda \notin \operatorname{Spek}(C).$$

$$\mathrm{Hom}_\lambda(\varphi) = \operatorname{Span}(e_1, \ldots, e_n), \qquad \mathrm{Eig}_\lambda(\varphi) = \operatorname{Span}(e_{r-d+1}, \ldots, e_r)$$

Der Hauptraum als Kern einer bestimmten Potenz von $\psi = \varphi - \lambda \operatorname{id}_V$ ist nur ein Spezialfall für folgende Überlegungen. Sei
$$\psi : V \longrightarrow V$$
ein beliebiger Endomorphismus, $n = \dim V$. Dann haben wir mit den Potenzen ψ^k $(k = 0, 1, \dots, n)$ eine aufsteigende, unvollständige Fahne von Kernen
$$0 \subseteq \ker \psi \subseteq \ker \psi^2 \subseteq \dots \subseteq \ker \psi^k \subseteq \dots \subseteq V \tag{9.1}$$
und eine absteigende, unvollständige Fahne von Bildern
$$V \supseteq \operatorname{im} \psi \supseteq \operatorname{im} \psi^2 \supseteq \dots \supseteq \operatorname{im} \psi^k \supseteq \dots \supseteq 0. \tag{9.2}$$
Es gilt nach der Dimensionsformel
$$\dim(\ker \psi^k) + \dim(\operatorname{im} \psi^k) = n, \tag{9.3}$$
aber nicht notwendigerweise
$$\ker \psi^k \cap \operatorname{im} \psi^k = 0,$$
also eventuell sind $\ker \psi^k$ und $\operatorname{im} \psi^k$ nicht komplementär.

Beispiel 9.3.3.
$$A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad \ker A = \operatorname{im} A = \mathbb{K} e_1.$$

Da die Dimension von V endlich ist, können die beiden Ketten nicht endlos auf- bzw. absteigen. Es gibt also zwei Zahlen
$$\begin{aligned} d &:= \min\{k \geq 0 \mid \ker(\psi^k) = \ker(\psi^{k+1})\}, \\ d' &:= \min\{k \geq 0 \mid \operatorname{im}(\psi^k) = \operatorname{im}(\psi^{k+1})\}. \end{aligned} \tag{9.4}$$
Wegen (9.3) ist natürlich $d = d'$. Und ab dem Index $k = d$ sind die Ketten stabil, d.h.
$$\begin{aligned} \ker(\psi^{d+i}) &= \ker(\psi^{d+i+1}), \\ \operatorname{im}(\psi^{d+i}) &= \operatorname{im}(\psi^{d+i+1}), \quad \text{für alle } i \geq 0. \end{aligned} \tag{9.5}$$
Denn wäre für ein $v \in V$ zwar $\psi^{d+i+1}(v) = 0$, aber $\psi^{d+i}(v) \neq 0$, so wäre mit $w := \psi^i(v) \neq 0$ gefunden mit $w \in \ker(\psi^{d+1})$, aber $w \notin \ker(\psi^d)$, was der Minimalität von d widerspräche. Die Stabilität der Kette der Bilder folgt wegen (9.3).

Wir setzen
$$K := \ker(\psi^d), \qquad B := \operatorname{im}(\psi^d). \tag{9.6}$$
Offenbar sind beide Unterräume ψ-invariant:
$$\psi(K) = K, \qquad \psi(B) = B. \tag{9.7}$$
Wir schreiben für die Einschränkungen
$$\psi_K : K \longrightarrow K, \qquad \psi_B : B \longrightarrow B. \tag{9.8}$$

Lemma 9.3.4.

(i) $\psi_K : K \to K$ *ist nilpotent mit Index* d.

(ii) $\psi_B : B \to B$ *ist ein Isomorphismus.*

(iii) $V \cong K \oplus B$.

(iv) $\dim K = \operatorname{algV}(\psi, 0) \geq d$.

Beweis. (i)
$$\psi_K^d(v)) = \psi^d(v) = 0 \qquad \text{für } v \in K = \ker(\psi^d).$$
Und wäre $\psi_k^{d-1} = 0$, so wäre $K \subseteq \ker(\ker \psi^{d-1})$, also $\ker(\psi^d) = \ker(\psi^{d-1})$. Damit wäre d nicht minimal in (9.3). Also ist der Nilpotenzindex genau d.

(ii) Offenbar ist ψ_B surjektiv nach Definition von B und d. Also ist es auch injektiv.

(iii) Ist $v \in K \cap B$, so ist $\psi^d(v) = 0$ und $v = \psi^d(w)$ für ein $w \in V$. Damit folgt $\psi^{2d}(w) = 0$, also $w \in \ker(\psi^{2d}) = \ker(\psi^d)$ nach Definition von d. Also ist $v = \psi^d(w) = 0$.

(iv) Da die Kette der Kerne (9.1) bis zu $\ker(\psi^d)$ strikt aufsteigt, gilt $m := \dim(K) \geq d$.
Für das charakteristische Polynom haben wir wegen der direkten Summenzerlegung $\psi = \psi_K \oplus \psi_B$:
$$\begin{aligned} \mathrm{CP}_\psi(t) &= \mathrm{CP}_{\psi_K}(t) \cdot \mathrm{CP}_{\psi_B}(t) \\ &= (-1)^m t^m \cdot \mathrm{CP}_{\psi_B}(t), \end{aligned}$$
weil ψ_K nilpotent ist. Weil ψ_B ein Isomorphismus ist, ist $\mathrm{CP}_{\psi_B}(0) = \mathrm{Det}(\psi_B) \neq 0$, also 0 keine Nullstelle von $\mathrm{CP}_{\psi_B}(t)$. Damit folgt $m = \mathrm{algV}(\psi, 0$.

□

Ist $\mathcal{B}_K$ eine Basis von K, $\mathcal{B}_B$ eine Basis von B, so ist $\mathcal{B} = \mathcal{B}_K \sqcup \mathcal{B}_B$ eine Basi von V. Und in dieser Basis ist
$$M_{\mathcal{B}\mathcal{B}}(\psi) = \left(\begin{array}{c|c} N & 0 \\ \hline 0 & C \end{array}\right)$$
mit einer nilpotenten $m \times m$-Matrix N vom Index d, und einer invertierbaren $(n-m) \times (n-m)$-Matrix C.

Beispiel 9.3.5. (1) ψ bijektiv: $K = 0$, $B = V$, $d = 0$, $m = 0$.

(2) ψ nilpotent: $K = V$, $B = 0$, $d = n$, $m = n$.

(3) $\mathrm{rg}(\psi) \geq n - m$.

(4)
$$A = \left(\begin{array}{ccc|c} 0 & 1 & 0 & \\ & 0 & 1 & 0 \\ & & 0 & \\ \hline & 0 & & C \end{array}\right)$$
$$K = \mathrm{Span}(e_1, e_2, e_3), \quad m = 3, \quad d = 2, \quad \mathrm{rg}(A) = n - 1.$$

Satz 9.3.6 (Hauptraumzerlegung)**.** *Es sei $\varphi : V \to V$ ein Endomorphismus, $n = \dim V$, mit zerfallendem charakteristischem Polynom*
$$\mathrm{CP}_\varphi(t) = (\lambda_1 - t)^{m_1} \cdot \ldots \cdot (\lambda_k - t)^{m_k},$$
wobei die $\lambda_1, \ldots, \lambda_k$ paarweise verschieden sind und $m_1 + \ldots + m_k = n$. Es sei
$$V_i := \mathrm{Hau}_{\lambda_i}(\varphi) = \ker((\varphi - \lambda_i t)^{m_i})$$
der Hauptraum zum Eigenwert λ_i $(i = 1, \ldots, k)$. Dann gilt:

(i) $\varphi(V_i) \subseteq V_i$ *und* $\dim V_i = m_i$.

(ii) $V = V_1 \oplus \ldots \oplus V_k$ *ist die direkte Summe der Haupträume.*

(iii) φ *ist eine Summe* $\varphi = \varphi_N + \varphi_D$, *wobei* φ_N *nilpotent ist,* φ_D *diagonalisierbar und beide vertauschen, d.h.* $\varphi_D \circ \varphi_N = \varphi_N \circ \varphi_D$.

Wählt man Basen $\mathcal{B}_i$ in V_i und setzt $\mathcal{B} = \mathcal{B}_1 \sqcup \ldots \sqcup \mathcal{B}_k$, so ist die zugehörige Matrix
$$M_{\mathcal{B}\mathcal{B}}(\varphi) = A = \begin{pmatrix} \lambda_1 \mathbb{1} + N_1 & & \\ & \ddots & \\ & & \lambda_r \mathbb{1} + N_r \end{pmatrix}$$

eine Blockmatrix mit $m_i \times m_i$-Blöcken und

$$\lambda_i \mathbb{1} + N_i = \begin{pmatrix} \lambda_i & & * \\ & \ddots & \\ & & \lambda_i \end{pmatrix}, \qquad N_i = \begin{pmatrix} 0 & & * \\ \vdots & \ddots & \\ 0 & \cdots & 0 \end{pmatrix},$$

und

$$M_{\mathcal{B}\mathcal{B}}(\varphi_D) = D(\underbrace{\lambda_1, \dots, \lambda_1}_{m_1}, \dots, \underbrace{\lambda_k, \dots, \lambda_k}_{m_k}) = \begin{pmatrix} \lambda_1 \mathbb{1} & & \\ & \ddots & \\ & & \lambda_k \mathbb{1} \end{pmatrix} = D \quad \text{(diagonal)}$$

$$M_{\mathcal{B}\mathcal{B}}(\varphi_N) = \begin{pmatrix} N_1 & & \\ & \ddots & \\ & & N_k \end{pmatrix} = N \quad \text{(nilpotent, obere Dreiecksmatrix).}$$

Anders gesagt: *Jede Matrix A mit zerfallendem charakteristischen Polynom ist ähnlich zu einer Blockmatrix $D+N$ mit D diagonal, N nilpotent und $DN = ND$.*

Beweis. Wir machen eine Induktion über k, $1 \leq k \leq n$.
Zum ersten Eigenwert λ_1 definieren wir

$$\psi = \varphi - \lambda_1 \operatorname{id}_V : V \longrightarrow V.$$

Das charakteristische Polynom ändert sich wie folgt:

$$\mathrm{CP}_\psi(t - \lambda_1) = \mathrm{CP}_\varphi(t),$$

also $\operatorname{algV}(\psi, 0) = \operatorname{algV}(\varphi, \lambda_1) = m_1$.
In der Notation des Lemmas ist $K = \operatorname{Hau}_{\lambda_1}(\varphi) = \ker \psi^d$, mit $d \leq m_1$, und wir haben mit $B = \operatorname{im} \psi^d$ die Zerlegung

$$V = \operatorname{Hau}_{\lambda_1}(\varphi) \oplus B.$$

Wir setzen $V_1 := \operatorname{Hau}_{\lambda_1}(\varphi)$ und notieren $\dim(V_1)) = m_1$ nach Lemma 9.3.4.
Beide Summanden K und B sind ψ-invariant, damit aber auch invariant unter $\varphi = \psi + \lambda_1 \operatorname{id}_V$.
Für $\varphi_B = \varphi|_B : B \to B$ gilt

$$\mathrm{CP}_{\varphi_B}(t) = (\lambda_2 - t)^{m_2} \cdot \ldots \cdot (\lambda_k - t)^{m_k}.$$

Damit können wir auf φ_B die Induktionsannahme anwenden, weil φ_B nur noch $k-1$ verschiedene Eigenwerte hat. Dies beweist die Aussagen (i) und (ii).
Die Aussage (iii) folgt nun auch, weill jedes

$$\varphi_i = \varphi|_{V_i} : V_i \longrightarrow V_i$$

ein zerfallendes charateristisches Polynom hat, also trigonalisierbar ist. Wählt man also Basen $\mathcal{B}_i$ in V_i, so hat φ_i obere Dreiecksgestalt,

$$M_{\mathcal{B}_i\mathcal{B}_i}(\varphi_i) = \begin{pmatrix} \lambda_i & & * \\ & \ddots & \\ & & \lambda_i \end{pmatrix} = D_i + N_i$$

mit $D_i = D(\lambda_i, \dots, \lambda_i) = \lambda_i \mathbb{1}_{m_i}$ diagonal und $N_i = \begin{pmatrix} 0 & & * \\ \vdots & \ddots & \\ 0 & \cdots & 0 \end{pmatrix}$ eine echte obere Dreiecksmatrix, also nilpotent.
Man setzt

$$\varphi_D = L_{\mathcal{B}\mathcal{B}}(D_1 \oplus \ldots \oplus D_k),$$
$$\varphi_N = L_{\mathcal{B}\mathcal{B}}(N_1 \oplus \ldots \oplus N_k),$$

wobei $\mathcal{B} = \mathcal{B}_1 \sqcup \ldots \sqcup \mathcal{B}_k$ eine Basis von V ist. Offenbar gilt auch $\varphi_D \circ \varphi_N = \varphi_N \circ \varphi_D$. □

9.4 Jordansche Normalform

Unsere nächste Normalform ist die vielleicht wichtigste Normalform für komplexe oder reelle Matrizen. Als Verbesserun der Hauptraumzerlegun wollen wir jetzt durch geschickte Wahl einer Basis den nilpotenten Teil in eine sehr einfache Gestalt bringen, die sich aus sogenannten Jordan-Blöcken zusammensetzt:

Definition 9.4.1. Für ein $d \geq 1$ und ein $\lambda \in \mathbb{K}$ heißt die $d \times d$-Matrix

$$J_d(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$$

Jordan-Block der Größe d zum Eigenwert λ.

Offenbar gilt für $J = J_d(\lambda)$:

(1) $\mathrm{Spek}(J) = \{\lambda\}$.

(2) J ist invertierbar genau dann, wenn $\lambda \neq 0$.

(3) $\mathrm{Det}(J) = \lambda^d$, $\mathrm{Spur}(J) = d\lambda$.

(4) $\mathrm{CP}_J(t) = (\lambda - t)^d = \mathrm{MP}_J(t)$.

(5) $\mathrm{Eig}_\lambda(J) = \mathbb{K}e_1$, $\mathrm{geomV}(J, \lambda) = 1$.

(6) $\mathrm{Hau}_\lambda(J) = \mathbb{K}^d$, $\mathrm{algV}(J, \lambda) = d$.

(7) $J_d(\lambda) = D(\lambda, \ldots, \lambda) + J_d(0)$ und $J_d(0)$ ist nilpotent mit Index d.

Wir schreiben $J_d = J_d(0)$ und beachten $J_1 = (0)$.
Die $n \times n$-Matrix

$$N = N(s_d, \ldots, s_1) = \begin{pmatrix} \boxed{\begin{matrix} J_d & & \\ & \ddots & \\ & & J_d \end{matrix}} & & & & \\ & \boxed{\begin{matrix} J_{d-1} & & \\ & \ddots & \\ & & J_{d-1} \end{matrix}} & & & \\ & & \ddots & & \\ & & & \boxed{\begin{matrix} J_2 & & \\ & \ddots & \\ & & J_2 \end{matrix}} & \\ & & & & \boxed{\begin{matrix} J_1 & & \\ & \ddots & \\ & & J_1 \end{matrix}} \end{pmatrix},$$

mit s_i Jordanblöcken J_i $(i = 1, \ldots, d)$, besteht aus

$$s = s_d + s_{d-1} + \ldots + s_2 + s_1$$

Jordanblöcken unterschiedlicher Größen, und es muss gelten

$$n = ds_d + (d-1)s_{d-1} + \ldots + 2s_2 + s_1.$$

Es gilt:

(1) N ist nilpotent mit Index d.

(2) $\mathrm{Spek}(N) = \{0\}$.

(3) $\mathrm{CP}_N(t) = (-1)^n t^n$, $\mathrm{algV}(N, 0) = n$.

(4) $\mathrm{Eig}_0(N) = \mathbb{K}e_1 + \mathbb{K}e_{d+1} + \ldots$, $\mathrm{geomV}(N, 0) = s$.

(5) $\mathrm{MP}_N(t) = t^d$ mit $d = \max\{k \mid s_k \neq 0\}$.

(6) $\mathrm{Hau}_0(N) = \mathbb{K}^n$.

Unser Ziel ist eine Normalform aus Jordan-Blöcken für Matrizen oder Endomorphismen mit zerfallendem charakteristischen Polynom. Wir beginnen mit einem nilpotenten Endomorphismus.

Lemma 9.4.2. *Es sei $\varphi : V \to V$ nilpotent vom Index d und sei $n = \dim V$. Dann gibt es eine Basis $\mathcal{B}$ von V, so dass*

$$M_{\mathcal{B}\mathcal{B}}(\varphi) = \begin{pmatrix} \boxed{\begin{matrix} J_d & & \\ & \ddots & \\ & & J_d \end{matrix}} & & & & \\ & \boxed{\begin{matrix} J_{d-1} & & \\ & \ddots & \\ & & J_{d-1} \end{matrix}} & & & \\ & & \ddots & & \\ & & & \boxed{\begin{matrix} J_2 & & \\ & \ddots & \\ & & J_2 \end{matrix}} & \\ & & & & \boxed{\begin{matrix} J_1 & & \\ & \ddots & \\ & & J_1 \end{matrix}} \end{pmatrix}$$

eine Blockmatrix mit s_k Jordan-Blöcken $J_k = J_k(0)$ (möglicherweise) unterschiedlicher Größen ist. Die Zahlen $s_1, \ldots, s_d \geq 0$ sind eindeutig bestimmt und erfüllen

$$ds_d + (d-1)s_{d-1} + \ldots + 2s_2 + s_1 = n.$$

Beweis. Wir schauen uns die Kerne

$$K_i := \ker(\varphi^i) \tag{9.9}$$

an und die strikt aufsteigende Kette

$$\begin{tikzcd}[column sep=small] 0 = K_0 & \subseteq & K_1 \arrow[dll, "\varphi"] & \subseteq & \ldots & \subseteq & K_{i-1} \arrow[dll, "\varphi"] & \subseteq & K_i \arrow[dll, "\varphi"] & \subseteq & \ldots & \subseteq & K_{d-1} & \subseteq & K_d = V \arrow[dll, "\varphi"] \\ 0 = K_0 & \subseteq & K_1 & \subseteq & \ldots & \subseteq & K_{i-1} & \subseteq & K_i & \subseteq & \ldots & \subseteq & K_{d-1} & \subseteq & K_d = V \end{tikzcd}$$

Man sieht leicht, dass

$$\varphi^{-1}(K_{i-1}) = K_i \quad \text{und} \quad \varphi(K_i) \subseteq K_i \tag{9.10}$$

gilt. Weiter gilt für jeden Unterraum $W \subseteq V$:

$$\text{Ist } W \cap K_i \text{ für ein } i \geq 1, \text{ so ist } \varphi|_W : W \to V \text{ injektiv.} \tag{9.11}$$

Denn dann ist ja erst recht $W \cap K_1 = W \cap \ker(\varphi) = 0$.
Deshalb kann man nun sukzessive einer Zerlegung von V finden wie folgt:

- Wir wählen $W_d \subseteq V$ mit $V = K_d = K_{d-1} \oplus W_d$;

- dann ist $\varphi(W_d) \subseteq K_{d-1}$ und $\varphi(W_d) \cap K_{d-2} = 0$ wegen (9.10) und (9.11);
- also gibt es ein $W_{d-1} \subseteq K_{d-1}$ mit $K_{d-1} = K_{d-2} \oplus W_{d-1}$ und $\varphi(W_d) \subseteq W_{d-1}$.

$$
\begin{tikzcd}
\cdots \arrow[r, hook] & K_{d-2} \arrow[r, hook] & K_{d-1} \arrow[r, hook] \arrow[dr, dash, dashed, "\text{komplementär}"'] & K_d \arrow[ddl, "\varphi"] \\
& & & W_d \arrow[u, hook] \arrow[dd, "\varphi"] \\
\cdots \arrow[r, hook] & K_{d-2} \arrow[r, hook] \arrow[dr, dash, dashed, "\text{komplementär}"'] & K_{d-1} & \\
& & W_{d-1} \arrow[u, hook] \arrow[r, phantom, "\supseteq"] & \varphi(W_d)
\end{tikzcd}
$$

Man verfährt so weiter bis hinunter zu K_1 und $K_0 = 0$ und erhält

$$
\begin{array}{ccccccccccccc}
V & = & K_d & & & & & & & & & & \\
 & & \phi\downarrow & & & & & & & & & & \\
V & = & K_{d-1} & \oplus & W_d & & & & & & & & \\
 & & \phi\downarrow & & \phi\downarrow & & & & & & & & \\
V & = & K_{d-2} & \oplus & W_{d-1} & \oplus & W_d & & & & & & \\
 & & \phi\downarrow & & \phi\downarrow & & \phi\downarrow & & & & & & \\
V & = & K_{d-3} & \oplus & W_{d-2} & \oplus & W_{d-1} & \oplus & W_d & & & & \\
 & & & & & & & & & & & & \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \ddots & & \\
 & & & & & & & & & & & & \\
V & = & K_1 & \oplus & W_2 & \oplus & W_3 & \oplus & \cdots & & \cdots & \oplus & W_d \\
 & & \phi\downarrow & & \phi\downarrow & & \phi\downarrow & & & & & & \phi\downarrow \\
V & = & K_0 & \oplus & W_1 & \oplus & W_2 & \oplus & W_3 & \oplus \cdots \oplus & W_{d-1} & \oplus & W_d
\end{array}
$$

Jede Zeile ist eine Zerlegung von V; die Pfeile geben die Wirkung von φ an.Insbesondere ist

$$V = W_1 \oplus W_2 \oplus \ldots \oplus W_d.$$

Wegen (9.11) wissen wir, dass φ auf der Kette der W_i injektiv ist:

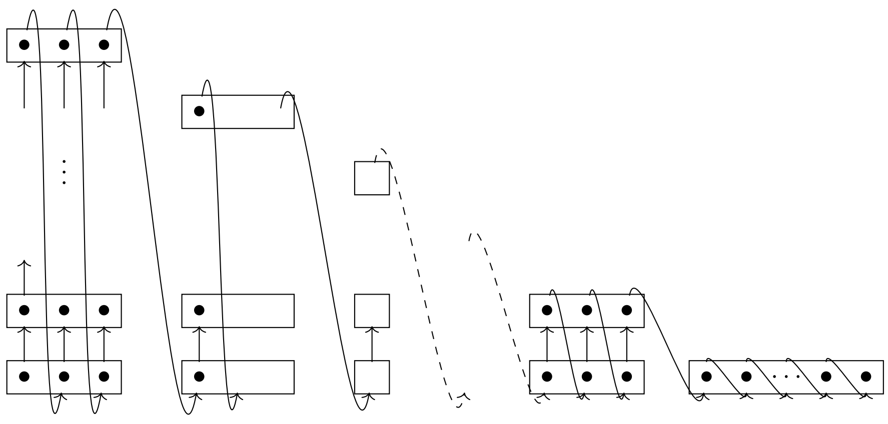
$$W_d \xrightarrow{\varphi} W_{d-1} \xrightarrow{\varphi} W_{d-2} \xrightarrow{\varphi} \ldots \xrightarrow{\varphi} W_2 \xrightarrow{\varphi} W_1$$

Wir wählen in W_d eine Basis $\mathcal{B}_d$; dann ist $\varphi(\mathcal{B}_d)$ linear unabhängig und kann durch $\mathcal{B}_{d-1}$ zu einer Basis von W_{d-1} ergänzt werden. Man verfährt so weiter und erhält eine Gesamtbasis von $W_d \oplus W_{d-1} \oplus \ldots \oplus W_1 = V$.

Teilbasen und Ergänzungen:

$$
\begin{array}{cccccccc}
\boxed{W_d} & \mathcal{B}_d & & & & & & \\
 & \downarrow & & & & & & \\
\boxed{W_{d-1}} & \varphi(\mathcal{B}_d) & \sqcup & \mathcal{B}_{d-1} & & & & \\
 & \downarrow & & \downarrow & & & & \\
\boxed{W_{d-2}} & \varphi^2(\mathcal{B}_d) & \sqcup & \varphi(\mathcal{B}_{d-1}) & \sqcup & \mathcal{B}_{d-2} & & \\
 & \downarrow & & \downarrow & & \downarrow & & \\
\vdots & \vdots & & \vdots & & \vdots & \ddots & \\
 & \downarrow & & \downarrow & & & & \\
\boxed{W_2} & \varphi^{d-2}(\mathcal{B}_d) & \sqcup & \varphi^{d-3}(\mathcal{B}_{d-1}) & \sqcup \ \ldots & & \ldots \ \sqcup & \mathcal{B}_2 \\
 & \downarrow & & \downarrow & & & & \downarrow \\
\boxed{W_1} & \varphi^{d-1}(\mathcal{B}_d) & & \varphi^{d-2}(\mathcal{B}_{d-1}) & \sqcup \ \ldots & & \ldots \sqcup & \varphi(\mathcal{B}_2) \sqcup \mathcal{B}_1
\end{array}
$$

Die Basen $\mathcal{B}_d, \mathcal{B}_{d-1}, \dots, \mathcal{B}_2, \mathcal{B}_1$ seien irgendwie geordnet; die weitergeschobenen Basen $\varphi^k(\mathcal{B}_i)$ seien wie $\mathcal{B}_i$ geordnet; für die Gesamtbasis $\mathcal{B} = \coprod_{k,i} \varphi^k(\mathcal{B}_i)$ erklären wir die Anordnung durch das Schema:



Dann hat φ in dieser Basis die versprochene Form

$$M_{\mathcal{B}\mathcal{B}}(\varphi) = J_d^{\oplus s_d} \oplus \ldots \oplus J_2^{\oplus s_2} \oplus J_1^{\oplus s_1}.$$

Die Dimension der Ergänzungen seien

$$\begin{aligned}
s_d &= \#\mathcal{B}_d &&= \dim W_d \\
s_{d-1} &= \#\mathcal{B}_{d-1} &&= \dim W_{d-1} - \dim W_d \\
&\vdots \\
s_2 &= \#\mathcal{B}_2 &&= \dim W_2 - \dim W_3 \\
s_1 &= \#\mathcal{B}_1 &&= \dim W_1 - \dim W_2
\end{aligned}$$

Diese Zahlen sind wegen $\dim W_i = \dim K_i - \dim K_{i-1}$ durch die Kerne K_i und damit durch φ eindeutig bestimmt.

□

Satz 9.4.3 (Jordan-Normalform)**.** *Es sei $\varphi : V \to V$ ein Endomorphismus, $n = \dim V$, und das charakteristische Polynom*

$$\mathrm{CP}_\varphi(t) = (\lambda_1 - t)^{m_1} \cdot \ldots \cdot (\lambda_k - t)^{m_k}$$

zerfalle mit $\lambda_1, \dots, \lambda_k$ verschieden, $m_1 + \ldots + m_k = n$.
Dann gibt es eine Basis $\mathcal{B}$ von V, so dass φ folgende Matrixdarstellung hat:

$$M_{\mathcal{B}\mathcal{B}}(\varphi) = \begin{pmatrix} \boxed{\lambda_1 \mathbb{1} + N_1} & & & \\ & \boxed{\lambda_2 \mathbb{1} + N_2} & & \\ & & \ddots & \\ & & & \boxed{\lambda_k \mathbb{1} + N_k} \end{pmatrix}$$

und jedes N_i ist selbst eine Blockmatrix aus Jordan-Blöcken

$$N_i = J_{d^{(i)}}(\lambda_i)^{s^{(i)}_{d^{(i)}}} \oplus \ldots \oplus J_2(\lambda_i)^{s_2^{(i)}} \oplus J_1(\lambda_i)^{s_1^{(i)}}$$

und die Zahlen $d^{(1)}, \dots, d^{(k)}$, sowie $s_1^{(i)}, \dots, s_{d_i}^{(i)}$ sind eindeutig bestimmt.

Beweis. Für $i = 1, \ldots, k$ sei der Hauptraum von λ_i

$$V_i := \mathrm{Hau}_{\lambda_i}(\varphi)$$

Wir wissen: $V = V_1 \oplus V_2 \oplus \ldots \oplus V_k$. Wir betrachten

$$\psi_i := (\varphi - \lambda_i \,\mathrm{id}_V) : V_i \longrightarrow V_i,$$

welches nilpotent mit Index $d_i \leq m_i$ ist.
Wir wenden das Lemma auf ψ_i an und erhalten eine Basis $\mathcal{B}_i$ von V_i, so dass

$$M_{\mathcal{B}_i\mathcal{B}_i}(\psi) = N_i$$

bereits die versprochene Gestalt hat. Mit der Basis $\mathcal{B} = \mathcal{B}_1 \sqcup \ldots \sqcup \mathcal{B}_k$ folgt dann die Behauptung. $\square$

Jordan-Normalform im Überblick

Sei $\varphi : V \to V$ ein Endomorphismus, $n = \dim V$.

(1)

$$\mathrm{CP}_{\varphi(t)} = (\lambda_1 - t)^{n_1} \cdot \ldots \cdot (\lambda_k - t)^{n_k}$$

$$n_1 + \ldots + n_k = n, \qquad \lambda_1, \ldots, \lambda_k \text{ verschieden Eigenwerte}$$

(2) Für jeden Eigenwert $\lambda = \lambda_i$ $(i = 1, \ldots, k)$ setzen wir

$$K_l(\lambda_i) = \ker((\varphi - \lambda_i \,\mathrm{id}_V)^l), \qquad d_i(\lambda_i) = d_i \text{ Nilpotenzindex}$$

$$0 = K_0(\lambda_i) \subsetneqq \underbrace{K_1(\lambda_i)}_{=\mathrm{Eig}_{\lambda_i}(\varphi)} \subsetneqq \ldots \subsetneqq \underbrace{K_{d_i}(\lambda_i)}_{=\mathrm{Hau}_{\lambda_i}(\varphi)} = \ldots = K_{n_i}(\lambda_i)$$

- $r_l(\lambda_i) := \dim K_l^{(i)}$
- $r_1(\lambda_i) = \mathrm{geomV}(\varphi, \lambda_i)$
- $d_i \leq n_i$

(3) $s_l(\lambda_i)$ ist die Anzahl der Jordan-Blöcke $J_l(\lambda_i)$ $(i = 1, \ldots, k;\ l = 1, \ldots, d_i)$

(4) Es gilt für $1 < l < n_i$:

- $r_l(\lambda_i) - r_{l-1}(\lambda_i) = s_l(\lambda_i) + s_{l+1}(\lambda_i) + \ldots + s_{d_i}(\lambda_i)$,
- $s_l(\lambda_i) = 2r_l(\lambda_i) - r_{l+1}(\lambda_i) - r_{l-1}(\lambda_i)$,
- $s_l(\lambda_i) = 0 \quad$ für $l > d_i$.

(5) Falls wir das Minimalpolynom kennen:

$$\mathrm{MP}_\varphi(t) = (t - \lambda_1)^{m_1} \cdot \ldots \cdot (\lambda_k - t)^{m_k}, \qquad 1 \leq m_i \leq n_i,$$

so wissen wir:

- $d_i = m_i$
- $s_l(\lambda_i) = 0 \quad$ für $l > m_i$.

Korollar 9.4.4. *Zwei komplexe Matrizen sind genau dann ähnlich, wenn sie das gleiche Spektrum haben und die Jordan-Invarianten $s_1(\lambda), s_2(\lambda), \ldots, s_d(\lambda), d = d(\lambda$ für alle Eigenwerte λ übereinstimmen.*

Korollar 9.4.5. *Eine komplexe Matrix ist genau dann diagonalisierbar, wenn für jeden Eigenwert λ gilt:*

$$0 = s_2(\lambda) = s_3(\lambda) = \ldots = s_d(\lambda), \quad d = d(\lambda).$$

Jordan-Normalform über $\mathbb{R}$

Über $\mathbb{K} = \mathbb{R}$ zerfallen nicht alle charakteristischen Polynome, weshalb sie eventuelll keine Jordansche Normalform besitzen (d.h. nicht ähnlich sind zu einer Blockmatrix aus Jordanblöcken $J_d(\lambda_i)$). Wir wissen aber, dass das charakteristische Polynom, wenn es auch nicht gänzlich in Linearfaktoren zerfällt, höchstens noch quadratische Faktoren hat.
Zunächst können wir zwei reelle $n \times n$-Matrizen A und B als komplexe $n \times n$-Matrizen betrachten und die Ähnlichkeitsfrage dort stellen.

Satz 9.4.6. *Zwei reelle $n \times n$-Matrizen A und B sind genau dann über $\mathbb{R}$ ähnlich, wenn sie über $\mathbb{C}$ ähnlich sind:*

$$\exists_{\Omega \in \mathrm{GL}_n(\mathbb{C})}\ \Omega A \Omega^{-1} = B \iff \exists_{\Theta \in \mathrm{GL}_n(\mathbb{R})} \Theta A \Theta^{-1} = B.$$

Beweis. Die Richtung „$\Leftarrow$“ ist offensichtlich. Sei also $\Omega A \Omega^{-1} = B$ für ein $\Omega \in \mathrm{GL}_n(\mathbb{C})$. Wir spalten Ω als

$$\Omega = \Omega_1 + i\Omega_2, \qquad \Omega_1, \Omega_2 \in \mathrm{Mat}_{n,n}(\mathbb{R})$$

in Realteil $\Omega_1 = \mathrm{Re}(\Omega)$ und Imaginärteil $\Omega_2 = \mathrm{Im}(\Omega)$ auf.
Dann ist

$$\Omega A = B\Omega$$

äquivalent zu

$$\Omega_1 A = B\Omega_1 \qquad \text{und} \qquad \Omega_2 A = B\Omega_2.$$

Das komplexe Polynom

$$f(z) = \mathrm{Det}(\Omega_1 + z\Omega_2)$$

ist nicht das Nullpolynom, denn für $z = i$ ist $f(i) = \mathrm{Det}(\Omega) \neq 0$. Also hat f Nicht-Nullstellen, und zwar auch reelle; es sei etwa $t_0 \in \mathbb{R}$ mit $f(t_0) \neq 0$ und wir setzen

$$\Theta := \Omega_1 + t_0\Omega_2 \in \mathrm{Mat}_{n,n}(\mathbb{R}).$$

Dann ist

$$\mathrm{Det}(\Theta) = \mathrm{Det}(\Omega_1 + t_0\Omega_2) = f(t_0) \neq 0,$$

also $\Theta \in \mathrm{GL}_n(\mathbb{R})$. Es gilt nun aber

$$\begin{aligned}\Theta A &= \Omega_1 A + t_0 \Omega_2 A \\ &= B\Omega_1 + t_0 B\Omega_2 = B\Theta.\end{aligned}$$

□

Beispiel 9.4.7. Sei $\varphi : V \to V$, $\dim V = n$ mit

$$\mathrm{MP}_\varphi(t) = (t - \lambda)^m, \qquad m \leq n.$$

φ hat also nur einen Eigenwert, nämlich λ.

Was sind die möglichen Jordan-Normalformen?

$\boxed{n=1}$ $\quad A = (\lambda)$

$\boxed{n=2}$

$$\begin{pmatrix}\lambda & 0\\ 0 & \lambda\end{pmatrix}$$
$m=1$, diagonalisierbar, $\mathrm{geomV}(\lambda)=2$

$$\begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix}$$
$m=2$, nicht-diagonalisierbar, $\mathrm{geomV}(\lambda)=1$

$\boxed{n=3}$

$$\begin{pmatrix}\lambda & & \\ & \lambda & \\ & & \lambda\end{pmatrix}$$
$m=1$, $s=3$, diagonalisierbar

$$\begin{pmatrix}\lambda & & \\ & \lambda & 1\\ & & \lambda\end{pmatrix}$$
$m=2$, $s=2$

$$\begin{pmatrix}\lambda & 1 & \\ & \lambda & 1\\ & & \lambda\end{pmatrix}$$
$m=3$, $s=1$

$\boxed{n=4}$

$$\begin{pmatrix}\lambda & & & \\ & \lambda & & \\ & & \lambda & \\ & & & \lambda\end{pmatrix}$$
$m=1$, $s=4$

$$\begin{pmatrix}\lambda & & & \\ & \lambda & & \\ & & \lambda & 1\\ & & 0 & \lambda\end{pmatrix}$$
$m=2$, $s=3$

$$\begin{pmatrix}\lambda & 1 & & \\ 0 & \lambda & & \\ & & \lambda & 1\\ & & 0 & \lambda\end{pmatrix}$$
$m=2$, $s=2$

$$\begin{pmatrix}\lambda & & & \\ & \lambda & 1 & \\ & & \lambda & 1\\ & & & \lambda\end{pmatrix}$$
$m=3$, $s=2$

$$\begin{pmatrix}\lambda & 1 & & \\ & \lambda & 1 & \\ & & \lambda & 1\\ & & & \lambda\end{pmatrix}$$
$m=4$, $s=1$

Beispiel 9.4.8.

$$A = \left(\begin{array}{cc|cc} 1 & 1 & & \\ -1 & 3 & & \\ \hline & & 2 & 0\\ & & 0 & 2\end{array}\right)$$

$\mathrm{CP}_A(t) = (2-t)^4$
$\mathrm{MP}_A(t) = (t-2)^2$

- nur ein Hauptraum

$$\psi = A - 2\mathbb{1} = \left(\begin{array}{cc|cc} -1 & 1 & & \\ -1 & 1 & & \\ \hline & & 0 & 0\\ & & 0 & 0\end{array}\right)$$

$\psi^2 = 0.$

- $\mathrm{geomV}_A(2) = 3$
- Es gibt 3 Jordan-Blöcke:

$$A \sim \begin{pmatrix}2 & 1 & & \\ & 2 & & \\ & & 2 & \\ & & & 2\end{pmatrix}$$

$$B = \begin{pmatrix}9 & -7 & 0 & 2\\ 7 & -5 & 0 & 2\\ 4 & -4 & 2 & 1\\ 0 & 0 & 0 & 2\end{pmatrix}$$

$\mathrm{CP}_B(t) = (2-t)^4$
$\mathrm{MP}_B(t) = (t-2)^2$

- nur ein Hauptraum

$$\psi = B - 2\mathbb{1} = \begin{pmatrix}7 & -7 & 0 & 2\\ 7 & -7 & 0 & 2\\ 4 & -4 & 0 & 1\\ 0 & 0 & 0 & 0\end{pmatrix}$$

$\psi^2 = 0.$

- $\mathrm{geomV}_B(2) = 2$
- Es gibt 2 Jordan-Blöcke:

$$B \sim \begin{pmatrix}2 & 1 & & \\ & 2 & & \\ & & 2 & 1\\ & & & 2\end{pmatrix}$$

Beispiel 9.4.9. $\mathbb{K} = \mathbb{R}$, $n = 6$

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

$$\mathrm{CP}_A(t) = (3-t)(2-t)^5$$

- zwei Eigenwerte: $\lambda_1 = 2$, $\lambda_2 = 3$
- zwei Haupträume:

$$\begin{aligned} 1 &\leq \mathrm{geomV}(3) = 1, \\ 1 &\leq \mathrm{geomV}(2) \leq 5 \end{aligned}$$

- $$\mathrm{rg}(A - 2\mathbb{1}) = \mathrm{rg} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 1 & 1 & 0 \\ 1 & -1 & 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -4 & 2 & 2 & 2 & 0 \\ 0 & 2 & -1 & -1 & -1 & 0 \end{pmatrix} = 3$$

- $\mathrm{geomV}(2) = 3$

Es gibt also 3 Jordan-Blöcke für $\lambda_2 = 2$. Für deren Größen gibt es folgende Möglichkeiten:

a) 5= 1+1+3
b) 5=1+2+2

a)

$$A \sim \begin{pmatrix} 2 & & & & & \\ & 2 & & & & \\ & & 2 & 1 & & \\ & & & 2 & 1 & \\ & & & & 2 & \\ & & & & & 3 \end{pmatrix} \implies \mathrm{MP}_A(t) = (t-3)(t-2)^3$$

Aber $\mathrm{rg}(A - 2\mathbb{1})^2 = 1$, also

$$\begin{aligned} &\dim\ \ker(A - 2\mathbb{1})^2 = 6 - 1 = 5 \\ &\dim\ \ker(A - 2\mathbb{1})^2 - \dim\ \ker(A - 2\mathbb{1}) = 5 - 3 = 2. \end{aligned}$$

Also gibt es 3 Jordan-Blöcke der Größe $s \geq 2$. Damit folgt: b) ist die richtige Normalform:

$$A \sim \begin{pmatrix} 2 & & & & & \\ & 2 & 1 & & & \\ & & 2 & & & \\ & & & 2 & 1 & \\ & & & & 2 & \\ & & & & & 3 \end{pmatrix} \implies \mathrm{MP}_A(t) = (t-3)(t-2)^3$$

10 Skalarprodukte

10.1 Motivation

Im allerersten Kapitel hatten wir uns mit dem $\mathbb{R}^2$ und dem $\mathbb{R}^3$ beschäftigt. Damals spielte der Begriff „Länge eines Vektors“ und „Winkel zwischen zwei Vektoren“ eine große Rolle, wie es ja für Anwendungen der Linearen Algebra auf Probleme der zwei- und dreidimensionalen Geometrie selbstverständlich ist. Nun haben wir eine ganze Menge Linearer Algebra gelernt und wenden uns wieder solchen „geometrischen“ Fragen zu:

(1) Gibt es auf einem $\mathbb{K}$-Vektorraum stets ein Skalarprodukt und damit einen Winkel- und Längenbegriff?

(2) Welche Endomorphismen sind dann winkel- und längentreu?

10.2 Bilinearformen

Definition 10.2.1. Es sei V ein Vektorraum über $\mathbb{K}$ mit beliebiger Dimension. Eine *bilineare Form*, d.h. eine Abbildung

$$\sigma : V \times V \longrightarrow \mathbb{K}$$

so dass für alle $x, y, z \in V, \alpha, \beta, \gamma, \delta \in \mathbb{K}$ gilt:

$$\begin{aligned} & \sigma(\alpha x + \beta y, z) = \alpha\sigma(x, z) + \beta\sigma(y, z), \\ \text{und} \qquad & \sigma(x, \gamma y + \delta z) = \gamma\sigma(x, y) + \delta\sigma(x, z), \end{aligned}$$

heißt

- *symmetrisch*, wenn $\sigma(x, y) = \sigma(y, x)$ für alle $x, y \in V$ gilt;
- *schief-symmetrisch* (oder *alternierend*), wenn $\sigma(x, y) = -\sigma(y, x)$ für alle $x, y \in V$ gilt;

Bilinearformen sind spezielle Multilinearformen; die Determinante haben wir als Beispiel einer alternierenden n-Form kennengelernt.

Bemerkung 10.2.2 (Quadratische Formen). [1] Einer symmetrischen Bilinearform σ auf V kann man eine Funktion

$$q_\sigma : V \longrightarrow \mathbb{K}, \qquad q_\sigma(x) = \sigma(x, x)$$

zuordnen. Es gilt

$$\begin{aligned} & q_\sigma(\lambda x) = \lambda^2 x, \quad \text{und} \\ & q_\sigma(x + y) - q_\sigma(x) - q_\sigma(y) = 2\sigma(x, y). \end{aligned}$$

Ist $\mathrm{char}(\mathbb{K}) \neq 2$, so kann man σ aus q_σ zurückgewinnen durch

$$\sigma(x, y) = \frac{1}{2}\left(q_\sigma(x + y) - q_\sigma(x) - q_\sigma(y)\right).$$

[1] wurde nicht in der Vorlesung besprochen

Definition 10.2.3. Eine Funktion $q : V \to \mathbb{K}$ heißt *quadratische Form*, wenn

$$q(\lambda x) = \lambda^2 q(x)$$

gilt für $\lambda \in \mathbb{K}$, $x \in V$ und

$$\sigma(x, y) := q(x + y) - q(x) - q(y)$$

eine symmetrische Bilinearform ist.

Beispiel 10.2.4. (1) $V = \mathbb{K}$: $q(x) = x^2$.

(2) $V = \mathbb{K}$, $\operatorname{char}(\mathbb{K}) = 2$: $q(x) = x^2 + x$.

Beispiel 10.2.5.

(1) $V = \mathbb{K}^n$, $\sigma(x, y) = \sum_{i=1}^n x_i y_i$, wobei $x = (x_1, \ldots, x_n)^\top$, $y = (y_1, \ldots, y_n)^\top$. Diese symmetrische Bilinearform nennt man die *Standardform auf* $\mathbb{K}^n$.

(2) $V = \mathbb{K}^n$, $\sigma_p(x, y) = \sum_{i=1} x_i y_i - \sum_{i=p+1}^n x_i y_i$ $(0 \leq p \leq n)$. Diese Form ist symmetrisch und für $p = n$ erhalten wir die Standardform.

(3) $V = \mathbb{K}^2$, $\sigma(x, y) = a_{11}x_1y_1 + a_{12}x_1y_2 + a_{21}x_2y_1 + a_{22}x_2y_2$.
Für $a_{12} = a_{21}$ erhalten wir eine symmetrische; und für $a_{12} = -a_{21}, a_{11} = a_{22} = 0$ eine schiefsymmetrische Form.

(4) $V = \operatorname{Mat}_{n,n}(\mathbb{K})$, $\sigma(A, B) = \operatorname{Spur}(A^\top B)$. Warum man hier $\operatorname{Spur}(A^\top B)$ anstatt einfach $\operatorname{Spur}(AB)$ nimmt, werden wir gleich sehen; beide Möglichkeiten führen jedenfalls zu symmetrischen Bilinearformen.

(5) $V = \mathcal{C}([-1, 1], \mathbb{R})$, der $\mathbb{R}$-Vektorraum der stetigen Funktionen $\mathbb{R} \supset I \to \mathbb{R}$.

$$\sigma(f, g) := \int_{-1}^{+1} f(t)g(t)\mathrm{d}t$$

ist eine symmetrische Bilinearform.

(6) $V = \mathcal{C}([-1, 1], \mathbb{R})$,

$$\sigma(f, g) := \int_{-1}^{+1} f(t)g(t)\frac{\mathrm{d}t}{\sqrt{1 - t^2}}$$

ist eine andere symmetrische Bilinearform auf demselben Vektorraum.

(7) $V = \operatorname{Pol}(\mathbb{R})$, der $\mathbb{R}$-Vektorraum der reellen Polynome;

$$\sigma(f, g) := \sum_{k=0}^{\infty} f^{(k)}(0)g^{(k)}(0) = \sum_{k=0}^{\infty} (k!)^2 f_k g_k,$$

worin $f^{(k)}(0) = k!f_k$ die k-te Ableitung von $f = f_0 + f_1x + f_2x^2 + \ldots$ an der Stelle 0 ist, ist eine symmetrische Bilinearform.

(8) $V = \operatorname{Flg}_{\mathrm{konv}}(\mathbb{R})$, der $\mathbb{R}$-Vektorraum der konvergenten reellen Folgen;

$$\sigma((a_n), (b_n)) := \lim_{n \to \infty} a_n b_n.$$

(9) $V = \operatorname{Fkt}(X, \mathbb{K})$, X eine endliche Menge;

$$\sigma(f, g) = \sum_{x \in X} f(x)g(x).$$

10.2.1 Darstellung einer Bilinearform durch eine Matrix

Es sei nun $\dim V = n$ endlich und $\mathcal{B} = (b_1, \ldots, b_n)$ eine Basis. Für die Darstellung zweier Vektoren $x = \sum_{i=1}^n x_i b_i$, $y = \sum_{j=1}^n y_j b_j$ in dieser Basis rechnet man für jede Bilinearform σ nach:

$$\begin{aligned}\sigma(x,y) &= \sigma(\sum_{i=1}^n x_i b_i, \sum_{j=1}^n y_j b_j)\\ &= \sum_{i=1}^n \sum_{j=1}^n x_i y_j \sigma(b_i, b_j).\end{aligned}$$

Also ist σ durch seine Werte $\sigma(b_i, b_j)$ auf Paaren (b_i, b_j) von Basisvektoren bestimmt. Diese Werte trägt man in eine $n \times n$-Matrix ein:

$$A = M_\mathcal{B}(\sigma) := (\sigma(b_i, b_j))_{i,j},$$

welche wir die *Darstellung von σ in der Basis $\mathcal{B}$ nennen.*

Achtung. Man beachte, dass der Index $\mathcal{B}$ nur einmal auftaucht; überhaupt ist dieser Darstellungsvorgang strikt zu unterscheiden von der Matrizendarstellung einer linearen Abbildung.

Hat man umgekehrt eine Matrix $A \in \mathrm{Mat}_{n,n}(\mathbb{K})$ gegeben, so ist offensichtlich

$$\sigma^A : \mathbb{K}^n \times \mathbb{K}^n \longrightarrow \mathbb{K}, \qquad \sigma^A(x,y) = x^\top A y$$

eine Bilinearform. Komponiert man noch mit der Koordinatenfunktion $\mathbb{K}_\mathcal{B} : V \to \mathbb{K}^n$, so erhält man eine Bilinearform $\sigma^A_\mathcal{B} = \sigma^A \circ (K_\mathcal{B} \times K_\mathcal{B})$.

$$\begin{tikzcd} V \times V \arrow[d, "\mathbb{K}_\mathcal{B} \times \mathbb{K}_\mathcal{B}"'] \arrow[dr, "\sigma^A_\mathcal{B} = L_\mathcal{B}"] & \\ \mathbb{K}^n \times \mathbb{K}^n \arrow[r] & \mathbb{K} \end{tikzcd}$$

Für $A = M_\mathcal{B}(\sigma)$ ist $\sigma^A_\mathcal{B} = \sigma$; und umgekehrt ist für $\sigma = \sigma^A_\mathcal{B}$ dann $M_\mathcal{B}(\sigma) = A$.
Es bezeichne $\mathrm{Sym}_n(\mathbb{K})$ bzw. $\mathrm{Alt}_n(\mathbb{K})$ den $\mathbb{K}$-Vektorraum der symmetrischen bzw. alternierenden/schiefsymmetrischen Matrizen. $\mathrm{Bil}(V)$ bzw. $\mathrm{Sym}(V)$ bzw. $\mathrm{Alt}(V)$ seien die $\mathbb{K}$-Vektorräume der Bilinearformen bzw. der symmetrischen sowie alternierenden Bilinearformen.

Satz 10.2.6. *Es sei $\mathcal{B}$ eine feste Basis. Dann sind $M_\mathcal{B}$ und $\sigma_\mathcal{B}$ zueinander inverse Isomorphismen von $\mathbb{K}$-Vektorräumen:*

$$\begin{tikzcd} \mathrm{Bil}(V) \arrow[r, shift left, "M_\mathcal{B}"] \arrow[r, phantom, "\cong"] & \mathrm{Mat}_{n,n}(\mathbb{K}) \arrow[l, shift left, "\sigma_\mathcal{B} = L_\mathcal{B}"] \end{tikzcd}$$

Entsprechen sich σ und A, so gilt:

(1) $\sigma \in \mathrm{Sym}(2, V) \iff A \in \mathrm{Sym}_n(\mathbb{K})$,

(2) $\sigma \in \mathrm{Alt}(2; V) \iff A \in \mathrm{Alt}_n(\mathbb{K})$.

Beweis. Dass $M_\mathcal{B}()$ und $\sigma_\mathcal{B}()$ Isomorphismen sind, ist klar. Ebenso ist klar, dass sie zueinander invers sind. Die Behauptungen (1) und (2) folgen aus der Rechnung oben. □

Beispiel 10.2.7. (In den ersten drei Beispielen ist $\mathcal{B} = \mathcal{E}$, die Standardbasis)

(1) Die zur Standarform gehörende Matrix ist die Einsmatrix !.

(2) Die zu σ_p gehörende Matrix ist $\left(\begin{array}{c|c} \mathbb{1}_p & 0 \\ \hline 0 & \mathbb{1}_{n-p} \end{array}\right)$

(3) Die Matrix ist $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$.

(4) Wir nehmen $\mathcal{B} = (\mathcal{E}_{ij})$ in lexikographischer Ordnung. Es ist

$$\sigma(E_{ij}, E_{kl}) = \mathrm{Spur}(E_{ij}^\top E_{kl}) = \mathrm{Spur}(E_{ji}E_{kl}) = \begin{cases} 0, & \text{falls } i \neq k, \\ \mathrm{Spur}(E_{jl}) = \begin{cases} 0, & j \neq l \\ 1, & j = l \end{cases} & \text{falls } i = k \end{cases}$$

Also ist $M_\mathcal{B}(\sigma) = \mathbb{1}_{n^2}$.

(5) Hier beschränken wir uns auf den Unterraum der Polynome und wählen die Basis $1, x, x^2, x^3, \ldots$. Dann ist

$$\sigma_5(x^i, x^j) = \frac{1 - (-1)^{i+j+1}}{i+j+1} = \begin{cases} \frac{2}{i+j+1}, & i+j \text{ gerade} \\ 0, & i+j \text{ungerade} \end{cases}$$

(6) Schränkt man σ_6 ein auf die Polynome, und beharrt auf der Basis $1, x, x^2, \ldots$, so muss man

$$\sigma_6(x^i, x^j) = \frac{1}{i+j+1} \int_{-1}^{1} \frac{x^{i+j+1}}{\sqrt{1-x^2}} \mathrm{d}x$$

berechnen, was nur für $i + j + 1 = 1, 2, 3$ elementar möglich ist. Wählt man hingegen als Basis $\mathcal{B}'$ die Tschebyscheff-Polynome

$$T_0(x) = 1,\ T_1(x) = x, \ldots, T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x), \ldots$$

so kann man alle Matrizeneinträge berechnen:

$$\sigma_6(T_i, T_j) = \int_{-1}^{1} T_i(x)T_j(x) \frac{\mathrm{d}x}{\sqrt{1-x^2}} = \begin{cases} 0, & i \neq j \\ \pi, & i = j = 0 \\ \frac{\pi}{2}, & i = j \neq 0 \end{cases}.$$

(7) Eingeschränkt auf $V_n = \mathrm{Pol}_n(\mathbb{R})$ ist $\sigma_7(x^i, x^j) = i!j!$.

(9) Ist $X = \{1, 2, \ldots, n\}$, so kann man als Basis $\mathcal{B} = (b_1, \ldots, b_n)$ die Funktionen $b_i(k) = \delta_{ik}$ nehmen; dann ist

$$\sigma_9(b_i, b_j) = \delta_{ij}.$$

10.2.2 Transformationsformel

Sei V' ein $\mathbb{K}$-Vektorraum mit Basis $\mathcal{B}'$, und V ein $\mathbb{K}$-Vektorraum mit Basis $\mathcal{B}$; ist σ eine Bilinearform auf V, so kann man diese mit jedem Homomorphismus $f : V' \to V$ zu einer Bilinearform $\sigma' = f^*(\sigma)$ auf V' „zurückziehen“, indem man setzt

$$\sigma'(x', y') = f^*(\sigma)(x', y') := \sigma(f(x'), f(y')).$$

Offenbar ist σ' symmetrisch (bzw. schief-symmetrisch), falls σ symmetrisch (bzw. schief-symmetrisch) ist. Es gilt

- $\mathrm{id}^*(\sigma) = \sigma$.
- $(g \circ f)^*(\sigma) = f^*(g^*(\sigma))$.
- $K_\mathcal{B}^*(\sigma^A) = \sigma$, falls $A = M_\mathcal{B}(\sigma)$.

Ist nun $F = M_{\mathcal{B}\mathcal{B}'}(f)$ und $A = M_\mathcal{B}(\sigma)$, so folgt

$$A' = M_{\mathcal{B}'}(\sigma') = F^\top AF,$$

denn für $x, y \in V$ ist $K_\mathcal{B}(x)^\top AK_\mathcal{B}(y) = \sigma(x, y)$, d.h. für $x = f(x'), y = f(y')$ erhält man

$$\begin{aligned} K_\mathcal{B}(f(x'))^\top AK_\mathcal{B}(f(y')) &= (FK_{\mathcal{B}'}(x'))^\top A(FK_{\mathcal{B}'}(y')) \\ &= K_{\mathcal{B}'}(x')^\top F^\top AFK_{\mathcal{B}'}(y'). \end{aligned}$$

10.2.3 Nicht-ausgeartete Bilinearformen

Definition 10.2.8. Eine Bilinearform $\sigma : V \times V \to \mathbb{K}$ heißt *nicht-ausgeartet*, wenn für jedes $x \neq 0$ ein $y \neq 0$ existiert, so dass $\sigma(x, y) \neq 0$ gilt.

Beispiel 10.2.9. (1) Für $V = \mathbb{K}^n$ und $0 \leq p \leq n$ ist

$$\sigma_p(x, y) = \sum_{i=0}^{p} x_i y_i - \sum_{i=p+1}^{n} x_i y_i$$

nicht ausgeartet: ist z.B. $x \neq 0$, und etwa $x_i \neq 0$, so setzt man $y = e_i$ und erhält $\sigma_p(x, e_i) = \pm x_i$ (je nachdem, ob $i \leq p$ oder $i > p$).

(2) Für $V = \mathrm{Mat}_{n,n}(\mathbb{K})$ ist $\sigma(A, B) = \mathrm{Spur}(A^\top B)$ nicht-ausgeartet: ist etwa $a_{ij} \neq 0$, so setze $B = E_{ij}$ und $(A^\top E_{ij}) = a_{ij} E_{ij}$, also $\sigma(A, E_{ij}) = a_{ij}$. Auch $\sigma'(A, B) = \mathrm{Spur}(AB)$ ist nicht-ausgeartet.

Sei $S = M_\mathcal{B}(\sigma)$ die Matrixdarstellung einer symmetrischen Bilinearform σ auf V mit der Basis $\mathcal{B}$.

Satz 10.2.10. *σ ist genau dann nicht-ausgeartet, wenn S invertierbar ist.*

Beweis. Es genügt $\sigma^S(x, y) = x^\top S y$ zu betrachten.
„$\Rightarrow$“ Ist S nicht invertierbar, so gibt es ein $x \neq 0$ mit $Sx = 0$, also $x^\top S^\top = x^\top S = 0$; dann gilt für alle y auch $x^\top S y = 0$.
„$\Leftarrow$“ Ist σ ausgeartet, so gibt es ein $x \neq 0$ mit $x^\top S y = 0$ für alle y; setzt man $y = e_1, \ldots, e_n$ ein, so folgt $(x^\top S)_i = 0$ für alle $i = 1, \ldots, n$. Also ist $x^\top S = Sx = 0$, und S ist nicht invertierbar. □

Definition 10.2.11. Es sei V ein $\mathbb{K}$-Vektorraum und σ eine symmetrische Bilinearform.

(i) Zwei Vektoren $u, w \in V$ *stehen orthogonal aufeinander*, wenn $\sigma(u, w) = 0$ gilt. Notation: $u \perp w$.

(ii) Zwei Unterräume $U, W \subseteq V$ *stehen orthogonal aufeinander*, wenn $\sigma(u, w) = 0$ gilt für alle $u \in U, w \in W$. Notation: $U \perp W$.

(iii) Zu einem Unterraum $U \subset V$ nennt man

$$U^\perp := \{w \in V \mid u \perp w \text{ für alle } u \in U\}$$

das *orthogonale Komplement* von U in V.

Die Namensgebung „orthogonales Komplement“ ist nur für sog. positiv-definite Formen σ auf endlichdimensionalen reellen (oder komplexen) Vektorräumen gerechtfertigt, siehe nächster Abschnitt. Sobald es Vektoren $x \neq 0$ mit $\sigma(x, x) = 0$ gibt (die man *isotrope* Vektoren nennt), werden die Verhältnisse kompliziert.

Definition 10.2.12. Ein Vektor $x \in V$ heißt *isotrop*, wenn $\sigma(x, x) = 0$. Der Vektorraum V heißt *isotrop*, wenn es ein isotropes $x \in V$ gibt. Anderenfalls, d.h. wenn $\forall x \in V : \sigma(x, x) \neq 0$ gilt, heißt V *anisotrop.*

10.3 Dualraum

Definition 10.3.1. Für jeden Vektorraum V ist sein *Dualraum* definiert als

$$V^* := \mathrm{Hom}_\mathbb{K}(V, \mathbb{K}).$$

Seine Elemente $\varphi : V \to \mathbb{K}$ nennt man *(Linear-)Formen.*

Wir wissen, dass $\mathrm{Hom}_\mathbb{K}(\_, \mathbb{K})$ ein kontravarianter Funktor ist: für ein $f : W \to V$ gibt es eine lineare Abbildung $f^* : W^* \to V^*$:

$$\begin{tikzcd}
V & V^* = \mathrm{Hom}_\mathbb{K}(V, \mathbb{K}) \arrow[d, "f^*"'] & \varphi \arrow[d, mapsto] \\
W \arrow[u, "f"] & W^* = \mathrm{Hom}_\mathbb{K}(W, \mathbb{K}) & f^*(\varphi) = \varphi \circ f
\end{tikzcd}$$

und es gilt:

(1) $\mathrm{id}_V^* = \mathrm{id}_{V^*}$,

(2) $(f \circ g)^* = g^* \circ f^*$.

Ist $\mathcal{B} = (b_1, \dots, b_n)$ eine Basis von V, so erhält man eine Basis $\mathcal{B}^* = (b_1^*, \dots, b_n^*)$ von V^*, welche *duale Basis* heißt, wie folgt:

$$b_k^* : V \longrightarrow \mathbb{K}, \qquad b_k^*(b_i) = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases}.$$

Es ist also $b_k^* = K_{\mathcal{B}}^k$, die k-te Koordinatenfunktion bezüglich $\mathcal{B}$. Man beachte: b_i^* hängt nicht nur von b_i ab, sondern von der gesamten Basis $\mathcal{B}$.
Die Matrixdarstellung von f und f^* in den Basenpaaren $(\mathcal{B}, \mathcal{A})$ bzw. $(\mathcal{A}^*, \mathcal{B}^*)$ erklärt endlich die Bedeutung des Transponierens:

Satz 10.3.2.

$$M_{\mathcal{A}^*\mathcal{B}^*}(f^*) = M_{\mathcal{B}\mathcal{A}}(f)^\top$$

Beweis. Wir schreiben $\mathrm{Mat}_{\mathcal{B}\mathcal{A}}(f) = C = (c_{ik})$ Ist $f(a_k) = \sum_{i=1}^n c_{ik} b_i$, so ist $f^*(b_i^*) = \sum_{i=1}^n c_{ik} a_k^*$. $\square$

Korollar 10.3.3. *Sei $f : V \to V$ ein Endomorphismus. Dann gilt:*

1) $\mathrm{Det}(f^*) = \mathrm{Det}(f)$,

2) $\mathrm{Spur}(f^*) = \mathrm{Spur}(f)$,

3) $\mathrm{CP}_f(t) = \mathrm{CP}_{f^*}(t)$ *und*

4) $\mathrm{MP}_f(t) = \mathrm{MP}_{f^*}(t)$.

Es gibt zwischen V und V^* eine wichtige Paarung, d.h. eine bilineare Abbildung

$$\begin{aligned} \sigma : V^* \times V &\longrightarrow \mathbb{K}, \\ (\varphi, v) &\longmapsto \varphi(v). \end{aligned}$$

Diese ist *nicht ausgeartet* im folgenden Sinn:

- Ist $\sigma(\varphi, v) = 0$ für alle v, so ist $\varphi = 0$.
- Ist $\sigma(\varphi, v) = 0$ für alle φ, so ist $v = 0$.

Korollar 10.3.4.

1) $\ker(f^*)$ *und* $\mathrm{im}(f)^*$ *sind komplementär in* V^*.

2) $\ker(f)^*$ *und* $\mathrm{im}(f^*)$ *sind komplementär in* W^*.

Insbesondere gilt:

$$\begin{aligned} f \text{ injektiv} &\iff f^* \text{ surjektiv.} \\ f \text{ surjektiv} &\iff f^* \text{ injektiv.} \\ f \text{ Isomorphismus} &\iff f^* \text{ Isomorphismus.} \end{aligned}$$

Beweis. (unabhängig von Satz 1, under Benutzung der dualen Basen): Es sei $\mathcal{A}'$ eine Basis von $\ker(f)$ und $\mathcal{A}''$ eine Ergänzung zu einer Basis $\mathcal{A} = \mathcal{A}' \sqcup \mathcal{A}''$ von W.

$$\begin{tikzcd}
0 & \subseteq & \mathrm{im}(f) & \subseteq & V & & V^* \arrow[d, "f^*"'] \arrow[drr] & \supseteq & \ker(f^*) \arrow[drr] & \supseteq & 0 \\
0 & \subseteq & \ker(f) \arrow[ull, "f"] & \subseteq & W \arrow[ull, "f"] \arrow[u, "f"'] & & W^* & \supseteq & \mathrm{im}(f^*) & \supseteq & 0
\end{tikzcd}$$

Dann ist $\mathcal{B}'' = f(\mathcal{A}'')$ eine Basis für $\mathrm{im}(f)$ und $\mathcal{B}'$ sei eine Ergänzung zu einer Basis $\mathcal{B} = \mathcal{B}' \sqcup \mathcal{B}''$ von V.
Es ist $\Phi_{\mathcal{A}}(\mathcal{A}''$ eine Basis von $\mathrm{im}(f^*)$ und $\Phi_{\mathcal{B}}(\mathcal{A}')$ eine Ergänzungsbasis; und $\Phi_{\mathcal{B}}(\mathcal{B})$ ist eine Basis für $\ker(f^*)$ und $\Phi_{\mathcal{B}}(\mathcal{B}'')$ eine Ergänzungsbasis. $\square$

Offenbar hat V^* die gleiche Dimension wie V. Ein möglicher Isomorphismus benutzt die duale Basis

$$\Phi_{\mathcal{B}} : V \xrightarrow{\cong}, \qquad b_k \longmapsto b_k^* = K_{\mathcal{B}}^k.$$

Dieser Isomorphismus mit der Matrix $M_{\mathcal{B}^*\mathcal{B}}(\Phi_{\mathcal{B}}) = \mathbb{1}$ ist *nicht natürlich*, d.h. nicht für alle $f : W \to V$ gilt

$$\Phi_A = f^* \circ \Phi_{\mathcal{B}} \circ f, \qquad \begin{tikzcd} V \arrow[r, "\Phi_{\mathcal{B}}", "\cong"'] & V^* \arrow[d, "f^*"] \\ W \arrow[u, "f"] \arrow[r, "\Phi_{\mathcal{A}}"'] \arrow[ur, phantom, "\neq"] & W^* \end{tikzcd} \;;$$

denn das würde nach Satz 10.3.2 für die darstellende Matrix $A = M_{\mathcal{B}\mathcal{A}}(f)$ bedeuten

$$\mathbb{1} = A^\top \cdot \mathbb{1} \cdot A, \qquad \text{also } A^\top = A^{-1},$$

was offenbar nicht für alle Matrizen gilt.

Man kann nun weiter von V^* zum Doppeldual V^{**} gehen:

$$\begin{tikzcd} V \arrow[r, "\Phi_{\mathcal{B}}", "\cong"'] \arrow[rr, bend left, "D_{\mathcal{B}}"] & V^* \arrow[r, "\Phi_{\mathcal{B}^*}", "\cong"'] \arrow[d, "f^*"'] & V^{**} \\ W \arrow[u, "f"] \arrow[r, "\cong", "\Phi_{\mathcal{A}}"'] \arrow[rr, bend right, "D_{\mathcal{A}}"'] & W^* \arrow[r, "\cong", "\Phi_{\mathcal{A}^*}"'] & W^{**} \arrow[u, "f^{**}"'] \end{tikzcd}$$

Man stellt überrascht fest:

Satz 10.3.5. *Der Isomorphismus $D_{\mathcal{B}} = \Phi_{\mathcal{B}^*} \circ \Phi_{\mathcal{B}} : V \to V^{**}$ ist natürlich, d.h. für alle $f : W \to V$ gilt:*

$$D_{\mathcal{B}} \circ f = f^{**} \circ D_{\mathcal{A}}.$$

Beweis. Übersetzt in Matrizen heißt das nur:

$$M_{\mathcal{B}^{**},\mathcal{B}^*}(\Phi_{\mathcal{B}^*}) M_{\mathcal{B}^*,\mathcal{B}}(\Phi_{\mathcal{B}}) M_{\mathcal{B},\mathcal{A}}(f) = \mathbb{1} \cdot \mathbb{1} \cdot A = A,$$
$$M_{\mathcal{B}^{**},\mathcal{A}^{**}}(f^{**}) M_{\mathcal{A}^{**},\mathcal{A}^*}(\Phi_{\mathcal{A}^*}) M_{\mathcal{A}^*,\mathcal{A}}(\Phi_{\mathcal{A})} = (A^\top)^\top \cdot \mathbb{1} \cdot \mathbb{1} = (A^\top)^\top = A.$$

□

Hält man bei einer Bilinearform $\sigma : V \times V \to \mathbb{K}$ ein Argument fest, so erhält man eine Linearform

$$\begin{matrix} \psi'_v : V \longrightarrow \mathbb{K}, \\ \psi'_v(x) := \sigma(v, x) \end{matrix} \qquad \text{bzw.} \qquad \begin{matrix} \psi''_w : V \longrightarrow \mathbb{K}, \\ \psi''_w(x) := \sigma(x, w) \end{matrix}$$

für jeweils festes $v \in V$ bzw. $w \in V$. Es gilt:

- Falls σ symmetrisch ist, dann ist $\psi'_v = \psi''_v$,
- Falls σ schiefsymmetrisch ist, dann ist $\psi'_v = -\psi''_v$.

Wegen

$$\begin{array}{lll} 1) & \psi'_0 = 0 & \psi''_0 = 0 \\ 2) & \psi'_{\lambda v} = \lambda \psi_v & \psi''_{\lambda w} = \lambda \psi''_w \\ 3) & \psi'_{v_1+v_2} = \psi' v_1 + \psi'_{v_2} & \psi''_{w_1+w_2} = \psi''_{w_1} + \psi''_{w_2}, \end{array}$$

sind

$$\begin{aligned} \Psi'_\sigma : V &\longrightarrow V^* \\ v &\longmapsto \psi'_v \end{aligned} \qquad \text{und} \qquad \begin{aligned} \Psi''_\sigma : V &\longrightarrow V^* \\ w &\longmapsto \psi''_w \end{aligned}$$

lineare Abbildungen. Offenbar gilt:

Satz 10.3.6. *Falls* $\dim(V) = n < \infty$ *und* $\sigma \in \mathrm{Bil}(V)$*, so sind äquivalent:*

(i) σ *ist nicht-ausgeartet.*

(ii) Ψ'_σ *und* Ψ''_σ *sind injektiv.*

(iii) Ψ'_σ *und* Ψ''_σ *sind Isomorphismen.*

Ist σ symmetrisch, so ist $\Psi'_\sigma = \Psi''_\sigma$. Es sei noch bemerkt, dass die Isomorphismen $\Phi_\mathcal{B}$ und $\Psi*_\sigma$ bzw. Ψ''_σ natürlich im Allgemeinen verschieden sind; allerdings gilt: ist σ symmetrisch und nicht-ausgeartet und $\mathcal{B}$ eine Orthonormalbasis bzg. σ, so ist $\Phi_\mathcal{B} = \Psi'_\sigma = \Psi''_\sigma$.

10.4 Euklidische Vektorräume

Das Standardskalarprodukt $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$ hat die besondere Eigenschaft:

$$\begin{array}{ll}(1) & \langle x, x\rangle \geq 0,\\ (2) & \langle x, x\rangle = 0 \iff x = 0.\end{array}$$

Die symmetrischen Bilinearformen

$$\sigma_p(x, y) = \sum_{i=1}^{p} x_i y_i - \sum_{i=p+1}^{n} x_i y_i$$

haben diese Eigenschaft nicht. Wir werden sehen, dass die obigen Eigenschaften genügen, um euklidische Geometrie zu betreiben.

Definition 10.4.1. Es sei V ein $\mathbb{R}$-Vektorraum. Eine symmetrische Bilinearform σ auf V heißt *positiv-definit*, falls $\sigma(x, x) > 0$ für alle $x \neq 0$ gilt. Eine solche Form σ nennt man auch *euklidisches Skalarprodukt*, und V nennt man dann einen *euklidischen Vektorraum*.

Eine positiv-definite, symmetrische Bilinearform (d.h. ein Skalarprodukt) ist stets nicht-ausgeartet.

Beispiel 10.4.2. (1) $V = \mathbb{R}^n$, $\sigma =$ Standardskalarprodukt.

(2) $V = \mathrm{Mat}_{n,n}(\mathbb{K})$, $\sigma(A, B) = \mathrm{Spur}(A^\top B)$ ist positiv-definit:

$$(A^\top A)_{ii} = \sum_{k=1}^{n} a^\top_{ik} a_{ki} = \sum_{k=1}^{n} a^2_{ki},$$

und ist etwa $a_{1i} \neq 0$, so ist

$$\mathrm{Spur}(A^\top A) \geq \sum_{k=1}^{n} a^2_{ki} \geq a^2_{1i} > 0.$$

Hingegen ist $\mathrm{Spur}(AB)$ nicht positiv-definit, denn für $A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ ist $A^2 = 0$, also $\mathrm{Spur}(A^2) = 0$, obwohl $A \neq 0$.

(3) $V = \mathcal{C}([-1, 1], \mathbb{R})$, $\sigma(f, g) = \int_{-1}^1 f(t)g(t)\mathrm{d}t$ ist positiv definit, denn

$$\int_{-1}^{1} f(t)^2 \mathrm{d}t > 0,$$

falls $f \neq 0$.

(4) $V = \mathbb{R}^n$,

$$\sigma_p(x, y) = \sum_{i=1}^{p} x_i y_i - \sum_{i=p+1}^{n} x_i y_i$$

ist für $1 \leq p < n$ nicht positiv-definit, denn es gibt dann isotrope Vektoren. (Für $p = 0$ würde man σ_0 *negativ-definit* nennen.)

Das entscheidend Neue ist nun, dass man in einem Euklidischen Vektorraum Längen und Winkel messen kann. Wir beginnen mit

Lemma 10.4.3 (Schwarz'sche Ungleichung)**.** *Sei σ ein Skalarprodukt. Dann gilt*

$$\sigma(x,y)^2 \leq \sigma(x,x) \cdot \sigma(y,y)$$

Beweis. Wir können $y \neq 0$ und damit $\sigma(y,y) > 0$ annehmen. Für alle $\lambda \in \mathbb{R}$ gilt

$$\begin{aligned} 0 &\leq \sigma(x-\lambda y, x-\lambda y) \\ &= \sigma(x,x) - \sigma(x,\lambda y) - \sigma(\lambda y, x) + \sigma(\lambda y, \lambda y) \\ &= \sigma(x,x) - 2\lambda\sigma(x,y) + \lambda^2\sigma(y,y). \end{aligned}$$

Man setzt $\lambda := \frac{\sigma(x,y)}{\sigma(y,y)}$ und erhält

$$0 \leq \sigma(x,x) - 2\frac{\sigma(x,y)^2}{\sigma(y,y)} + \frac{\sigma(x,y)^2}{\sigma(y,y)^2}\sigma(y,y).$$

Daraus durch Multiplikation mit $\sigma(y,y)$ eben

$$0 \leq \sigma(x,x)\sigma(y,y) - \sigma(x,y)^2.$$

□

Diese Ungleichung ist nun fundamental für die beiden kommenden Begriffe.

Definition 10.4.4.

(i) $\|x\| = \|x\|_\sigma := \sqrt{\sigma(x,x)}$ heißt *Länge* von x.

(ii)
$$\cos\sphericalangle(x,y) = \cos\sphericalangle_\sigma(x,y) := \frac{\sigma(x,y)}{\|x\|_\sigma \|y\|_\sigma}$$
heißt *Kosinus des Winkels* zwischen x und y, $x, y \neq 0$.

Lemma 10.4.5.

(i) $\|x\| \geq 0$*, und* $\|x\| = 0$ *nur für* $x = 0$*.*

(ii) $\|\lambda x\| = |\lambda|\|x\|$*.*

(iii) $\|x+y\| \leq \|x\| + \|y\|$ *(*Dreiecksungleichung*).*

Beweis.

(i) ist klar, weil σ kein isotropen Vektoren zulässt.

(ii) klar.

(iii) Man rechnet nach
$$\sigma(x+y, x+y) = \sigma(x,x) + 2\sigma(x,y) + \sigma(y,y),$$
wegen der Schwarzschen Ungleichung folgt
$$\begin{aligned} \|x+y\|^2 &= \|x\|^2 + 2\sigma(x,y) + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2. \end{aligned}$$

□

Ein weiterer wichtiger Begriff ist:

Definition 10.4.6. Eine Teilmenge $\mathcal{B} = \{b_1, \dots, b_m, \dots\} \subset V$ heißt *Orthonormalsystem*, falls gilt:

$$\sigma(b_i, b_j) = \delta_{ij} = \begin{cases} 0, & \text{für } i \neq j \\ 1, & \text{für } i = j \end{cases}.$$

Die Forderung $\sigma(b_i, b_i) = 1$ lässt sich durch *Normierung* immer erreichen: für $x \neq 0$ hat $x' = \frac{x}{\|x\|}$ immer die Länge 1. $\mathcal{B}$ heißt *Orthonormalbasis*, wenn es Basis und Orthonormalsystem ist.

Lemma 10.4.7. *Ist $\mathcal{B} = \{b_1, \dots\}$ ein Orthonormalsystem, so ist $\mathcal{B}$ linear unabhängig.*

Beweis. Angenommen

$$\alpha_1 b_1 + \ldots + \alpha_m b_m = 0$$

sei eine Relation. Man bildet das Skalarprodukt mit jedem einzelnen $b_1, b_2, \dots$ usw. und erhält

$$\begin{aligned} 0 &= \sigma(\alpha_1 b_1 + \ldots + \alpha_m b_m, b_k) = \\ &= \alpha_1 \sigma(b_1, b_k) + \alpha_2 \sigma(b_2, b_k) + \ldots + \alpha_k \sigma(b_k, b_k) + \ldots + \alpha_m \sigma(b_m, b_k) \\ &= \alpha_k. \end{aligned}$$

□

Damit erhalten wir für jedes $x \in V$ eine *Fourier-Darstellung* bzgl. $\mathcal{B}$:

Lemma 10.4.8. *Ist $\mathcal{B} = (b_1, \dots, b_n)$ eine geordnete Orthonormalbasis von V, so gilt für jedes $x \in V$*

$$x = \sigma(x, b_1) b_1 + \ldots + \sigma(x, b_n) b_n.$$

Insbesondere ist also $K^i_{\mathcal{B}}(x) = \sigma(x, b_i)$ die Koordinate von x bzgl. der Basis $\mathcal{B}$.

Beweis. Es sei $x = \alpha_1 b_1 + \ldots + \alpha_n b_n$ die Darstellung von x in der Basis $\mathcal{B}$. Daraus erhalten wir für $k = 1, 2, \dots, n$ die Gleichungen

$$\begin{aligned} \sigma(x, b_k) &= \sigma(\alpha_1 b_1 + \ldots + \alpha_k b_k + \ldots + \alpha_n b_n, b_k) \\ &= \alpha_1 \sigma(b_1, b_k) + \ldots + \alpha_k \sigma(b_k, b_k) + \ldots + \alpha_n \sigma(b_n, b_k) \\ &= \alpha_k \end{aligned}$$

□

Besonders wichtig ist nun das Orthonormalisierungsverfahren von Gram-Schmidt:

Satz 10.4.9. *Sei V ein euklidische Vektorraum der Dimension n und σ sei das Skalarprodukt. Zu jeder geordneten Basis $\mathcal{B} = (b_1, \dots, b_n)$ gibt es eine neue Basis $\mathcal{B}' = (b'_1, \dots, b'_n)$, so dass gilt:*

(1) $\mathcal{B}'$ ist eine Orthonormalbasis.

(2) Für $m = 1, \dots, n$ gilt für die Teilbasen $\mathcal{B}_m = (b_1, \dots, b_m)$ bzw. $\mathcal{B}'_m = (b'_1, \dots, b'_m)$:

$$\operatorname{Span}(\mathcal{B}_m) = \operatorname{Span}(\mathcal{B}'_m) =: V_m \tag{10.1}$$

$$\operatorname{Det}(M_{\mathcal{B}'_m, \mathcal{B}_m}(\mathrm{id}_{V_m})) > 0. \tag{10.2}$$

Beweis. Wir setzen für $k = 1, 2, \dots, n$ nacheinander

$$\tilde{b}_k := b_k - \sum_{l=1}^{k-1} \sigma(b_k, b'_l) b'_l, \qquad \text{und} \tag{10.3}$$

$$b'_k := \frac{\tilde{b}_k}{\sqrt{\sigma(\tilde{b}_k, \tilde{b}_k)}} = \frac{\tilde{b}_k}{\|\tilde{b}_k\|_\sigma}. \tag{10.4}$$

Offenbar ist $\mathcal{B}' = (b_1', b_2', \dots, b_m')$ ein System von normierten Vektoren wegen (10.4). Für die Orthogonalität rechnet man nach ($1 \le i \le k-1$):

$$\begin{aligned}\sigma(\tilde{b}_k, b_i') &= \sigma(b_k - \sum_{l=1}^{k-1} \sigma(b_k, b_l')b_l', b_i') \\ &= \sigma(b_k, b_i') - \sum_{l=1}^{k-1} \sigma(b_k, b_l')\sigma(b_l', b_i').\end{aligned}$$

Nach Induktionsvoraussetzung können wir

$$\sigma(b_l', b_i') = \delta_{li}, \qquad 1 \le l, i \le k-1$$

annehmen. Also bleibt in der Summe nur der Term $l = i$ stehen:

$$\begin{aligned}\sigma(\tilde{b}_k, b_i') &= \sigma(b_k, b_i') - \sigma(b_k, b_i')\underbrace{\sigma(b_i', b_i')}_{=1} \\ &= \sigma(b_k, b_i') - \sigma(b_k, b_i') \\ &= 0.\end{aligned}$$

Es steht $\tilde{b}_k$ also orthogonal auf allen $b_1', \dots, b_{k-1}'$, also auch b_k'. Damit ist $\mathcal{B}_m'$ ein Orthogonalsystem, also auch $\mathcal{B}' = \mathcal{B}_n'$.
Die Behauptung (10.1) ist offensichtlich, da in (3) von b_k eine Linearkombination der $b_i', \dots, b_{k-1}'$ subtrahiert wird. Die Behauptung (10.2) folgt aus

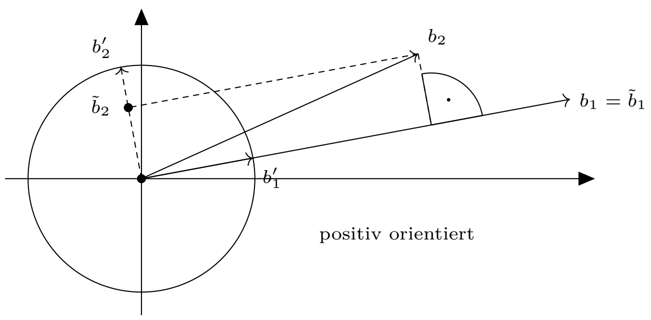
$$M_{\mathcal{B}_m', \mathcal{B}_m}(\mathrm{id}_{V_m}) = \begin{pmatrix} \frac{1}{\beta_1} & -\frac{\sigma(b_2, b_1')}{\beta_2} & \cdots & -\frac{\sigma(b_m, b_1')}{\beta_m} \\ & \frac{1}{\beta_2} & \cdots & -\frac{\sigma(b_m, b_2')}{\beta_m} \\ & & \ddots & \vdots \\ & & & \frac{1}{\beta_m} \end{pmatrix}$$

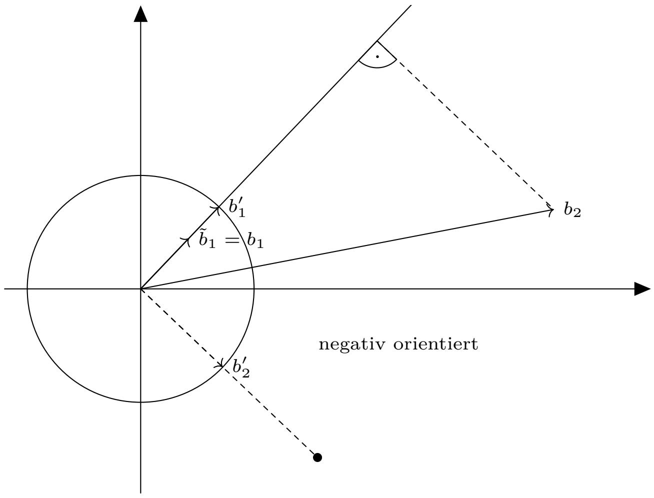
mit $\beta_k = \|\tilde{b}_k\|_\sigma$. □

Bemerkung 10.4.10. 1) $\mathcal{B}$ und $\mathcal{B}'$ gehören zur gleichen Orientierungsklasse, weil wegen (10.2) der Basiswechsel positive Determinante hat.

2) $\mathcal{B}'$ ist nach dem Algorithmus von Gram-Schmidt eine Funktion von $\mathcal{B}$.

Skizzen zum Beweis
$V = \mathbb{R}^2$, σ Standardskalarprodukt





Beispiel 10.4.11. Sei $V = \mathbb{R}^3$, σ das Standardskalarprodukt und $\mathcal{B}$ gegeben durch

$$b_1 = \begin{pmatrix}1\\0\\1\end{pmatrix}, \qquad b_2 = \begin{pmatrix}-1\\-1\\2\end{pmatrix}, \qquad b_3 = \begin{pmatrix}4\\0\\1\end{pmatrix}$$

1)

$$\begin{aligned}\tilde{b}_1 &= b_1\\ b_1' &= \frac{b_1}{\|b_1\|} = \begin{pmatrix}\frac{1}{\sqrt{2}}\\0\\\frac{1}{\sqrt{2}}\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix}.\end{aligned}$$

2)

$$\begin{aligned}\tilde{b}_2 &= b_2 - \underbrace{\sigma(b_2, b_1')}_{(-1)\frac{1}{\sqrt{2}}+(2)\frac{1}{\sqrt{2}}=\frac{1}{\sqrt{2}}} = \begin{pmatrix}-\frac{3}{2}\\-1\\\frac{3}{2}\end{pmatrix}\\ b_2' &= \frac{1}{\sqrt{22}}\begin{pmatrix}-3\\-2\\3\end{pmatrix}.\end{aligned}$$

3)

$$\begin{aligned}\tilde{b}_3 &= b_2 - \underbrace{\sigma(b_3, b_1')}_{\frac{5}{\sqrt{2}}} b_1 - \underbrace{\sigma(b_3, b_2')}_{-\frac{9}{\sqrt{22}}} b_2 = \begin{pmatrix}\frac{3}{11}\\-\frac{9}{11}\\-\frac{3}{11}\end{pmatrix}\\ b_3' &= \frac{\tilde{b}_3}{\sqrt{\frac{9}{121}+\frac{81}{121}+\frac{9}{121}}} = \frac{11}{\sqrt{99}}\begin{pmatrix}\frac{3}{11}\\-\frac{9}{11}\\-\frac{3}{11}\end{pmatrix} = \frac{1}{\sqrt{11}}\begin{pmatrix}1\\-3\\-1\end{pmatrix}.\end{aligned}$$

Beispiel 10.4.12. $V = \mathbb{R}^2$

$$S = \begin{pmatrix}a & b\\ b & c\end{pmatrix}, \qquad S = S^\top$$

$$\begin{aligned}\sigma(x,y) &= x^\top S y = (x_1, x_2)\begin{pmatrix}a & b\\ b & c\end{pmatrix}\begin{pmatrix}y_1\\y_2\end{pmatrix}\\ &= ax_1y_1 + b(x_1y_2 + x_2y_1) + cx_2y_2\end{aligned}$$

$$\|x\|_\sigma = \sqrt{ax_1^2 + 2bx_1x_2 + cx_2^2}$$

Speziell für

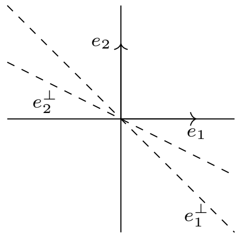
$$S = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

ist also

$$\begin{aligned} \sigma(x,y) &= x_1y_1 + x_1y_2 + x_2y_1 + 2x_2y_2 \\ \|x\|_\sigma &= \sqrt{x_1^2 + 2x_1x_2 + 2x_2^2} \end{aligned}$$

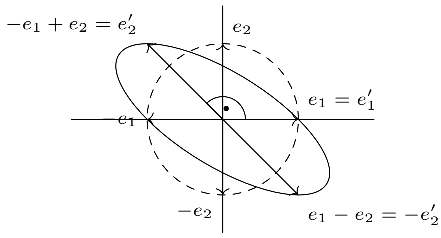
$$\|e_1\|_\sigma = 1, \qquad \|e_2\|_\sigma = \sqrt{2}$$

$$e_1^\perp = \{y \mid y_1 + y_2 = 0\}, \qquad e_2^\perp = \{y \mid y_1 + 2y_2 = 0\}$$



$$\begin{aligned} \tilde{e}_1 &= e_1 \\ e_1' &= e_1 \end{aligned}$$

$$\begin{aligned} \tilde{e}_2 &= e_2 - \underbrace{\sigma(e_2, e_1')}_{1} e_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \\ e_2' &= \frac{\tilde{e}_2}{\|\tilde{e}_2\|_\sigma} = \underbrace{\frac{1}{\sqrt{1 + 2\cdot(-1)\cdot 1 + 2}}}_{=1} \tilde{e}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \end{aligned}$$



Man beachte:

$$\begin{aligned} E &= \{(x_1, x_2) \mid \|x\|_\sigma = 1\} \\ &= \{(x_1, x_2) \mid x_1^2 + 2x_1x_2 + 2x_2^2 = 1\} \qquad \textit{Ellipse} \end{aligned}$$

$$(\pm 1, 0) \in E, \qquad \left(0, \pm\frac{1}{\sqrt{2}}\right) \in E, \qquad \pm(1,-1) \in E.$$

Anwendungen

(1) *Lotrechte Projektion.* Sei $U \subseteq V$ Unterraum, $\mathcal{A} = (a_1, \ldots, a_m)$ Orthonormalbasis für U

$$P_U : V \to V$$

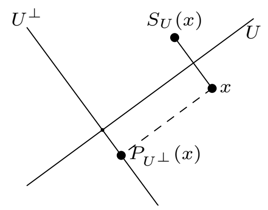
$$P_U(x) := \sum_{i=1}^m \sigma(x, a_i) a_i$$

Die lotrechte Projektion hat folgende Eigenschaften:

1) $P_U(x) \in U$, $P_V = \mathrm{id}_V$, $P_0 = 0$.
2) $x - P_U(x) \in U^\perp$ (Lot von x auf U)
3) $P_U \circ P_U = P_U$ Projektion (idempotent)
4) $P_U(x) = x \iff x \in U = \operatorname{im} P_U$
5) $P_U(x) = 0 \iff x \in U^\perp = \ker P_U$
6) $x - P_U(x) = P_{U^\perp}(x)$
7) P_U hängt nicht von der Wahl von $\mathcal{A}$ ab, sondern nur von σ und $U \subseteq V$.
8) $\mathrm{rg}(P_U) = \dim U$.

(2) *Spiegelungen.* Sei $U \subseteq V$

$$R_U : V \longrightarrow V, \qquad R_U(x) = x - 2P_{U^\perp}(x)$$



1) $S_U(x) = x \iff x \in U = \mathrm{Fix}(S_U)$ („Spiegel")
2) $S_U \circ S_U = \mathrm{id}_V$ Involution
3) $S_U(x) = -x \iff x \in U^\perp$
4) $\dim U = 0$: $\quad S_U(x) = x$, $S_U = \mathrm{id}_V$
   $\dim U = n$: $\quad S_U(x) = -x$, $S_U = -\mathrm{id}_V$

10.5 Skalarprodukte, Normen, Metriken

(1) Ein Skalarprodukt σ auf einem reellen Vektorraum V induziert auf V eine *Norm* durch

$$N(x) := \|x\|_\sigma, \qquad N : V \longrightarrow \mathbb{R},$$

welche folgende Axiome erfüllt:

- $N(x) \geq 0$, und $N(x) = 0 \iff x = 0$,
- $N(\lambda x) = |\lambda| N(x)$,
- $N(x+y) \leq N(x) + N(y)$.

(2) Eine Norm auf V induziert eine *Metrik* durch

$$d(x,y) := N(x-y), \qquad d : V \times V \longrightarrow \mathbb{R},$$

welche folgende Axiome erfüllt:

- $d(x,y) \geq 0$, und $d(x,y) = 0 \iff x = y$,
- $d(x,y) = d(y,x)$,
- $d(x,y) \leq d(x,z) + d(z,y)$.

(3) Eine Metrik auf V induziert eine *Topologie* auf V, welche folgende Eigenschaft hat:

- die Addition $V \times V \longrightarrow V$, $(x,y) \longmapsto x+y$, und
- die Skalierung $\mathbb{R} \times V \longrightarrow V$, $(\lambda, x) \longmapsto \lambda x$

sind stetig.

Wir haben also folgende Kette:

$$\text{Skalarprodukt} \rightsquigarrow \text{Norm} \rightsquigarrow \text{Metrik} \rightsquigarrow \text{Topologie}$$

Aber es gibt Topologien, die nicht von Metriken induziert sind, und Metriken, die nicht von Normen induziert sind, und Normen, die nicht von Skalarprodukten induziert sind.
Man nennt zwei Normen N_1, N_2 auf einem Vektorraum V äquivalent, wenn es zwei Konstanten $\alpha_1, \alpha_2 > 0$ gibt mit

$$N_1(x) \leq \alpha_2 N_2(x), \qquad N_2(x) \leq \alpha_1 N_1(x), \qquad x \in V.$$

Und zwei Metriken d_1, d_2 auf V (oder einer Menge) heißen äquivalent, wenn es zwei positive Konstanten $\alpha_1, \alpha_2 > 0$ gibt mit

$$d_1(x,y) \leq \alpha_2 d_2(x,y), \qquad d_2(x,y) \leq \alpha_1 d(x,y), \qquad x, y \in V$$

Offenbar induzieren äquivalente Normen auch äquivalente Metriken; und äquivalente Metriken induzieren dieselbe Topologie. Für Geometrie und Topologie ist wichtig:

Satz 10.5.1. *Auf einem endlich-dimensionalen reellen Vektorraum sind je zwei Normen äquivalent.*

Zum Beweis, dass jede Norm zur Standardnorm äquivalent ist, benutzt man die Kompaktheit der Sphäre $\mathbb{S} = \{x \mid \|x\| = 1\}$ und die Stetigkeit der Norm. Die Funktion $f : \mathbb{S} \to \mathbb{R}$, $f(x) = N(x)$ nimmt auf $\mathbb{S}$ ein Minimum $a > 0$ und ein Maximum $b > 0$ an. Für ein beliebiges $x \in V$, $x \neq 0$ folgt wegen $N(x) = \|x\| N\left(\frac{x}{\|x\|}\right)$ dann

$$a\|x\| \leq N(x) \leq b\|x\|.$$

10.6 Orthogonale Komplemente

Wir haben das orthogonale Komplement eines Unterraumes $U \subseteq V$ definiert als

$$U^\perp := \{x \in V \mid \sigma(x,u) = 0 \text{ für alle } u \in U\}$$

Nur in einem euklidischen Vektorraum verdient es den Namen Komplement.

Lemma 10.6.1. *In einem euklidischen Vektorraum V gilt für Unterräume:*

(1) $V^\perp = 0$, $0^\perp = V$, $(U^\perp)^\perp = U$.

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

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

(4) $U_1 \perp U_2 \implies U_1 \cap U_2 = 0$.

(5) $U + U^\perp = V$, $U \cap U^\perp = 0$.

Beweis. (1) Nur für $x = 0$ gilt $\sigma(x,y) = 0$ für alle $y \in V$, denn σ ist nicht-ausgeartet. $0^\perp = V$ ist für jede symmetrische Bilinearform. Die dritte Gleichung $(U^\perp)^\perp = U$ folgt aus der Definition.

(2) Die Inklusion „$\subseteq$" ist klar; und aus $\sigma(x,u_1) = \sigma(x,u_2) = 0$ für alle $u_1 \in U_1, u_2 \in U_2$ folgt

$$\sigma(x, u_1 + u_2) = \sigma(x,u_1) + \sigma(x,u_2) = 0$$

für alle $u_1 + u_2 \in U_1 + U_2$.

(3) Ist $(b_1, \ldots, b_k)$ eine Orthonormalbasis für $U_1 \cap U_2$ und $(b_1, \ldots, b_k, b_{k+1}, \ldots, b_l)$ eine Orthonormalbasis für U_1 und $(b_1, \ldots, b_k, b_{l+1}, \ldots, b_m)$ eine Orthonormalbasis für U_2, und $(b_1, \ldots, b_k, \ldots, b_l, \ldots, b_m, \ldots, b_n)$ eine Orthonormalbasis für V, so ist $(b_{k+1}, \ldots, b_n)$ eine für $(U_1 \cap U_2)^\perp$, und $(b_{l+1}, \ldots, b_n)$ eine für $U_1^\perp$ und $(b_{k+1}, \ldots, b_l, b_{m+1}, \ldots, b_n)$ eine für $U_2^\perp$.

(4) Genau wie (3).

(5) Genau wie (3).

$\square$

10.7 Hilbert-Räume

In unendlich-dimensionalen reellen Vektorräumen V mit Skalarprodukt σ hat man oft folgende Situation:

1) $\dim_{\mathbb{R}} V$ ist überabzählbar.
2) es gibt ein abzählbares System $\mathcal{B} = (b_1, b_2, \ldots)$ von Vektoren mit
   - $\sigma(b_i, b_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$,
   - für jedes $x \in V$ und jedes $\epsilon > 0$ gibt es ein $n \in \mathbb{N}$ mit
$$\|x - \sum_{i=1}^{n} \sigma(x, b_i) b_i\| < \epsilon.$$

Es ist also $V' = \operatorname{Span}(\mathcal{B})$ ein dichter Unterraum abzählbarer Dimension. Man nennt ein solches V einen *Hilbert-Raum.*

Beispiel 10.7.1. $V = \mathcal{C}^0([-1,1], \mathbb{R})$ Raum der stetigen Funktionen,

$$\sigma(f, g) = \int_{-1}^{1} f(t) g(t) \mathrm{d}t$$

$$b_k = \begin{cases} \sin(kt), & k \text{ ungerade} \\ \cos(kt), & \text{gerade} \end{cases}.$$

10.8 Unitäre Vektorräume

Im letzten Abschnitt haben wir uns ausschließlich mit reellen Vektorräumen V mit einer symmetrischen nicht-ausgearteten Bilinearform σ beschäftigt, die zudem auch positiv-definit war. Wie steht es mit komplexen Vektorräumen? – Um wieder Längen definieren zu können, sollte $\sigma(x, x)$ eine nicht-negative reelle Zahl sein, und nicht eine beliebige komplexe Zahl. Das Standardskalarprodukt auf dem $\mathbb{C}^n$,

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

hat offenbar diese Eigenschaft. Aber es ist *keine* Bilinearform, denn es gilt

$$\langle x, \lambda y \rangle = \overline{\lambda} \langle x, y \rangle \qquad \text{und}$$
$$\langle y, x \rangle = \overline{\langle x, y \rangle}.$$

Man verzichtet deshalb für komplexe Vektorräume auf die $\mathbb{C}$-Linearität im zweiten Argument und auch auf die Symmetrie und fordert nur folgendes:

Definition 10.8.1. Es sei V ein $\mathbb{C}$-Vektorraum und $\sigma : V \times V \to \mathbb{C}$ eine Funktionmit den Eigenschaften:

(S1) $\sigma(x_1 + x_2, y) = \sigma(x_1, y) + \sigma(x_2, y)$, $x_1, x_2, y \in V$.

(S2) $\sigma(\lambda x, y) = \lambda \sigma(x, y)$, $\lambda \in \mathbb{C}, x, y \in V$.

(S3) $\sigma(x, y_1 + y_2) = \sigma(x, y_1) + \sigma(x, y_2)$, $x, y_1, y_2 \in V$.

(S4) $\sigma(x, \lambda y) = \overline{\lambda} \sigma(x, y)$, $\lambda \in \mathbb{C}$, $x, y \in V$.

(S5) $\sigma(y, x) = \overline{\sigma(x, y)}$, $x, y \in V$.

(S6) $\sigma(x, x)$ ist reell, $\sigma(x, x) \geq 0$, und $\sigma(x, x) = 0$ nur für $x = 0$.

Ein σ mit den Eigenschaften (S1-S4) heißt *Sesquilinearform*; zusammen mit (S5) und (S6) heißt σ *hermitesch.* Ein V mit einer hermiteschen Sesquilinearform heißt *unitärer Vektorraum.*

Beispiel 10.8.2. (1) $V = \mathbb{C}^n$, $\sigma = \langle -, - \rangle$ Standardform

(2) $V = \mathbb{C}^n$, $\sigma(x, y) = x^\top S\overline{y}$, wobei S eine *hermitesche Matrix* ist, d.h. $S^\top = \overline{S}$ ($\overline{y}$ bzw. $\overline{S}$ ist die komplexe Konjugation), und S muss positiv-definit sein. (vgl. späteres Kriterium).

(3) $V = \mathrm{Mat}_{n,n}(\mathbb{C})$, $\sigma(A, B) = \mathrm{Spur}(A^\top \overline{B})$.

(4) $V = \mathcal{C}^0([-1, 1], \mathbb{C})$, $\sigma(f, g) = \int_{-1}^1 f(t)\overline{g(t)}\mathrm{d}t$.

Für unitäre Vektorräume gelten nun alle grundlegenden Eigenschaften, die wir auch für euklidische Vektorräume bewiesen haben:

1. $\|x\|_\sigma = \|x\| = \sqrt{\sigma(x, x)}$ ist eine Norm.
2. Schwarzsche Ungleichung:
$$|\sigma(x, y)| \le \|x\|_\sigma \cdot \|y\|_\sigma.$$
3. Gram-Schmidt-Orthonormalisierung.
4. Transformationsformel bei Basiswechsel mit $F = M_{\mathcal{B}'\mathcal{B}}(\mathrm{id}_V)$:
$$M_{\mathcal{B}'}(\sigma) = F^\top \cdot M_{\mathcal{B}}(\sigma) \cdot \overline{F}.$$

Beweis der „komplexen“ Schwarzschen Ungleichung: Sei $\lambda \in \mathbb{C}$:

$$\begin{aligned} 0 &\le \sigma(x - \lambda y, x - \lambda y) \\ &= \sigma(x, x) - \lambda\sigma(x, y) - \underbrace{\overline{\lambda}\sigma(y, x)}_{\overline{\lambda\sigma(x,y)}} + \overline{\lambda}\lambda\sigma(y, y) \\ &= \underbrace{\sigma(x, x)}_{\in\mathbb{R}} - 2\underbrace{\mathrm{Re}(\lambda\sigma(x, y)}_{\in\mathbb{R}}) + |\lambda|^2 \underbrace{\sigma(y, y)}_{\in\mathbb{R}} \end{aligned}$$

Speziell für $\lambda = \frac{\overline{\sigma(x,y)}}{\sigma(y,y)}$:

$$0 \le \sigma(x, x)\sigma(y, y) - 2\underbrace{\mathrm{Re}(\overline{\sigma(x, y)}\sigma(x, y))}_{|\sigma(x,y)|^2} + \underbrace{\overline{\overline{\sigma(x, y)}} \cdot \overline{\sigma(x, y)}}_{|\sigma(x,y)|^2},$$

also
$$|\sigma(x, y)| \le \|x\|_\sigma \cdot \|y\|_\sigma$$

10.9 Orthogonale und unitäre Abbildungen

Es sei V zunächst ein Vektorraum über enem beliebigen Körper $\mathbb{K}$, und σ sei eine Bilinearform. Wir betrachten alle Endomorphismen $f : V \to V$, die σ-*invariant* sind, d.h.

$$\sigma(f(x), f(y)) = \sigma(x, y)$$

gilt für alle $x, y \in V$. Sie bilden eine Teilmenge $\mathrm{End}(V, \sigma) \subseteq \mathrm{End}(V)$ aller Endomorphismen, und wir schreiben
$$\mathrm{Aut}(V, \sigma) = \mathrm{End}(V, \sigma) \cap \mathrm{Aut}(V)$$
Diese Menge ist offenbar eine Gruppe bezüglich Komposition von Abbildungen.
(Allgemeiner könnte man zwei $\mathbb{K}$-Vektorräume V, V' mit jeweiliger Bilinearform σ, σ' und alle Abbildungen $f : V \to V'$ mit $\sigma'(f(x), f(y)) = \sigma(x, y)$ betrachten.)
Von besonderem Interesse ist natürlich der Fall eines euklidischen bzw. unitären Vektorraumes.

Definition 10.9.1. Für einen euklidischen bzw. unitären Vektorraum V mit Skalarprodukt σ nennt man einen Endomorphismus $f : V \to V$ *orthogonal* bzw. *unitär*, falls σ invariant unter f bleibt.
Die Gruppe dieser Endomorphismen heißt *orthogonale Gruppe* $\mathrm{O}(V) = \mathrm{Aut}(V, \sigma)$ bzw. *unitäre Gruppe* $\mathrm{U}(V) = \mathrm{Aut}(V, \sigma)$.

Satz 10.9.2. *Für eine orthogonale bzw. unitäre Abbildung $f : V \to V$ gilt:*

(i) $\|f(x)\| = \|x\|$ *(*längentreu*).*

(ii) $\sphericalangle(f(x), f(y)) = \sphericalangle(x, y)$ *(*winkeltreu*). ($\mathbb{K} = \mathbb{R}$)*

(iii) Für alle Eigenwerte λ von f gilt: $|\lambda| = 1$.

Beweis. (i) $\|f(x)\| = \sqrt{\sigma(f(x), f(x))} = \sqrt{\sigma(x, x)} = \|x\|$.

(ii)

$$\sphericalangle(fx, fy) = \arccos\left(\frac{\sigma(fx, fy)}{\|fx\|\|fy\|}\right) = \arccos\left(\frac{\sigma(x, y)}{\|x\|\|y\|}\right) = \sphericalangle(x, y).$$

(iii) Gilt $f(x) = \lambda x$, $x \neq 0$, so folgt aus (i):

$$\|x\| = \|f(x)\| = \|\lambda x\| = |\lambda|\|x\|,$$

also $|\lambda| = 1$.

□

Ganz wichtig ist nun eine Schlussfolgerung aus (i):

Korollar 10.9.3. *Eine orthogonale bzw. unitäre Abbildung f ist stets invertierbar, und f^{-1} ist ebenfalls orthogonal bzw. unitär.*

Bemerkung 10.9.4. $\mathrm{End}(V, \sigma) = \mathrm{Aut}(V, \sigma)$ für ein euklidisches bzw. hermitesches Skalarprodukt.
Es sei $V = \mathbb{K}^n$, $\mathbb{K}$ beliebiger Körper, und die symmetrische Bilinearform σ auf V sei durch die symmetrische Matrix S gegeben, d.h. $\sigma(x, y) = x^\top Sy$. Ist nun f der durch die Matrix A (in der Standardbasis $\mathcal{E} = (e_1, \dots, e_n)$) gegebene Endomorphismus, so ist die Invarianz

$$\sigma(fx, fy) = \sigma(x, y), \qquad x, y \in V$$

äquivalent zur Gleichung

$$(Ax)^\top S(Ay) = x^\top A^\top SAy = x^\top Sy.$$

Setzt man $x = e_i, y = e_j$ $(i = 1, \dots, n; j = 1, \dots, n)$, so erhält man

$$(A^\top SA)_{ij} = S_{ij}$$

für alle i, j, also die Matrixgleichung

$$A^\top SA = S,$$

also

$$\mathrm{Aut}(V, \sigma) \cong \{A \in \mathrm{GL}_n(\mathbb{K}) \mid A^\top SA = S\}.$$

Beispiel 10.9.5. Beispiele und wichtige Notation:

1) $V = \mathbb{R}^n$, $\sigma = \langle -, - \rangle$, d.h. $S = \mathbb{1}$.
Die Gruppe $\mathrm{Aut}(V, \sigma) = \{A \in \mathrm{GL}_n(\mathbb{R}) \mid A^\top A = \mathbb{1}\}$ heißt *orthogonale Gruppe* (des $\mathbb{R}^n$) und wird mit $\mathrm{O}(n)$ bezeichnet. Diese Matrizen heißen *orthogonal*.

2) Die Untergruppe $\mathrm{SO}(n) = \{A \in \mathrm{GL}_n(\mathbb{R}) \mid A^\top A = \mathbb{1},\ \mathrm{Det}(A) = 1\}$ heißt *spezielle orthogonale Gruppe* des $\mathbb{R}^n$

3) $V = \mathbb{C}^n$, $\sigma = \langle -, - \rangle$, d.h. $S = \mathbb{1}$.
Die Gruppe $Aut(V, \sigma) = \{A \in \mathrm{GL}_n(\mathbb{C}) \mid A^\top \overline{A} = \mathbb{1}\}$ heißt *unitäre Gruppe* und wird mit $\mathrm{U}(n)$ bezeichnet. Diese Matrizen heißen *unitär*.

4) Die Untergruppe $\mathrm{SU}(n) = \{A \in \mathrm{GL}_n(\mathbb{C}) \mid A^\top \overline{A} = \mathbb{1},\ \mathrm{Det}(A) = 1\}$ heißt *spezielle unitäre Gruppe* des $\mathbb{C}^n$.

Notation: $V = \mathbb{K}$-Vektorraum, $\sigma =$ Bilinearform, $\sigma : V \times V \to \mathbb{K}$ (z.B. $\mathbb{K} = \mathbb{R}$: σ reelles Skalarprodukt; $\mathbb{K} = \mathbb{C}$ hermitesches Skalarprodukt)

(1)

$$\mathrm{Aut}(V, \sigma) := \{f \in \mathrm{Aut}(V) \mid \sigma(f(x), f(y)) = \sigma(x, y),\ \text{für alle } x, y \in V\}$$

(2) $V = \mathbb{K}^n$, $\sigma(x, y) = \overline{x}^\top Sy$, wobei $\mathbb{K} \to \mathbb{K}, \lambda \mapsto \overline{\lambda}$ ein Automorphismus der Ordnung 2 des Körpers $\mathbb{K}$ sei.

$$\mathrm{Aut}^-(V, \sigma) := \{A \in \mathrm{GL}_n(\mathbb{K}) \mid \overline{Ax}^\top SAy = \overline{x}^\top Sy\} = \{A \in \mathrm{GL}_n(\mathbb{K}) \mid \overline{A}^\top SA = S\}$$

Satz 10.9.6. *Es sei V ein euklidischer bzw. unitärer Vektorraum endlicher Dimension, und $\mathcal{B}$ sei eine Orthonormalbasis. Ein Endomorphismus $f : V \to V$ ist genau dann orthogonal bzw. unitär, wenn die Matrix $A = M_{\mathcal{B}\mathcal{B}}(f)$ orthogonal bzw. unitär ist.*

Beweis. Wir schreiben alles in der Basis $\mathcal{B}$: $A = M_{\mathcal{B}\mathcal{B}}(f)$, $M_\mathcal{B}(\sigma) = \mathbb{1}$, und $\tilde{x} = M_\mathcal{B}(x), \tilde{y} = M_\mathcal{B}(y)$. Dann ist $\sigma(fx, fy) = \sigma(x, y)$ äquivalent zu $(A\tilde{x})^\top \mathbb{1}(\overline{A\tilde{y}}) = \tilde{x}^\top \overline{\tilde{y}}$. Dies wiederum ist äquivvalent zur Matrizengleichung $A^\top \overline{A} = \mathbb{1}$. □

Korollar 10.9.7. *(i) Für einen orthogonalen bzw. unitären Endomorphismus f gilt $|\operatorname{Det}(f)| = 1$.*

(ii) Für eine orthogonale bzw. unitäre Matrix A gilt $|\operatorname{Det}(A)| = 1$.

Beweis. Aus $A^\top \overline{A} = \mathbb{1}$ folgt

$$1 = \operatorname{Det}(A^\top \overline{A}) = \operatorname{Det}(A^\top) \operatorname{Det}(\overline{A}) = \operatorname{Det}(A) \cdot \overline{\operatorname{Det}(A)} = |\operatorname{Det}(A)|^2.$$

□

Satz 10.9.8. *f ist genau dann orthogonal bzw. unitär, wenn $f\mathcal{B} = (fb_1, \ldots, fb_n)$ eine Orthonormalbasis ist für jede Orthonormalbasis $\mathcal{B} = (b_1, \ldots, b_n)$.*

Beweis. „⇒“ Es gilt dann $\sigma(fb_i, fb_j) = \sigma(b_i, b_j) = \delta_{ij}$, also ist $f\mathcal{B}$ eine Orthonormalbasis.
„⇐“ Sei $x = \sum_i x_i b_i$, $y = \sum_j y_j b_j$. Dann gilt einerseits

$$\sigma(x, y) = \sum_{i,j} \sigma(b_i, b_j) = \sum_{i,j} x_i y_j \delta_{ij},$$

und andererseits

$$\begin{aligned}
\sigma(fx, fy) &= \sigma(\sum_i x_i f(b_i), \sum_j y_j f(b_j)) \\
&= \sum_{i,j} x_i y_j \sigma(fb_i, fb_j) \\
&= \sum_{i,j} x_i y_j \delta_{ij},
\end{aligned}$$

denn nach Voraussetzung ist $f\mathcal{B} = (fb_1, \ldots)$, wieder eine Orthonormalbasis. □

Satz 10.9.9. *Ist f orthogonal bzw. unitär und sind λ_1, λ_2 zwei verschiedene Eigenwerte, so stehen die Eigenräume $\mathrm{Eig}_{\lambda_1}(f)$ und $\mathrm{Eig}_{\lambda_2}(f)$ orthogonal aufeinander.*

Beweis. Für $0 \neq x_1 \in \mathrm{Eig}_{\lambda_1}(f)$, $0 \neq x_2 \in \mathrm{Eig}_{\lambda_2}(f)$ erhalten wir

$$\sigma(x_1, x_2) = \sigma(fx_1, fx_2) = \sigma(\lambda_1 x_1, \lambda_2 x_2) = \lambda_1 \overline{\lambda_2} \sigma(x_1, x_2),$$

sowohl im reellen als auch im komplexen Fall. Wäre nun $\sigma(x_1, x_2) \neq 0$, so folgte $\lambda_1 \overline{\lambda_2} = 1$, also (wegen $|\lambda_1| = |\lambda_2| = 1$) sofort $\lambda_1 = \overline{\lambda_2}^{-1} = \lambda_2$, im Widerspruch zur Voraussetzung. □

Wir werden nun einige weitere Beispiele behandeln, nämlich die Gruppen $\mathrm{SO}(1) \subset \mathrm{O}(1)$, $\mathrm{SO}(2) \subset \mathrm{O}(2)$, $\mathrm{SU}(2)$ und ein Beispiel einer $\mathrm{Aut}(\mathbb{R}^2, \sigma)$ wobei σ kein Skalarprodukt ist.

Beispiel 10.9.10.

$$\begin{aligned}
\mathrm{O}(1) &= \{\mathrm{id}, -\mathrm{id}\} \cong \mathbb{Z}/2 \\
\mathrm{SO}(1) &= 1
\end{aligned}$$

Beispiel 10.9.11.

$$\mathrm{O}(2) \quad \cong \quad \underbrace{\{\text{Drehungen}\}}_{\text{Det}=+1} \quad \sqcup \quad \underbrace{\{\text{Drehspiegelungen}\}}_{\text{Det}=-1}$$

$$D_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \qquad S_u D_\alpha$$

$$\mathrm{SO}(2) \cong \{\text{Drehungen}\} = \{D_\alpha \mid 0 \leq \alpha < \pi\} \cong \mathbb{S}^1,$$

wobei der letzte Isomorphismus gegeben ist durch $D_\alpha \mapsto e^{\alpha i}$.

Beispiel 10.9.12. $\mathrm{SU}(2) \cong \mathbb{S}^3$:
Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SU}(2) \subset \mathbb{C}^4$. Die Bedingungen an A sind:

(1) $\| \begin{pmatrix} a \\ c \end{pmatrix} \| = a\overline{a} + c\overline{c} = |a|^2 + |c|^2 = 1.$

(2) $\| \begin{pmatrix} b \\ d \end{pmatrix} \| = b\overline{b} + d\overline{d} = |b|^2 + |d|^2 = 1.$

(3) $\langle \begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix} \rangle = a\overline{b} + c\overline{d} = 0.$

(4) $\mathrm{Det}(A) = ad - cb = 1$

Mit

$$\begin{aligned} a &= a_1 + \sqrt{-1}a_2, \\ b &= b_1 + \sqrt{-1}b_2, \qquad \text{usw.} \end{aligned}$$

sind (1) und (2) äquivalent zu $\begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix} \in \mathbb{S}^3$.
Und (3) und (4) ergeben ein LGS über $\mathbb{C}$ in den Unbekannten b, d:

$$\begin{aligned} \overline{a} \cdot \boxed{b} + \overline{c} \cdot \boxed{d} &= 0 \\ -c \cdot \boxed{b} + a \cdot \boxed{d} &= 1 \end{aligned}$$

$$L = \begin{pmatrix} \overline{a} & \overline{c} \\ -c & a \end{pmatrix}, \qquad \mathrm{Det}(L) = \overline{a}a + c\overline{c} = 1.$$

Eindeutige Lösungen nach Cramerscher Regel:

$$b = \frac{\begin{vmatrix} 0 & \overline{c} \\ 1 & a \end{vmatrix}}{\mathrm{Det}(b)} = -\overline{c}, \qquad d = \frac{\begin{vmatrix} \overline{a} & 0 \\ -c & 1 \end{vmatrix}}{\mathrm{Det}(b)} = \overline{a}$$

Also: $A = \begin{pmatrix} a & -\overline{c} \\ c & \overline{a} \end{pmatrix}$ für beliebiges $\begin{pmatrix} a \\ c \end{pmatrix} \in \mathbb{S}^3$. Damit ist

$$\mathrm{SU}(2) \longrightarrow \mathbb{S}^3, \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \longmapsto \begin{pmatrix} a \\ c \end{pmatrix}$$

eine Bijektion.

Beispiel 10.9.13. $\mathrm{SO}(3) \cong \mathbb{R}P^3$:

$$A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \in \mathrm{SO}(3) \subset \mathbb{R}^9$$

Bedingungen: Wir bezeichnen die Spalten von A mit $a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, c = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$.

(1) $\|a\| = \|b\| = \|c\| = 1$ (3 Bedingungen)

(2) $\langle a, b\rangle = \langle b, c\rangle = \langle c, a\rangle = 0$ (3 Bedingungen)

$9 - 6 = 3 = \dim \mathrm{SO}(3)$

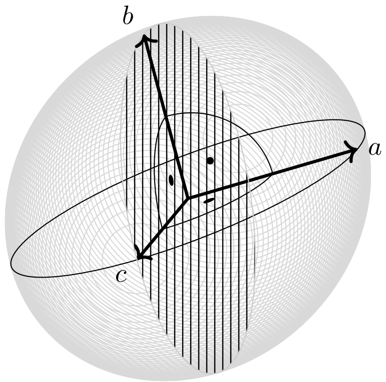
Geometrisch (siehe Zeichnung):

(1) $a \in \mathbb{S}^2$.

(2) $b \in \mathbb{S}^2 \cap \mathrm{Span}(a)^\perp$

(3) $c \in \mathbb{S}^2 \cap \mathrm{Span}(a, b)^\perp$.

Für c gibt es nur eine Wahl, nämlich das Kreuzprodukt $c = a \times b$ zweier Vektoren in $\mathbb{R}^3$. Wenn man die Neigungswinkel („*Eulersche Winkel*") von a, b, c zu den Standardachsen e_1, e_2, e_3 zu Hilfe nimmt, kann man $\mathrm{SO}(3) \cong \mathbb{R}P^3$ zeigen.



Beispiel 10.9.14. Nun sei $\mathbb{K} = \mathbb{R}$, $V = \mathbb{R}^2$ und $\sigma(x, y) = x^\top S y$ für $S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

$$\begin{aligned} \sigma(x, y) &= x_1 y_1 - x_2 y_2 \\ \sigma(e_1, e_1) &= +1 \\ \sigma(e_2, e_2) &= -1 \end{aligned}$$

σ ist

- symmetrisch
- weder positiv- noch negativ-definit („indefinit")

Was ist $\mathrm{Aut}(V, \sigma)$?

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad A^\top S A = S$$

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\text{also} \qquad \begin{pmatrix} a^2 - c^2 & ab - cd \\ ab - cd & b^2 - d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Damit erhalten wir

$$a^2 - c^2 = 1 \tag{10.5}$$

$$b^2 - d^2 = -1 \tag{10.6}$$

$$ab - cd = 0 \tag{10.7}$$

Aus Gleichung (10.7) erhält man $a^2 b^2 = c^2 d^2$, also unter Verwendung von (10.6) und (10.5)

$$0 = a^2 b^2 - c^2 d^2 = a^2(d^2 - 1) - d^2(a^2 - 1),$$

also $a^2 = d^2$.

Lösung 1: $a = d$; *Lösung 2*: $a = -d$.
Für $a \leq -1$ oder $a \geq 1$

$$\text{Lösung 1:} \qquad A = \begin{pmatrix} a & \pm\sqrt{a^2-1} \\ \pm\sqrt{a^2-1} & a \end{pmatrix} \qquad \det A = +1$$

$$\text{Lösung 2:} \qquad A = \begin{pmatrix} a & \pm\sqrt{a^2-1} \\ \mp\sqrt{a^2-1} & -a \end{pmatrix} \qquad \det A = -1$$

Mit $a = \frac{1}{\sqrt{1-t^2}}$ $(-1 < t < 1)$ erhalten wir für Lösung 1:

$$A(t) = \begin{pmatrix} \frac{1}{\sqrt{1-t^2}} & \frac{t}{\sqrt{1-t^2}} \\ \frac{t}{\sqrt{1-t^2}} & \frac{1}{\sqrt{1-t^2}} \end{pmatrix}, \qquad \operatorname{Det} A(t) = 1,$$

$$A(s)A(t) = A\left(\frac{s+t}{1+st}\right)$$

Additionstheorem des Tangens hyperbolicus:

$$\rho(s+t) = \frac{\rho(s)+\rho(t)}{1+\rho(s)\rho(t)}.$$

$$\rho(t) = \frac{e^t - e^{-t}}{e^t + e^{-t}} = \frac{\sinh(t)}{\cosh(t)} = \tanh(t)$$

$$\begin{aligned} \cosh(x+y) &= \cosh(x)\cosh(y) + \sinh(x)\sinh(y) \\ \sinh(x+y) &= \sinh(x)\sinh(y) + \cosh(x)\cosh(y) \end{aligned}$$

$$\cosh^2(x) - \sinh^2(x) = 1$$

$$1 = \sigma(v,v) = v_1^2 - v_2^2$$

Also ist

$$\mathbb{R} \longrightarrow \operatorname{Aut}(V,\sigma), \qquad t \longmapsto A(\rho(t))$$

ein Homomorphismus von Gruppen.

Wir kommen nun zu den sogenannten Normalformensätzen. Für einen unitären Endomorphismus gilt nun:

Satz 10.9.15. *Zu jedem unitären Endomorphismus $f : V \to V$ eines unitären Vektorraumes V endlicher Dimension gibt es eine Orthonormalbasis $\mathcal{B}$ aus Eigenvektoren, so dass gilt:*

$$A = M_{\mathcal{B}\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Beweis. $n = 1$: Es muss $f(x) = \lambda x$ für alle $x \in V$ und ein $\lambda \in \mathbb{S}^1$ gelten; jedes b_1 mit $\|b_1\| = 1$ ist eine Orthonormalbasis.
$n \geq 2$: Das charakteristische Polynom $\mathrm{CP}_f(t) = (\lambda_1 - t) \cdot \ldots \cdot (\lambda_n - t)$ zerfalle und b_1 sei ein Eigenvektor zum Eigenwert λ_1, und es sei $\|b_1\| = 1$. Entscheidend ist nun, dass es zu dem invarianten Unterraum $\operatorname{Span}(b_1)$ ein invariantes Komplement gibt, nämlich das orthogonale Komplement $U = \operatorname{Span}(b_1)^\perp$. Die Einschränkung $f'|_U : U \to U$ ist unitär, und $\dim U = \dim V - 1 = n - 1$. Ist also $\mathcal{B}' = (b_2, \ldots, b_n)$ eine Orthonormalbasis von U aus Eigenvektoren von f', so ist $\mathcal{B} = (b_1, b_2, \ldots, b_n)$ die gesuchte Basis. $\square$

Korollar 10.9.16. *Jeder unitäre Endomorphismus ist diagonalisierbar.*

Für eine unitäre Matrix $B \in U(n)$ besagt Satz 10.9.15, dass es eine invertierbare Matrix $W \in \mathrm{GL}_n(\mathbb{C})$ gibt, so dass $A = WBW^{-1} = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$, also konjugiert ist zu einer Diagonalmatrix. Nun wollen wir noch sehen, dass A kongruent ist zu einer Diagonalmatrix und zwar vermöge einer unitären Matrix. Vorher wollen wir Ähnlichkeit und Kongruenz noch klar gegeneinander abheben: Ähnlichkeit kennen wir von Diagonalisierbarkeitsfragen, und Kongruenz aus der Transformationsformel für Bilinearformen.

Definition 10.9.17. Es seien $A, B \in \mathrm{Mat}_{n,n}(\mathbb{K})$.

(i) A und B heißen vermöge $W \in \mathrm{GL}_n(\mathbb{K})$ *ähnlich* zueinander (*konjugiert* zueinander), falls

$$A = WBW^{-1}$$

gilt.

(ii) A und B heißen vermöge $W \in \mathrm{GL}_n(\mathbb{K})$ *kongruent*, falls

$$A = W^\top B\overline{W}$$

gilt. (Hier ist $\overline{W} = W$ zu setzen, falls $\mathbb{K} \neq \mathbb{C}$)

Korollar 10.9.18. *Zu jeder Matrix $B \in \mathrm{U}(n)$ gibt es eine Matrix $W \in \mathrm{U}(n)$ mit*

$$A = W^\top B\overline{W} = \mathrm{Diag}(\lambda_1, \dots, \lambda_n).$$

Beweis. Die Spalten $W_1, \dots, W_n$ von Ω seien die Eigenvektoren von B, die nach dem Satz eine Orthonormalbasis $\mathcal{B} = (w_1, \dots, w_n)$ bilden. Also ist $\Omega \in \mathrm{U}(n)$, d.h. $\overline{\Omega}^\top = \Omega^{-1}$. Es gilt auch $\Omega = M_{\mathcal{B}\mathcal{E}}(\mathrm{id}_{\mathbb{C}^n})$ für den Basiswechsel von der Standardbasis $\mathcal{E} = (e_1, \dots, e_n)$ des $\mathbb{C}^n$ zu $\mathcal{B} = (w_1, \dots, w_n)$. Es folgt nun mit $B = M_{\mathcal{E}\mathcal{E}}(f)$, $fx = Bx$ sofort:

$$\begin{aligned} \Omega B\overline{\Omega}^\top &= M_{\mathcal{B}\mathcal{E}}(\mathrm{id}) M_{\mathcal{E}\mathcal{E}}(f) \overline{M_{\mathcal{B}\mathcal{E}}(\mathrm{id})}^\top \\ &= M_{\mathcal{B}\mathcal{E}}(\mathrm{id}) M_{\mathcal{E}\mathcal{E}}(f) M_{\mathcal{B}\mathcal{E}}(\mathrm{id})^{-1} \\ &= M_{\mathcal{B}\mathcal{E}}(\mathrm{id}) M_{\mathcal{E}\mathcal{E}}(f) M_{\mathcal{E}\mathcal{B}}(\mathrm{id}) \\ &= M_{\mathcal{B}\mathcal{B}}(f) = A = \mathrm{Diag}(\lambda_1, \dots, \lambda_n). \end{aligned}$$

Nun setzt man noch $W = \overline{\Omega}$. □

Bemerkung 10.9.19. Für ein orthogonales bzw. unitäre W sind die Bedingungen $A = WBW^{-1}$ und $A = W^\top B\overline{W}$ äquivalent, denn $W^{-1} = \overline{W}^\top$. Diesen Sonderfall wird man „Hauptachsentransformation“ nennen. Im Allgemeinen sind die beiden Äquivalenzrelationen aber verschieden.

Beispiel 10.9.20. Um also eine unitäre Matrix B zu diagonalisieren, sucht man zunächst einen Eigenvektor b_1 zu einem Eigenwert λ_1

$$(B - \lambda_1 \mathbb{1})b_1' = 0,$$

den man dann normiert: $b_1 := \frac{b_1'}{\|b_1'\|}$.
Dann sucht man einen zweiten Eigenvektor b_2' zu λ_2, der aber orthogonal zu b_1' sein soll:

$$\begin{aligned} (B - \lambda_2 \mathbb{1})b_2' &= 0, \\ \langle b_1', b_2' \rangle &= 0. \end{aligned}$$

Dieser wird dann auch normiert $b_2 := \frac{b_2'}{\|b_2'\|}$.
Im nächsten Schritt sucht man einen Eigenvektor b_3' zu λ_3, der nun orthogonal zu b_1' und b_2' sein soll:

$$\begin{aligned} (B - \lambda_3 \mathbb{1})b_3' &= 0, \\ \langle b_1', b_3' \rangle &= 0, \\ \langle b_2', b_3' \rangle &= 0. \end{aligned}$$

Der letzte Eigenvektor zu λ_n soll dann orthogonal zu $b_1', \dots, b_{n-1}'$ sein.

Nun kommen wir zu dem entsprechenden Satz über reelle Matrizen.

Satz 10.9.21. *Zu jedem orthogonalen Endomorphismus $f : V \to V$ eines euklidischen Vektorraumes V endlicher Dimension gibt es eine Orthonormalbasis $\mathcal{B}$, so dass gilt:*

$$A = M_{\mathcal{B}\mathcal{B}}(f) = \begin{pmatrix} \underbrace{\boxed{\begin{matrix} +1 & & \\ & \ddots & \\ & & +1 \end{matrix}}}_{p} & & & & 0 \\ & \underbrace{\boxed{\begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix}}}_{q} & & & \\ & & \boxed{A_1} & & \\ 0 & & & \ddots & \\ & & & & \boxed{A_r} \end{pmatrix},$$

wobei $p \geq 0, q \geq 0, r \geq 0$ und $p + q + 2r = n$ gilt, und die A_k Drehmatrizen von der Form

$$\begin{aligned} A_k &= \begin{pmatrix} \cos\alpha_k & -\sin\alpha_k \\ \sin\alpha_k & \cos\alpha_k \end{pmatrix}, \qquad 0 < \alpha_k < 2\pi, \qquad \alpha_k \neq \pi \\ \textit{oder} \qquad A_k &= \begin{pmatrix} -\cos\alpha_k & -\sin\alpha_k \\ -\sin\alpha_k & \cos\alpha_k \end{pmatrix} \end{aligned}$$

sind.

Beweis. Man verfährt wieder induktiv und spaltet 1-dimensionale bzw. 2-dimensionale invariante Unterräume ab, deren orthogonale Komplemente nun auch invariant sind. Im ersten Fall erhält man einen Eigenwert $\lambda = \pm 1$ (und einen Eigenvektor b mit $\|b\| = 1$) in der Basis. Im zweiten Fall erhält man eine Ebene, in der man eine Orthonormalbasis wählt. □

Korollar 10.9.22. *Zu jeder Matrix $B \in \mathrm{O}(n)$ gibt es eine Matrix $W \in \mathrm{GL}_n(\mathbb{R})$ mit $A = WBW^{-1}$ von der obigen Gestalt.*

Korollar 10.9.23. *Zu jeder Matrix $B \in \mathrm{O}(n)$ gibt es eine orthogonale Matrix $W \in \mathrm{O}(n)$ mit $A = W^\top BW$ von der obigen Gestalt.*

Beweis. Man übernimmt den Beweis wie im Beweis von Korollar 10.9.18; der einzige Unterschied besteht darin, dass nur m Vektoren der Orthonormalbasis $\mathcal{B}$ Eigenvektoren sind. Dies spielt aber für den Beweis keine Rolle. □

11 Hauptachsentransformation

11.1 Symmetrische Matrizen

Wir benutzen symmetrische Matrizen dazu, symmetrische Bilinearformen darzustellen:

$$\boxed{\text{symmetrische Bilinearformen}} \qquad \boxed{\text{symmetrische Matrizen}}$$

$$\begin{array}{rcl} \operatorname{Sym}(2,V) & \underset{L_{\mathcal{B}}=\sigma_{\mathcal{B}}}{\overset{M_{\mathcal{B}}}{\rightleftarrows}} \cong & \operatorname{Sym}_n(\mathbb{K}) \\[1em] \sigma & \longmapsto & M_{\mathcal{B}}(\sigma) = S \\[1em] \sigma(x,y) = K_{\mathcal{B}}(x)^\top S K_{\mathcal{B}}(y) & \longleftarrow\!\shortmid & S \end{array}$$

Will man symmetrische Bilinearformen klassifizieren, so muss man also symmetrische Matrizen studieren.

Definition 11.1.1. Zwei Matrizen A und B heißen *kongruent* wenn es eine invertierbare Matrix $\Omega \in \mathrm{GL}_n(\mathbb{K})$ gibt mit $\Omega^\top A \Omega = B$.
Im euklidischen Fall $\mathbb{K} = \mathbb{R}$ (wo σ eine Bilinearform ist) ist „A kongruent B, vermöge einer orthogonalen Matrix Ω“ (*orthogonal-kongruent*) gleichbedeutend mit „A äquivalent B vermöge einer orthogonalen Matrix Ω“ (*orthogonal-äquivalent*).
Im unitären Fall $\mathbb{K} = \mathbb{C}$ (wo σ nur eine Sesquilinearform ist) lautet die Definition: A kongruent B, wenn es eine invertierbare Matrix $\Omega \in \mathrm{GL}_n(\mathbb{C})$ gibt mit $\Omega^\top A \overline{\Omega} = B$. Analog zum euklidischen Fall ist unitär-kongruent gleichbedeutend mit unitär-äquivalent.

Beispiel 11.1.2.

1) $\mathbb{K} = \mathbb{C}$: Ist A unitär, so ist es unitär-kongruent zu einer Diagonalmatrix (d.h. unitär-diagonalisierbar), nach Korollar 10.9.18.

2) $\mathbb{K} = \mathbb{R}$: Ist A orthogonal, so ist es orthogonal kongruent zu einer Diagonalmatrix (d.h. orthogonal-diagonalisierbar) genau dann, wenn das charakteristische Polynom zerfällt, nach Satz (10.9.21).
Eine andere notwendige Bedingung für orthogonale Diagonalisierbarkeit ist: $A = A^\top$, denn ist

   - $\Omega A \Omega^{-1} = D$ diagonal und $\Omega^{-1} = \Omega^\top$ orthogonal,

   so folgt:

$$\begin{aligned} A^\top &= (\Omega^{-1} D \Omega)^\top = \Omega^\top D^\top (\Omega^{-1})^\top \\ &= \Omega^{-1} D \Omega = A. \end{aligned}$$

3) $\mathbb{K} = \mathbb{R}$: Wie man an einer Drehmatrix $D = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$ sieht, sind nicht alle orthogonalen Matrizen symmetrisch.

4) $\mathbb{K} = \mathbb{C}$: mit der gleichen Rechnung folgt auch (vgl. nächstes Lemma)

$$A \text{ unitär} \iff A \text{ unitär-diagonalisierbar mit nur reellen Eigenwerten} \iff \overline{A}^\top = A \text{ (hermitesch)}.$$

Lemma 11.1.3.

(i) Alle Eigenwerte einer reellen symmetrischen Matrix A sind reell, wenn man $\mathrm{CP}_A(t)$ also komplexes Polynom auffaßt. Insbesondere zerfällt $\mathrm{CP}_A(t)$ über $\mathbb{R}$.

(ii) Alle Eigenwerte einer komplexen hermitischen Matrix A sind reell.

Beweis. (i) folgt aus (ii).

(ii) Ist $0 \neq x \in \mathbb{C}^n$ ein Eigenvektor zm Eigenwert $\lambda \in \mathbb{C}$, so gilt:

$$\lambda\langle x, x\rangle = \langle \lambda x, x\rangle = \langle Ax, x\rangle = (Ax)^\top \overline{x} = x^\top A^\top \overline{x} = x^\top \overline{A}\overline{x} = x^\top \overline{(Ax)} = \langle x, Ax\rangle = \langle x, \lambda x\rangle = \overline{\lambda}\langle x, x\rangle.$$

Weil $\langle x, x\rangle \neq 0$, muss $\lambda = \overline{\lambda}$ sein.

□

Lemma 11.1.4. *Sind $\lambda_1, \lambda_2 \in \operatorname{Spek}(A)$ zwei verschiedene Eigenwerte einer reellen symmetrischen bzw. komplexen hermiteschen Matrix A, so gilt $\operatorname{Eig}_{\lambda_1}(A) \perp \operatorname{Eig}_{\lambda_2}(A)$.*

Beweis. Für $x_1 \in \operatorname{Eig}_{\lambda_1}(A), x_2 \in \operatorname{Eig}_{\lambda_2}(A)$ gilt

$$\begin{aligned}\overline{\lambda_2}\langle x_1, x_2\rangle &= \langle x_1, \lambda x_2\rangle = \langle x_1, Ax_2\rangle \\ &= x_1^\top \overline{Ax_2} = x_1^\top \overline{A}\overline{x_2} = x_1^\top A^\top \overline{x_2} = (Ax_1)^\top \overline{x_2} \\ &= \langle Ax_1, x_2\rangle = \langle \lambda_1 x_1, x_2\rangle = \lambda_1\langle x_1, x_2\rangle.\end{aligned}$$

Wegen $\lambda_1 \neq \lambda_2$ und $\lambda_1 = \overline{\lambda_1}$, $\lambda_2 = \overline{\lambda_2}$ (nach Lemma 11.1.3) folgt $\langle x_1, x_2\rangle = 0$. □

Lemma 11.1.5. *Ist $U \subseteq \mathbb{R}^n$ bzw. $U \subseteq \mathbb{C}^n$ invariant unter A, so auch $U^\perp$.*

Beweis. Für $u \in U$, $x \in U^\perp$ gilt:

$$\begin{aligned}\langle u, Ax\rangle &= u^\top \overline{Ax} = u^\top \overline{A}\overline{x} = u^\top A^\top \overline{x} = (Au)^\top \overline{x} \\ &= \langle Au, x\rangle.\end{aligned}$$

Ist $Au \in U$, so ist $\langle Au, x\rangle = 0$. Damit ist $\langle u, Ax\rangle = 0$, also $Ax \in U^\perp$. □

Satz 11.1.6 (Hauptsatz über reelle symmetrische Matrizen)**.** *Für ein $A \in \operatorname{Mat}_{n,n}(\mathbb{R})$ sind äquivalent:*

(i) A ist symmetrisch.

(ii) A ist orthogonal-diagonalisierbar (d.h. es gibt ein $W \in \operatorname{O}(n)$ mit $WAW^{-1} = \operatorname{Diag}(\lambda_1, \dots, \lambda_n)$).

(iii) Es gibt eine Orthogonalbasis des $\mathbb{R}^n$ aus Eigenvektoren von A.

Beweis. (ii) und (iii) sind offenbar äquivalent. Oben haben wir schon *(ii)* ⇒ *(i)* gezeigt; bleibt nur noch *(i)* ⇒ *(ii)* zu zeigen.
Nach Lemma 11.1.3 zerfällt das charakteristische Polynom

$$\operatorname{CP}_A(t) = (\lambda_1 - t) \cdot \ldots \cdot (\lambda_n - t)$$

in Linearfaktoren mit $\lambda_i \in \mathbb{R}$. Wir wählen einen Eigenvektor b_1 mit $\|b_1\| = 1$. Dann ist $U = \operatorname{Span}(b_1)$ invariant und nach Lemma 11.1.5 auch $U^\perp$. Ist $w_2, w_3, \dots, w_n$ irgendeine Basis von $U^\perp$, dann ist die Matrix $A' = (a'_{ij})$ mit

$$a'_{ij} := w_i^\top A w_j \qquad (2 \leq i, j \leq n)$$

offenbar symmetrisch. Nach Induktionsvoraussetzung mit $\dim U^\perp = n-1$ gibt es eine Orthonormalbasis $b_2, \dots, b_n$ von $U^\perp \cong \mathbb{R}^{n-1}$ aus Eigenvektoren von A'_j, diese sind auch Eigenvektoren von A in $\mathbb{R}^n$. Damit ist $\mathcal{B} = (b_1, b_2, \dots, b_n)$ die gesuchte Basis. □

Verzichtet man auf die Orthogonalität des diagonalisierenden Basiswechsels W, so erhält man den

Satz 11.1.7 (Trägheitssatz von Sylvester)**.** *Für eine reelle symmetrische Matrix A gibt es ein $R \in \operatorname{GL}_n(\mathbb{R})$ mit*

$$RAR^{-1} = \operatorname{Diag}(\underbrace{+1, \dots, +1}_{p}, \underbrace{-1, \dots, -1}_{q}, \underbrace{0, \dots, 0}_{n-p-q}).$$

Bemerkung 11.1.8. p und q sind eindeutig durch A bestimmt, als Dimension der Eigenräume zu $+1$ bzw. -1. Es ist $p + q = \operatorname{rg}(A)$. Man nennt $p - q$ die *Signatur* oder den *Trägheitsindex* von A.

Beweis. Nach dem Hauptsatz wissen wir bereits

$$A = W \operatorname{Diag}(\lambda_1, \ldots, \lambda_n) W^{-1}$$

mit $W^{-1} = W^\top$ orthogonal.
Nun sortiert man die Eigenwerte durch eine Permutation $\pi \in \mathfrak{S}_n$ in die positiven $\lambda_{\pi(1)}, \ldots, \lambda_{\pi(p)} > 0$, die negativen $\lambda_{\pi(p+1)}, \ldots, \lambda_{\pi(p+q)} < 0$ und in die Nullen $\lambda_{\pi(p+q+1)}, \ldots, \lambda_{\pi(n)} = 0$. Es sei $P = P_\pi$ die zugehörige Permutationsmatrix; P ist orthogonal, also $P^\top = P^{-1}$.
Dann normiert man noch die Diagonalmatrix:

$$N := \operatorname{Diag}\left(\frac{1}{\sqrt{\lambda_{\pi(1)}}}, \ldots, \frac{1}{\sqrt{\lambda_{\pi(p)}}}, \frac{1}{\sqrt{-\lambda_{\pi(p+1)}}}, \ldots, \frac{1}{\sqrt{-\lambda_{\pi(p+q)}}}, 1, \ldots, 1 \right).$$

N ist natürlich symmetrisch, aber nur dann orthogonal, wenn $|\lambda_{\pi(1)}| = \ldots = |\lambda_{\pi(p+q)}| = 1.$ gilt. Zusammen haben wir

$$NPWAW^{-1}P^{-1}N^{-1} = \operatorname{Diag}(+1, \ldots, +1, -1, \ldots, -1, 0, \ldots, 0)$$

und $R = NPW \in \mathrm{GL}_n(\mathbb{R})$, aber im Allgemeinen nicht orthogonal. □

Bemerkung 11.1.9. Sei $\sigma(x,y) = x^\top Ay$ durch die symmetrische Matrix $A \in \mathrm{GL}_n(\mathbb{R})$ gegeben, und sei $RAR^{-1} = \operatorname{Diag}(+1, \ldots, -1, \ldots, 0, \ldots)$ wie oben. Dann ist σ

- positiv-definit auf $R^{-1} \operatorname{Span}(e_1, \ldots, e_p) R$,
- negativ-definit auf $R^{-1} \operatorname{Span}(e_{p+1}, \ldots, e_{p+q}) R$,
- isotrop auf $R^{-1} \operatorname{Span}(e_{p+q}, \ldots, e_n) R$.

Beispiel 11.1.10. Die Bilinearformen $\sigma_p(x,y) = \sum_{i=1}^p x_i y_i - \sum_{i=p+1}^n x_i y_i$ auf $V = \mathbb{K}^n$ sind nicht-ausgeartet $(p + q = n)$ und haben Signatur $p - q = p - (n - p) = 2p - n$.

11.2 Adjungierte Abbildung

Die Transponierte $A^\top$ einer $(m \times n)$-Matrix A ist die Matrix der dualen Abbildung:

$$M_{\mathcal{B}',\mathcal{B}}(f)^\top = M_{\mathcal{B}^*,\mathcal{B}'^*}(f^*)$$

in der Situation

$$\begin{tikzcd} V, \mathcal{B} \arrow[r, "\Phi_{\mathcal{B}}", "\cong"'] \arrow[d] & V^*, \mathcal{B}^* \\ W, \mathcal{B}' \arrow[r, "\Phi_{\mathcal{B}'}", "\cong"'] & W^*, \mathcal{B}'^* \arrow[u, "f^*"'] \end{tikzcd}$$

mit dem Isomorphismu $\Phi_{\mathcal{B}}$, der durch $b_i \mapsto b_i^* = K_{\mathcal{B}}^i$ gegeben ist $(b_i \in \mathcal{B}, b_i^* \in \mathcal{B}^*)$.

Besitzt V ein nicht-ausgeartete symmetrische Bilinearform σ (und W ein σ'), so gibt es einen weiteren Isomorphismus

$$\Psi_\sigma : V \longrightarrow V^*, \qquad \Psi_\sigma(x) : y \longmapsto \sigma(x,y).$$

Für die Paarung

$$\varepsilon_\sigma : V^* \times V \longrightarrow \mathbb{K}, \qquad (\alpha, v) \longmapsto \varepsilon_\sigma(\alpha, v) := \alpha(v)$$

gibt es eine *Adjunktionsformel*

$$\varepsilon_{\sigma'}(\beta, f(v)) = \beta(f(v)) = f^*(\beta)(v) = \varepsilon_\sigma(f^*(\beta), v)$$

für $\beta \in W^*, v \in V$.

$$\begin{tikzcd} V \arrow[r, "\Psi_\sigma"] \arrow[d, shift left, "f"] & V^* \\ W \arrow[r, "\Psi_{\sigma'}"] \arrow[u, bend left, "f^{\mathrm{ad}}"] & W^* \arrow[u, "f^*"'] \end{tikzcd}$$

Für die lineare Abbildung
$$f^{\mathrm{ad}} := \Psi_\sigma^{-1} \circ f^* \circ \Psi_{\sigma'} : W \longrightarrow V$$
gilt also $\sigma(f^{\mathrm{ad}}(x'), y) = \sigma'(x', f(y))$.

Definition 11.2.1. Zwei lineare Abbildungen $f : V \to W$ und $g : W \to V$ heißen *adjungiert zueinander*, falls
$$\sigma(x, g(x')) = \sigma'(f(x), x').$$
für $x \in V, x' \in W$ gilt.

Lemma 11.2.2. *(i) Für zwei euklidische Vektorräume gibt es zu jedem $f : V \to W$ genau eine adjungierte Abbildung $f^{\mathrm{ad}} = \Psi_\sigma^{-1} \circ f^* \circ \Psi_{\sigma'} : W \to V$.*

(ii) Für zwei unitäre Vektorräume gibt es zu jedem $f : V \to W$ genau eine adjungierte Abbildung $f^{\mathrm{ad}} = \Psi_\sigma^{-1} \circ f^* \circ \Psi_{\sigma'} : W \to V$, wobei hier $\Psi_\sigma(v)(y) = \sigma(y, x)$ zu setzen ist.

Offenbar ist $M_{\mathcal{B},\mathcal{B}'}(f^{\mathrm{ad}}) = M_{\mathcal{B}',\mathcal{B}}(f)^\top$ für zwei Orthonormalbasen $\mathcal{B} \subseteq V$, $\mathcal{B}' \subseteq W$. Und es gilt:

1) $0^{\mathrm{ad}} = 0$, $\mathrm{id}^{\mathrm{ad}} = \mathrm{id}$,

2) $(f_2 \circ f_1)^{\mathrm{ad}} = f_1^{\mathrm{ad}} \circ f_2^{\mathrm{ad}}$, $(f^{\mathrm{ad}})^{\mathrm{ad}} = f$,

3) $(\lambda f)^{\mathrm{ad}} = \overline{\lambda} f^{\mathrm{ad}}$, $(f_1 + f_2)^{\mathrm{ad}} = f_1^{\mathrm{ad}} + f_2^{\mathrm{ad}}$.

Lemma 11.2.3. *(i)* $\mathrm{im}(f^{\mathrm{ad}}) = \ker(f)^\perp$.

(ii) $\ker(f^{\mathrm{ad}}) = \mathrm{im}(f)^\perp$.

Definition 11.2.4. Ein Endomorphismus $f : V \to V$ heißt *selbstadjungiert*, wenn $f = f^{\mathrm{ad}}$ gilt.

Offenbar ist f genau dann selbstadjungiert, wenn $M_{\mathcal{B}\mathcal{B}}(f) = A$ eine symmetrische (hermitesche) Matrix ist (für $\mathcal{B}$ eine Orthonormalbasis). Also haben wir:

Satz 11.2.5. *Für eine selbstadjungierten Endomorphismus $f : V \to V$ eines euklidischen bzw. unitären Vektorraumes V gibt es eine Orthonormalbasis aus Eigenvektoren.*

11.3 Normale Abbildungen

Definition 11.3.1. Es sei V ein euklidischer bzw. unitären Vektorraum. Ein Endomorphismus $f : V \to V$ heißt *normal*, falls $f \circ f^{\mathrm{ad}} = f^{\mathrm{ad}} \circ f$ gilt.

Beispiel 11.3.2.

1) f selbstadjungiert $\implies$ f normal.

2) $\mathbb{K} = \mathbb{C}$: f unitär $\implies$ f unitär-diagonalisierbar $\implies$ f normal.

3) $\mathbb{K} = \mathbb{C}$: f orthogonal-diagonalisierbar $\implies$ f normal.

Lemma 11.3.3. *Sei $f : V \to V$ normal.*

(i) $\ker(f) = \ker(f^{\mathrm{ad}})$.

(ii) $\mathrm{im}(f) = \mathrm{im}(f^{\mathrm{ad}})$

Beweis. Nach Lemma 10.3.4 sind (i) und (ii) äquivalent. Für ein $x \in \ker(f)$ gilt
$$0 = \sigma(fx, fx) = \sigma(x, (f^{\mathrm{ad}} \circ f)x) = \sigma(x, (f \circ f^{\mathrm{ad}})x) = \sigma(f^{\mathrm{ad}}x, f^{\mathrm{ad}}x).$$
Also ist $f^{\mathrm{ad}}(x) = 0$, und demnach $\ker(f) \subseteq \ker(f^{\mathrm{ad}})$. Die umgekehrte Inklusion folgt durch Vertauschen der Rollen von f und f^{ad}. □

Lemma 11.3.4. *Für ein normales $f : V \to V$ gilt*

(i) $\overline{\mathrm{Spek}(f)} = \mathrm{Spek}(f^{\mathrm{ad}})$,

(ii) $\mathrm{Eig}_\lambda(f) = \mathrm{Eig}_{\overline{\lambda}}(f^{\mathrm{ad}})$.

Beweis. Für $\varphi = f - \lambda\,\mathrm{id}$ und $\varphi^{\mathrm{ad}} = f^{\mathrm{ad}} - \overline{\lambda}\,\mathrm{id}$ gilt

$$\begin{aligned}\varphi^{\mathrm{ad}} \circ \varphi &= f^{\mathrm{ad}} \circ f + \lambda\overline{\lambda}\,\mathrm{id} - \overline{\lambda} f - \lambda f^{\mathrm{ad}} \\ &= f \circ f^{\mathrm{ad}} + \lambda\overline{\lambda}\,\mathrm{id} - \lambda f^{\mathrm{ad}} - \overline{\lambda} f = \varphi^{\mathrm{ad}} \circ \varphi.\end{aligned}$$

Also ist φ normal. Nach Lemma 11.3.3 ist

$$\mathrm{Eig}_\lambda(f) = \ker(\varphi) = \ker(\varphi^{\mathrm{ad}}) = \mathrm{Eig}_{\overline{\lambda}}(f^{\mathrm{ad}}).$$

□

Lemma 11.3.5. *Sei $f : V \to V$ normal und $U = \mathrm{Eig}_\lambda(f)$ ein Eigenraum. Dann ist $U^\perp$ invariant.*

Beweis. Für ein $u \in U$, $x \in U^\perp$ gilt:

$$\sigma(fx, u) = \sigma(x, f^{\mathrm{ad}}u) = \sigma(x, \overline{\lambda}u) = \lambda\sigma(x, u) = 0.$$

Also ist $f(x) \in U^\perp$. □

Satz 11.3.6. *Für ein $A \in \mathrm{Mat}_{n,n}(\mathbb{C})$ sind äquivalent:*

(i) *A ist unitär-diagonalisierbar.*

(ii) *Es gibt eine Orthonormalbasis aus Eigenvektoren von A.*

(iii) *A ist normal.*

11.4 Kegelschnitte

Es sei hier $\mathbb{K}$ ein Körper mit $\mathrm{char}(\mathbb{K}) \neq 2$. Eine symmetrische Bilinearform σ auf einem $\mathbb{K}$-Vektorraum V bestimmt eine *quadratische Form*

$$q = q_\sigma : V \longrightarrow \mathbb{K}, \qquad q(x) = \sigma(x, x),$$

d.h. eine Funktion mit den Eigenschaften

(Q1) $q(\lambda x) = \lambda^2 q(x)$,

(Q2) $\frac{1}{2}(q(x+y) - q(x) - q(y))$ ist eine symmetrische Bilinearform.

Beispiel 11.4.1. 1) $V = \mathbb{K}^n$, $S \in \mathrm{Mat}_{n,n}(\mathbb{K})$ symmetrisch: $q(x) = x^\top Sx$.

2) konkret für $S_p = \left(\begin{array}{c|c} \mathbb{1}_p & 0 \\ \hline 0 & -\mathbb{1}_q \end{array}\right)$, $p + q = n$,

$$q_{S_p}(x) = \sigma_p(x, x) = \sum_{i=1}^{p} x_i^2 - \sum_{i=p+1}^{n} x_i^2,$$

Signatur $p - q = 2p - n$; vgl. Beispiel 10.2.5(2).

Wir betrachten nun etwas allgemeiner Funktionen $F : \mathbb{K}^n \to \mathbb{K}$ der folgenden Form

$$F(x) = F(x_1, \ldots, x_n) = \underbrace{\sum_{i,j} a_{ij}x_ix_j}_{\text{quadratischer Anteil}} + \underbrace{\sum_{k=1}^{n} b_kx_k}_{\text{linearer Anteil}} + \underbrace{c}_{\text{konstanter Anteil}}$$

mit $a_{ij} \in \mathbb{K}$, $b_k \in \mathbb{K}$, $c \in \mathbb{K}$; es ist $F \in \mathbb{K}[x_1, \ldots, x_n]$ ein allgemeines Polynom in n Variablen vom Grad ≤ 2. Wir schreiben kompakter

$$F(x) = x^\top Ax + bx + c$$

mit $A = (a_{ij}) \in \mathrm{Mat}_{n,n}(\mathbb{K})$, $b = (b_1, \ldots, b_n)^\top \in \mathbb{K}^n$, $c \in \mathbb{K}$.

- Ohne F zu ändern, können wir A durch die symmetrische Matrix $\frac{1}{2}(A + A^\top)$ ersetzen, denn für jedes Paar (i,j) haben wir
$$a_{ij}x_ix_j + a_{ji}x_jx_i = \frac{a_{ij}+a_{ji}}{2}x_ix_j + \frac{a_{ji}+a_{ij}}{2}x_jx_i.$$
Also nehmen wir A als symmetrisch an, und haben q_A als quadratischen Anteil.
- Der lineare Anteil $x \mapsto bx$ ist die zu b gehörende Linearform, also $\Psi_\sigma(b) \in (\mathbb{K}^n)^*$ für $\sigma = \langle -,- \rangle$ das Standardskalarprodukt auf $\mathbb{K}^n$.

Beispiel 11.4.2. aus der Physik:

1) Trägheitstensor
2) Dieelektrizitätszahlen (Ohmsches Gesetz)
3) Elastizitätsmoduln (Hooksches Gesetz)

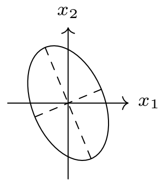
Definition 11.4.3. Die Niveaumenge
$$Q(F) = \{x \in \mathbb{K}^n \mid F(x) = 0\}$$
heißt *Quadrik* oder *Hyperfläche zweiten Grades* oder *Kegelschnitt.*

Bemerkung 11.4.4. Ziel ist eine Klassifikation aller Funktionen F im Fall $\mathbb{K} = \mathbb{R}$. Durch

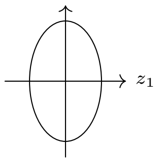
- eine Drehung des Koordinatensystems (Hauptachsentransformation $x \mapsto y = Wx$ $(W \in \mathrm{SO}(n))$)
- eine Translation des Koordinatensystems $y \mapsto z = y + P$

soll F in eine Normalform gebracht werden, in welcher die Gestalt (geometrisch, analytisch, ...) von $Q(F)$ leicht ablesbar ist.

Beispiel 11.4.5.
$$A = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}, \qquad b = 0, \qquad c = -1.$$
$$F(x_1,x_2) = 4x_1^2 + 2x_1x_2 + 2x_2^2 - 1$$
$$Q(F) = \{x \in \mathbb{R}^2 \mid F(x) = 0\} \quad \text{Ellipse}$$



$$\lambda_1 = 3 + \sqrt{2} > 0$$
$$\lambda_2 = 3 - \sqrt{2} > 0$$
$$W = \begin{pmatrix} c & -s \\ s & c \end{pmatrix}, \qquad c = \frac{1+\sqrt{2}}{\sqrt{(1+\sqrt{2})^2+1}}, \qquad s = \frac{1}{\sqrt{(1+\sqrt{2})^2+1}}.$$
$$W^\top AW = \begin{pmatrix} 3+\sqrt{2} & 0 \\ 0 & 3-\sqrt{2} \end{pmatrix} \text{ diagonal}$$



$$Q(F') = \left\{ z \in \mathbb{R}^2 \mid \frac{z_1^2}{\alpha_1^2} + \frac{z_2^2}{\alpha_2^2} = 1 \right\}$$
$$\alpha_1 = \sqrt{\frac{1}{3+\sqrt{2}}}, \qquad \alpha_2 = \sqrt{\frac{1}{3-\sqrt{2}}}$$

Satz 11.4.6 (Klassifikation der Kegelschnitte)**.** *Zu jedem $F : \mathbb{R}^n \to \mathbb{R}$, $F(x) = x^\top Ax + bx + c$ mit $A^\top = A$, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$ und $A \neq 0$ gibt es*

- $W \in \mathrm{SO}(n)$

- $P \in \mathbb{R}^n$
- $0 < k \leq n$,

so dass für die neuen Variablen $y = Wx$ zund $z = y + P = Wx + P$ eine der drei folgenden Alternativen gilt:

(I) $F(x) = \sum_{i=1}^k \lambda_i z_i^2$,

(II) $F(x) = \sum_{i=1}^k \lambda_i z_i^2 + \delta$,

(III) $F(x) = \sum_{i=1}^k \lambda_i z_i^2 + \beta z_{k+1}$,

wobei die $\lambda_1, \ldots, \lambda_k$ Eigenwerte von A sind und $\delta, \beta \neq 0$.

Bemerkung 11.4.7. Die transformierte Funktion F' ist

(I) $F'(z) = z^\top Dz$,

(II) $F'(z) = z^\top Dz + \delta$,

(III) $F'(z) = z^\top Dz + \beta z_{k+1}$.

Es gibt eine Bijektion zwischen den Niveaumengen

$$Q(F') \xrightarrow{\cong} Q(F), \qquad z \longmapsto x = W^{-1}(z - P).$$

Beweis. (1) A werde nach dem Hauptsatz über symmetrische Matrizen orthogonal diagonalisiert durch $W \in \mathrm{O}(n)$, also

$$WAW^\top = D = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n).$$

Für $x = W^{-1}y$ ergibt sich

$$\begin{aligned} F(x) = F(W^{-1}y) &= (W^{-1}y)^\top A(W^{-1}y) + b^\top W^{-1}y + c \\ &= y^\top (W^\top)^{-1} AW^{-1}y + b^\top W^\top y + c \\ &= y^\top WAW^\top y + (Wb)^\top y + c \\ &= y^\top Dy + (Wb)^\top y + c \\ &= \sum_{i=1}^n \lambda_i y_i^2 + \sum_{i=1}^n b_i' y_i + c \end{aligned}$$

mit $b' = Wb$.

(2) *Fall: A invertierbar.* Es ist also $\lambda_1, \ldots, \lambda_n \neq 0$; wir setzen $z_i = y_i + \frac{b_i'}{2\lambda_i}$ und erhalten durch quadratische Ergänzung

$$\begin{aligned} F(x) &= \sum_i \lambda_i (z_i - \frac{b_i'}{2\lambda_i})^2 + \sum_i b_i'(z_i - \frac{b_i'}{2\lambda_i}) + c \\ &= \sum_i \left(\lambda_i z_i^2 - 2\lambda_i \frac{b_i'}{2\lambda_i} z_i + \lambda_i \frac{(b_i')^2}{4\lambda_i^2} \right) + \sum_i \left(b_i z_i - \frac{(b_i')^2}{2\lambda_i} \right) + c \\ &= \sum_i \lambda_i z_i^2 \underbrace{- \frac{1}{2} \sum_i \frac{(b_i')^2}{\lambda_i} + c}_{\delta} \\ &= z^\top Dz + \delta. \end{aligned}$$

Dies ist Form (I) für $\delta = 0$, und Form (II) für $\delta \neq 0$.

(3) *Fall: A nicht-invertierbar.* Wir können annehmen (bis auf eine Variablenvertauschung, d.h. eine orthogonale Permutationsmatrix), dass $\lambda_1, \dots, \lambda_k \neq 0$, $\lambda_{k+1} = \dots = \lambda_n = 0$, also $\mathrm{rg}(A) = k$ für ein $k = 0, 1, \dots, n-1$.
Wir substituieren wie oben $z_i = y_i + \frac{b_i'}{2\lambda_i}$ $(i = 1, \dots, k)$ und erhalten:

$$\begin{aligned}
F(x) &= \sum_{i=1}^{k} \lambda_i y_i^2 + \sum_{i=1}^{k} b_i' y_i + \sum_{i=k+1}^{n} b' i y_i + c \\
&= \sum_{i=1}^{k} \lambda_i \left(z_i - \frac{b_i'}{2\lambda_i} \right)^2 + \sum_{i=1}^{k} b_i' \left(z_i - \frac{b_i'}{2\lambda_i} \right) + \sum_{i=k+1}^{n} b_i' \left(z_i - \frac{b_i'}{2\lambda_i} \right) + c \\
&= \sum_{i=1}^{k} \left(\lambda_i z_i^2 - 2\lambda_i \frac{b_i'}{2\lambda_i} z_i + \lambda_i \frac{(b_i')^2}{4\lambda_i^2} \right) + \sum_{i=1}^{k} \left(b_i' z_i - \frac{(b_i')^2}{2\lambda_i} \right) + \sum_{i=k+1}^{n} \left(b_i' z_i - \frac{(b_i')^2}{2\lambda_i} \right) + c \\
&= \sum_{i=1}^{k} \lambda_i z_i^2 + \sum_{i=k+1}^{n} b_i' z_i \underbrace{- \frac{1}{2} \sum_{i=1}^{n} \frac{(b_i')^2}{\lambda_i} + c}_{\delta}
\end{aligned}$$

Sind alle $b_i' = 0$, so ist dies Form (I). Falls nicht, so sei l minimal mit $b_l' \neq 0$, $k < l \leq n$.

Falls nicht, so sei l minimal mit $b_l' \neq 0$, $k < l \leq n$. Nach einer Variablenvertauschung, d.h. einer orthogonalen Permutationsmatrix) können wir $l = k + 1$ annehmen.
Die Substitution

$$u_i = z_i \quad (i \neq k+1), \qquad u_{k+1} = z_{k+1} + \frac{1}{b_{k+1}'} \sum_{j=k+2}^{n} b_j' z_j$$

eliminiert die Variablen $z_{k+2}, \dots, z_n$:

$$\begin{aligned}
F(x) &= \sum_{i=1}^{k} \lambda_i z_i^2 + b_{k+1}' z_{k+1} + \sum_{j=k+2}^{n} b_j' z_j + \delta \\
&= \sum_{i=1}^{k} \lambda_i u_i^2 + b_{k+1}' u_{k+1} - \frac{b_{k+1}'}{b_{k+1}'} \sum_{j=k+2}^{n} b_j' z_j + \sum_{j=k+2}^{n} b_j' z_j + \delta \\
&= \sum_{i=1}^{k} \lambda_i u_i^2 + b_{k+1}' u_{k+1} + \delta
\end{aligned}$$

Ist $\delta = 0$, so ist dies Fall (III). Anderenfalls eliminiert eine letzte Substitution

$$v_i = u_i \quad (i \neq k+1), \qquad v_{k+1} = u_{k+1} - \frac{\delta}{b_{k+1}'}$$

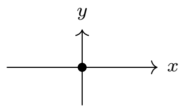
den konstanten Term und wir haben Fall (III).

Die Substitutionen sind alle orthogonal oder affin, bis auf $z \mapsto u$; hier überlege man, dass die ersten $z_1, \dots, z_k$ festbleiben und man im Koordinatenraum $z_1 = \dots = z_k = 0$ des $\mathbb{R}^n$ eine orthonormale Teilbasis wählen kann, welche mit dem Vektor $a$mit den z-Koordinaten $(0, \dots, 0, 1, \frac{b_{k+2}'}{b_{k+1}'}, \dots, \frac{b_n'}{b_{k+1}'}$ beginnt.
W ergibt sich aus der Komposition aller orthogonalen Substitutionen, P aus den affinen. Falls nötig, schalten wir die Spiegelung $S = \mathrm{Diag}(-1, +1, \dots, +1)$ nach, damit $SW \in \mathrm{SO}(n)$ gilt. □

Klassifikation für $n = 2$

(I.1) $k = 2, \lambda_1, \lambda_2 > 0, \lambda_1 x_1^2 + \lambda_2 x_2^2 = 0$

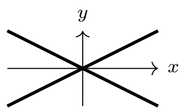
$$\boxed{\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = 0}$$



Punkt

(I.2) $k = 2, \lambda_1 > 0, \lambda_2 < 0$

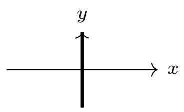
$$\boxed{\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2} = 0}$$



Geradenkreuz

(I.3) $k = 1, \lambda_1 > 0, \lambda_2 = 0$

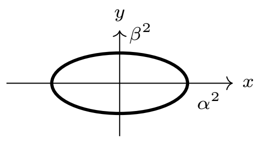
$$\boxed{\frac{x^2}{\alpha^2} = 0}$$



Gerade

(II.1) $k = 2, \lambda_1, \lambda_2 > 0, \delta = -1$

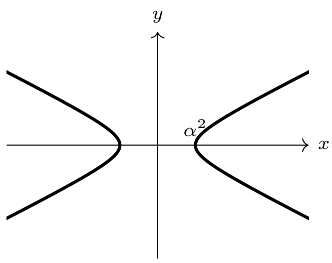
$$\boxed{\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = 1}$$



Ellipse

(II.2) $k = 2, \lambda_1 > 0, \lambda_2 < 0, \delta = -1$

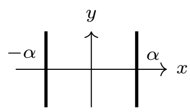
$$\boxed{\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2} = 1}$$



Hyperbel

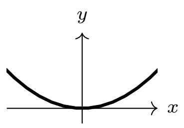
(II.3) $k = 1, \lambda_1 > 0, \lambda_2 = 0, \delta = -1$

$$\boxed{\frac{x^2}{\alpha^2} = 1}$$



(III.1) $k = 1, \lambda_1 > 0, \lambda_2 = 0, \beta = -1$

$$\boxed{\frac{x^2}{\alpha^2} - y = 0}$$



Parabel

In allen anderen Fällen ist $Q = \emptyset$.

Klassifikation für $n = 3$

	λ_1	λ_2	λ_3	δ	β	Achsenabschnittsform	$Q(F)$
$k = 3$							
I	$+$	$+$	$+$	$-$	0	$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} + \frac{z^2}{\gamma^2} - 1 = 0$	Ellipsoid
II	$+$	$+$	$-$	0	0	$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} - \frac{z^2}{\gamma^2} = 0$	elliptischer Doppelkegel
II	$+$	$+$	$-$	$-$	0	$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} - \frac{z^2}{\gamma^2} - 1 = 0$	einschaliges Hyperboloid
$k = 2$							
III	$+$	$+$	0	0	$-$	$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} - \frac{z}{\gamma^2} = 0$	elliptisches Paraboloid
III	$+$	$-$	0	0	$-$	$\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2} - \frac{z}{\gamma^2} = 0$	hyperbolisches Paraboloid (Sattelfläche
$k = 2$							
II	$+$	$+$	0	$-$	0	$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} - 1 = 0$	elliptischer Zylinder
$k = 1$							
III	$+$	0	0	0	$-$	$\frac{x^2}{\alpha^2} - y = 0$	parabolischer Zylinder
$k = 2$							
II	$+$	$-$	0	$-$	0	$\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2} - 1 = 0$	hyperbolischer Zylinder

Alle restlichen Fälle sind ausgeartet oder Q ist leer.

Stichwortverzeichnis