Mathematik II für Physiker*

Sommersemester 2014

Stefan Teufel
Mathematisches Institut
Uni Tübingen

30. Juni 2014

* Diese vorläufige Version des Skriptums ist nur zum Gebrauch parallel zum Besuch der Vorlesung gedacht. Das Studium des Skripts kann den Besuch der Vorlesung **nicht** ersetzen! Falls Sie Fehler finden, teilen Sie mir diese (auch die offensichtlichen) bitte mit!

Inhaltsverzeichnis

1 Vektorräume

1.1 Beispiel. **Die Menge der n-Tupel reeller Zahlen**

$$\mathbb{R}^n := \{(x_1, x_2, \dots, x_n) \,|\, x_j \in \mathbb{R}\}$$

Definiert man die **Summe** zweier Elemente aus $\mathbb{R}^n$ durch

$$(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) := (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)\,,$$

so erhält man wieder ein Element aus $\mathbb{R}^n$. Ebenso erklärt man für $\alpha \in \mathbb{R}$ und $(x_1, \dots, x_n) \in \mathbb{R}^n$ die **Multiplikation**

$$\alpha\,(x_1, \dots, x_n) := (\alpha x_1, \alpha x_2, \dots, \alpha x_n) \in \mathbb{R}^n$$

komponentenweise.

Die Rechenregeln für reelle Zahlen implizieren sofort die entsprechenden Regeln für die Rechenoperationen im $\mathbb{R}^n$. Für die **Addition** gilt:

(1)	$(x+y)+z = x+(y+z)$	$\forall x, y, z \in \mathbb{R}^n$	(Assoziativgesetz)
(2)	$x+y = y+x$	$\forall x, y \in \mathbb{R}^n$	(Kommutativgesetz)
(3)	$x+0 = x$	$\forall x \in \mathbb{R}^n$ und $0 := (0, \dots, 0)$	(Existenz der Null)
(4)	$x+(-x) = 0$	$\forall x \in \mathbb{R}^n$ und $-x := (-x_1, \dots, -x_n)$	(Existenz des Inversen)

Für die **Muliplikation** mit reellen Zahlen gilt:

(5)	$\alpha(\beta x) = (\alpha\beta)x$	$\forall x \in \mathbb{R}^n, \alpha, \beta \in \mathbb{R}$
(6)	$1x = x$	$\forall x \in \mathbb{R}^n$

Weiterhin sind Addition und Multiplikation "verträglich":

(7)	$\alpha(x+y) = \alpha x + \alpha y$	$\forall x, y \in \mathbb{R}^n, \alpha \in \mathbb{R}$	(Distributivgesetz 1)
(8)	$(\alpha+\beta)x = \alpha x + \beta x$	$\forall x \in \mathbb{R}^n, \alpha, \beta \in \mathbb{R}$	(Distributivgesetz 2)

1.2 Beispiel. **Die Menge der komplexwertigen stetigen Funktionen auf $[0,1]$**

$$C([0,1]) := \{f : [0,1] \to \mathbb{C} \,|\, f \text{ ist stetig}\}$$

Definiert man die **Summe** zweier Funktionen $f, g \in C([0,1])$ durch

$$(f+g)(x) := f(x) + g(x) \qquad \forall x \in [0,1]$$

so erhält man wieder ein Element aus $C([0,1])$. Ebenso erklärt man für $\alpha \in \mathbb{C}$ und $f \in C([0,1])$ die **Multiplikation**

$$(\alpha f)(x) := \alpha f(x) \qquad \forall x \in [0,1]$$

punktweise. Auch für $C([0,1])$ gelten die acht Rechenregeln, die wir für $\mathbb{R}^n$ gefunden haben. Es bezeichne 0 die Funktion $0(x) = 0$ für alle $x \in [0,1]$ und für $f \in C([0,1])$ sei $-f := (-1)f$. Dann gilt für $f, g, h \in C([0,1])$ und $\alpha, \beta \in \mathbb{C}$:

(1)	$(f+g)+h = f+(g+h)$	(Assoziativgesetz)
(2)	$f+g = g+f$	(Kommutativgesetz)
(3)	$f+0 = f$	(Existenz der Null)
(4)	$f+(-f) = 0$	(Existenz des Inversen)
(5)	$\alpha(\beta f) = (\alpha\beta)f$	
(6)	$1f = f$	
(7)	$\alpha(f+g) = \alpha f + \alpha g$	(Distributivgesetz 1)
(8)	$(\alpha+\beta)f = \alpha f + \beta f$	(Distributivgesetz 2)

Zumindest bezüglich dieser acht Regeln verhalten sich also Funktionen wie n-Tupel von Zahlen. Wir hätten genausogut differenzierbare oder einfach ganz beliebige Funktionen betrachten können und wären zum gleichen Ergebnis gekommen.

1.3 Bemerkung. In der Mathematik gibt es viele sehr verschiedene Mengen deren Elemente man in natürlicher Weise addieren bzw. mit Zahlen multiplizieren kann, und zwar so, dass diese acht Regeln erfüllt sind.

Deshalb betrachtet man solche Mengen abstrakt, also unabhängig von der konkreten Form der Elemente und ihren konkreten zusätzlichen Eigenschaften (wie "Länge eines Vektors im $\mathbb{R}^n$" oder "Stetigkeit einer Funktion").

1.4 Notation. Im folgenden bezeichne $\mathbb{K}$ einen Körper, wobei wir immer $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ im Kopf haben sollten.

1.5 Definition. <u>**Vektorraum**</u>

Ein Tripel $(V, +, \cdot)$ bestehend aus einer Menge V, einer Abbildung (genannt Addition)

$$+ : V \times V \to V, \quad (u, v) \mapsto u + v$$

und einer Abbildung (genannt skalare Multiplikation)

$$\cdot : \mathbb{K} \times V \to V, \quad (\alpha, u) \mapsto \alpha u$$

heißt **Vektorraum über** $\mathbb{K}$, wenn die folgenden acht Axiome gelten:

(A1)	$(u+v)+w = u+(v+w)$	$\forall u, v, w \in V$	(Assoziativgesetz)
(A2)	$u+v = v+u$	$\forall u, v \in V$	(Kommutativgesetz)
(A3)	Es gibt ein Element $0 \in V$ mit $u + 0 = u$	$\forall u \in V$	(Existenz der Null)
(A4)	Zu jedem $u \in V$ gibt es ein Element $-u \in V$ mit $u + (-u) = 0$		(Existenz des Inversen)
(A5)	$\alpha(\beta u) = (\alpha\beta)u$	$\forall \alpha, \beta \in \mathbb{K},\ u \in V$	
(A6)	$1u = u$	$\forall u \in V$	
(A7)	$\alpha(u+v) = \alpha u + \alpha v$	$\forall \alpha \in \mathbb{K},\ u, v \in V$	(Distributivgesetz 1)
(A8)	$(\alpha+\beta)u = \alpha u + \beta u$	$\forall \alpha, \beta \in \mathbb{K},\ u \in V$	(Distributivgesetz 2)

Insbesondere ist also $(V, +)$ eine abelsche Gruppe.

1.6 Bemerkung. Wir betrachten die lineare Algebra für diesen axiomatisch definierten Vektorraumbegriff und nicht nur für den $\mathbb{R}^n$ oder den $\mathbb{C}^n$.
Vorteil: Alles was wir ausgehend von den Axiomen zeigen können gilt für eine Vielzahl verschiedenster konkreter Vektorräume.
Nachteil (ein sehr kleiner): Einige Aussagen, die z.B. für den $\mathbb{R}^n$ offensichtlich sind, müssen für abstrakte Vektorräume erst bewiesen werden. Die folgenden Bemerkungen sind Beispiele dafür.

1.7 Bemerkung. **Eindeutigkeit der Null**
In einem Vektorraum gibt es stets nur einen Nullvektor.

Beweis. Seien 0 und $\tilde{0}$ Nullvektoren, so ist

$$0 \stackrel{(A3)}{=} 0 + \tilde{0} \stackrel{(A2)}{=} \tilde{0} + 0 \stackrel{(A3)}{=} \tilde{0}\,.$$

□

1.8 Bemerkung. **Eindeutigkeit des Inversen**
In einem Vektorraum gibt es zu jedem u nur ein inverses Element $-u$.

Beweis. Sei $u + a = 0$ und $u + b = 0$, dann gilt

$$a \stackrel{(A3)}{=} a + 0 \stackrel{(Ann.)}{=} a + (u + b) \stackrel{(A1)}{=} (a + u) + b \stackrel{(A2)}{=} (u + a) + b \stackrel{(Ann.)}{=} 0 + b \stackrel{(A2)}{=} b + 0 \stackrel{(A3)}{=} b$$

□

1.9 Bemerkung. Auch viele weitere aus dem $\mathbb{R}^n$ bekannte Rechenregeln folgen aus den Vektorraumaxiomen, z.B.

$$0 \cdot u = 0, \quad (-1)u = -u, \quad \alpha \cdot 0 = 0 \qquad \forall u \in V, \alpha \in \mathbb{K}\,.$$

Beweis. Übungen. □

1.10 Bemerkung. Keines der Axiome ist überflüssig. Die Forderung $1 \cdot u = u$ stellt z.B. die Verträglichkeit von Addition in V und Addition in $\mathbb{K}$ sicher:

$$u + u = 1 \cdot u + 1 \cdot u = (1 + 1) \cdot u = 2 \cdot u\,.$$

Definiert man in $\mathbb{R}^2$ die skalare Multiplikation durch

$$\alpha \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ 0 \end{pmatrix},$$

so sind alle Axiome außer $1 \cdot x = x$ erfüllt.

1.11 Notation. (a) Im folgenden schreiben wir statt $u + (-v)$ einfach $u - v$.

(b) Einige Doppeldeutigkeiten bei Bezeichnungen sind ohne unhandliche Notation nicht zu vermeiden, z.B. werden das Nullelement $0 \in \mathbb{K}$ des Körpers und der Nullvektor $0 \in V$ mit dem selben Symbol 0 bezeichnet. Auch schreiben wir V für den Vektorraum und meinen eigentlich das Tripel $(V, +, \cdot)$.

Falls eine Teilmenge $U \subset V$ eines $\mathbb{K}$-Vektorraums V selbst wieder ein $\mathbb{K}$-Vektorraum ist, so spricht man von einem Untervektorraum (oder kurz Unterraum oder Teilraum).

1.12 Definition. **Untervektorraum**
Sei V ein Vektorraum über $\mathbb{K}$. Eine nichtleere Teilmenge $U \subset V$ heißt **Untervektorraum von** V, wenn für alle $u, v \in U$ und alle $\alpha \in \mathbb{K}$ gilt :

$$u + v \in U \quad \text{und} \quad \alpha u \in U\,.$$

1.13 Bemerkung. Ein Untervektorraum $U \subset V$ enthält immer den Nullvektor und mit jedem $u \in U$ auch den Vektor $-u$. Insbesondere ist also ein Untervektorraum selbst wieder ein Vektorraum!

Beweis. Das folgt sofort aus $0 \cdot u = 0$ und $(-1) \cdot u = -u$, was wiederum gemäß Bemerkung 1.9 aus den Axiomen folgt. □

1.14 Beispiele. (a) $\{0\}$ und V sind Unterräume von V.

(b) Sind U_1 und U_2 Unterräume von V, so ist auch $U_1 \cap U_2$ ein Unterraum von V.

(c) Im $\mathbb{R}^3$ sind die Unterräume außer $\{0\}$ und $\mathbb{R}^3$ alle Geraden und Ebenen durch den Ursprung.

1.15 Warnung. Was die algebraischen Manipulationen betrifft, kann man in abstrakten Vektorräumen rechnen wie im $\mathbb{R}^n$ und die Intuition aus dem "Anschauungsraum" $\mathbb{R}^3$ wird im folgenden auch immer wieder das Verständnis erleichtern. Aber man darf nicht vergessen, dass Konzepte wie "Länge" eines Vektors bzw. der "Winkel" zwischen zwei Vektoren in einem abstrakten Vektorraum zunächst nicht vorkommen.

1.16 Definition. **Linearkombination und lineare Hülle**

Sei V ein $\mathbb{K}$-Vektorraum und seien $v_1, \dots, v_n \in V$. Jeder Vektor der Form

$$\alpha_1 v_1 + \ldots + \alpha_n v_n = \sum_{j=1}^{n} \alpha_j v_j \quad \text{mit} \quad \alpha_1, \dots, \alpha_n \in \mathbb{K}$$

heißt **Linearkombination** der Vektoren $v_1, \dots, v_n$.

Die Menge aller Linearkombinationen aus $v_1, \dots, v_n$ heißt ihre **lineare Hülle** oder ihr **Aufspann**, bezeichnet mit

$$\mathrm{Span}\{v_1, \dots, v_n\}\,.$$

1.17 Bemerkung. $\mathrm{Span}\{v_1, \dots, v_n\}$ ist ein Teilraum von V.

Beweis. Übungsaufgabe. □

1.18 Definition. **Erzeugendensystem**

Die Vektoren $v_1, \dots, v_n \in V$ heißen **Erzeugendensystem** für einen Teilraum $U \subset V$, falls

$$U = \mathrm{Span}\{v_1, \dots, v_n\}$$

1.19 Beispiele.

- Für jeden Vektor $v \neq 0$ des $\mathbb{R}^3$ ist

$$\mathrm{Span}\{v\} = \{tv \,|\, t \in \mathbb{R}\}$$

die Ursprungsgerade mit Richtungsvektor v.

- Sind u, $v \in \mathbb{R}^3$ nicht parallel, so ist

$$\mathrm{Span}\{u, v\} = \{su + tv \,|\, s, t \in \mathbb{R}\}$$

die von u und v aufgespannte Ursprungsebene.

- Sei $p_k(x) = x^k$, $k \in \mathbb{N}$, so ist

$$\mathrm{Span}\{p_0, p_1, \dots, p_n\}$$

der Vektorraum der Polynome vom Grad $\leq n$.

1.20 Definition. Sei $M \subset V$ eine nichtleere Teilmenge, dann bezeichnet $\mathrm{Span}M$ die Menge aller Linearkombinationen aus je **endlich vielen** Vektoren aus M.

1.21 Beispiel. Der Vektorraum $P_{\mathbb{K}}$ aller Polynome mit Koeffizienten in $\mathbb{K}$ ist

$$\mathrm{Span}\{p_0, p_1, p_2, \dots\}$$

1.22 Definition. Lineare Unabhängigkeit

Es heißen n Vektoren $v_1, \ldots, v_n$ eines $\mathbb{K}$-Vektorraumes V **linear abhängig**, wenn es Skalare $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ gibt, die nicht alle Null sind, so dass

$$\alpha_1 v_1 + \ldots + \alpha_n v_n = 0\,.$$

Andernfalls heißen $v_1, \ldots, v_n$ **linear unabhängig**.

Merke: $v_1, \ldots, v_n \in V$ sind linear unabhängig, genau dann wenn

$$\alpha_1 v_1 + \ldots + \alpha_n v_n = 0 \quad \Rightarrow \quad \alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$$

1.23 Bemerkung. Ein einzelner Vektor v heißt also linear unabhängig, falls $v \neq 0$ ist.

1.24 Satz. Für $n \geq 2$ sind die Vektoren $v_1, \ldots, v_n$ genau dann linear abhängig, wenn wenigstens einer von ihnen eine Linearkombination der übrigen ist.

Beweis. „$\Rightarrow$“: (d.h.linear abhängig $\Rightarrow$ Linearkombination).

Seien $v_1, \ldots, v_n$ linear abhängig, dann gibt es $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ mit $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$ und $\alpha_m \neq 0$ für ein $m \in \{1, \ldots, n\}$. Dann kann man aber nach v_m auflösen,

$$v_m = \sum_{k \neq m} \left(-\frac{\alpha_k}{\alpha_m} \right) v_k\,,$$

also ist v_m eine Linearkombination der anderen Vektoren.

„$\Leftarrow$“: Ist einer der Vektoren Linearkombination der anderen, etwa $v_1 = \alpha_2 v_2 + \ldots + \alpha_n v_n$, so folgt $(-1)v_1 + \alpha_2 v_2 + \ldots \alpha_n v_n = 0$. Also sind $v_1, \ldots, v_n$ linear abhängig, da $\alpha_1 = -1 \neq 0$. □

1.25 Definition. Eine beliebige nichtleere Teilmenge M von V heißt linear unabhängig, wenn je endlich viele Vektoren aus M linear unabhängig sind.

1.26 Beispiele. (a) Im $\mathbb{K}^n$ sind die kanonischen Basisvektoren

$$e_1 = (1, 0, \ldots, 0) =: \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

linear unabhängig. Denn aus $\alpha_1\, e_1 + \ldots + \alpha_n\, e_n = 0$ folgt $\alpha_1 = \ldots = \alpha_n = 0$.

(b) Die Monome $p_k(x) = x^k, k = 0, \ldots, n$ sind linear unabhängig für jedes $n \in \mathbb{N}$. Denn

$$\begin{aligned} \alpha_0\, p_0 + \ldots + \alpha_n\, p_n = 0 \quad &\Rightarrow \quad \alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n = 0 \quad \forall x \in \mathbb{R} \\ &\Rightarrow \quad \alpha_0 = \alpha_1 = \ldots = \alpha_n = 0\,. \end{aligned}$$

(c) Die Menge $M = \{p_0, p_1, \ldots\}$ ist eine linear unabhängige Menge im Vektorraum $P_{\mathbb{K}}$ der Polynome. Denn jede endliche Teilmenge $T \subset M$ ist in $\{p_0, p_1, \ldots, p_n\}$ enthalten für n hinreichend groß. Und Teilmengen linear unabhäniger Mengen sind linear unabhängig (dies folgt direkt aus Satz 1.24).

1.27 Definition. **Basis eines Vektorraums**
Ein n-Tupel $(v_1, \ldots, v_n)$ von Vektoren des $\mathbb{K}$-Vektorraums V heißt eine **Basis von** V, wenn $v_1, \ldots, v_n$ linear unabhängig sind und

$$\operatorname{Span}\{v_1, \ldots, v_n\} = V$$

gilt.

1.28 Satz. Es ist $(v_1, \ldots, v_n)$ genau dann eine Basis von V, wenn es zu jedem $v \in V$ genau ein $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$ gibt, für das

$$v = \alpha_1 v_1 + \ldots + \alpha_n v_n$$

gilt.

Beweis. „$\Rightarrow$“: Sei $(v_1, \ldots, v_1)$ Basis von V dann gibt es wegen $\operatorname{Span}\{v_1, \ldots, v_n\} = V$ zumindest ein $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$ mit $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$.
Falls auch $(\beta_1, \ldots, \beta_n) \in \mathbb{K}^n \quad v = \beta_1 v_1 + \ldots + \beta_n v_n$ erfüllt, folgt aus

$$0 = v - v = (\alpha_1 - \beta_1)v_1 + \ldots + (\alpha_n - \beta_n)v_n$$

wegen der linearen Unabhängigkeit

$$(\alpha_1 - \beta_1) = \ldots = (\alpha_n - \beta_n) = 0$$

also

$$(\alpha_1, \ldots, \alpha_n) = (\beta_1, \ldots, \beta_n)$$

und somit die Eindeutigkeit.

„$\Leftarrow$“: Die Annahme enthält insbesondere, dass $\operatorname{Span}\{v_1, \ldots, v_n\} = V$. Die Eindeutigkeit der Darstellung in $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$ für $v = 0$ impliziert die lineare Unabhängigkeit von $v_1, \ldots, v_n$. $\square$

1.29 Beispiel. Die **kanonische Basis** des $\mathbb{K}^n$ ist

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots \quad e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

1.30 Bemerkung. In Verallgemeinerung von Definition 1.27 definiert man auch:
Eine Teilmenge $M \subset V$ eines $\mathbb{K}$-Vektorraums V heißt **Basis von** V, falls M linear unabhängig ist und $\operatorname{Span} M = V$ gilt.

1.31 Beispiel. Die Monome $M = \{p_0, p_1, \ldots\}$ bilden eine Basis des Vektorraums $P_{\mathbb{K}}$ der Polynome.

1.32 Bemerkung. Man kann zeigen, dass jeder Vektorraum außer $V = \{0\}$ eine Basis besitzt.

Der Begriff der Dimension eines Vektorraumes beruht auf der folgenden Beobachtung:

1.33 Satz. Besitzt ein Vektorraum V eine Basis aus n Vektoren, so besteht auch jede andere Basis aus n Vektoren.

1.34 Definition. Dimension eines Vektorraums

Besitzt ein Vektorraum V eine Basis aus n Vektoren, so ist seine **Dimension** gleich n und wir schreiben

$$\dim V = n\,.$$

Für $V = \{0\}$ setzt man $\dim V = 0$ und falls V keine endliche Basis besitzt, so setzt man $\dim V = \infty$ und bezeichnet V als **unendlichdimensional**.

Die Aussage des Satzes 1.33 ergibt sich direkt aus folgendem Lemma.

1.35 Lemma. Sei $b_1, \dots, b_n$ ein Erzeugendensystem für V und $a_1, \dots, a_m \in V$ seien linear unabhängig. Dann gilt $m \le n$.

Beweis. Statt

$$a_1, \dots, a_m \in V \text{ linear unabhängig} \quad \Rightarrow \quad m \le n\,,$$

zeigen wir die äquivalente Aussage

$$m > n \quad \Rightarrow \quad a_1, \dots, a_m \in V \text{ linear abhängig}\,,$$

also, dass je m Vektoren $a_1, \dots, a_m$ linear abhängig sind, falls $m > n$ gilt.

Induktion nach n:
$\mathbf{n = 1}$ (Induktionsanfang)
Der Vektor b_1 erzeuge V. Für Vektoren $a_1, \dots, a_m$, $m > 1$, gilt dann $a_j = \alpha_j b_j$, $j = 1, \dots, m$. Falls eins der a_j der Nullvektor ist, so sind $a_1, \dots, a_m$ linear abhängig. Ansonsten ist beispielsweise $\alpha_2 a_1 + (-\alpha_1) a_2 = \alpha_2 \alpha_1 b_1 - \alpha_1 \alpha_2 b_1 = 0$ eine nichttriviale Linearkombination zu Null und wieder folgt die lineare Abhängigkeit.

$\mathbf{n - 1 \Rightarrow n}$: (Induktionsschritt)
Die Behauptung sei für $n-1$ richtig und in $V = \mathrm{Span}\{b_1, \dots, b_n\}$ seien Vektoren $a_1, \dots, a_m$ gegeben $(m > n)$. Dann gibt es Zahlen $\alpha_{jk} \in \mathbb{K}$ mit

$$a_j = \sum_{k=1}^{n} \alpha_{jk} b_k \qquad (j = 1, \dots, m)\,.$$

Seien nicht alle α_{jk} gleich Null, etwa $\alpha_{11} \neq 0$. Ansonsten wären alle $a_j = 0$ und somit $a_1, \dots, a_m$ linear abhängig.
Die $m-1$ Vektoren

$$c_j := a_j - \frac{\alpha_{j1}}{\alpha_{11}}\, a_1 = \sum_{k=1}^{n} \frac{\alpha_{11}\alpha_{jk} - \alpha_{j1}\alpha_{1k}}{\alpha_{11}} b_k \quad (j = 2, \dots, m)$$

liegen in $\mathrm{Span}\{b_2, \dots, b_n\}$, da der Koeffizent von b_1 Null ist. Wegen $m - 1 > n - 1$ sind nach Induktionsvoraussetzung die $c_2, \dots, c_m$ linear abhängig. Also gibt es Zahlen $\lambda_2, \dots, \lambda_m$ die nicht alle Null sind und für die gilt

$$0 = \sum_{j=2}^{m} \lambda_j\, c_j = \sum_{j=2}^{m} \lambda_j (a_j - \frac{\alpha_{j1}}{\alpha_{11}} a_1) = \sum_{j=1}^{m} \lambda_j a_j \quad \text{mit} \quad \lambda_1 := -\frac{1}{\alpha_{11}} \sum_{j=2}^{m} \lambda_j\, \alpha_{j1}$$

Also sind $a_1, \dots, a_m$ linear abhängig. □

1.36 Satz. Basisergänzungssatz

Sei $b_1, \dots, b_n$ ein Erzeugendensystem von V. Falls $a_1, \dots, a_m \in V$ linear unabhängig sind und keine Basis von V bilden, so lassen sich die $a_1, \dots, a_m$ durch Hinzunahme geeigneter b_k zu einer Basis von V ergänzen.

1.37 Satz. **Basisauswahlsatz**
Besitzt der Vektorraum $V \neq \{0\}$ ein endliches Erzeugendensystem $b_1, \ldots, b_n$, so läßt sich aus diesem eine Basis für V auswählen.

Beweis. von Satz 1.36
Sei $A = \{a_1, \ldots, a_m\}$, $M = \{a_1, \ldots, a_m, b_1, \ldots, b_n\}$. Wir betrachten die Menge

$$\mathcal{T} = \{S \subset M \mid A \subset S \quad \text{und} \quad \operatorname{Span} S = V\},$$

also die Menge aller Teilmengen von M, die alle $a_1, \ldots, a_m$ enthalten und ganz V erzeugen. Die Menge $\mathcal{T}$ ist endlich und nicht leer, denn $M \in \mathcal{T}$.

Sei $j = \min\{\,|S|\mid S \in \mathcal{T}\}$, dann gibt es ein $S_0 \in \mathcal{T}$ mit $|S_0| = j$.

Wir zeigen nun, dass S_0 linear unabhängig ist. Daraus folgt dann, dass S_0 eine Basis ist, da $S_0 \in \mathcal{T}$ ja $\operatorname{Span} S_0 = V$ impliziert.

a) Wäre ein $b_k \in S_0$ Linearkombination der restlichen Vektoren in S_0, so wäre S_0 nicht minimal, d.h. es wäre $S_0 \setminus \{b_k\} \in \mathcal{T}$ und $|S_0 \setminus \{b_k\}| < |S_0|$.

b) Wäre ein $a_j \in S_0$ Linearkombination der übrigen Vektoren in S_0, so müßte dabei eines der b_n, sagen wir b_k, einen von Null verschiedenen Vorfaktor haben, da die $a_1, \ldots, a_n$ linear unabhängig sind. Dieses b_k wäre dann eine Linearkombination der anderen Vektoren in S_0 und wir wären wieder bei a). □

Beweis. (von Satz 1.37)
Genau wie der Beweis von Satz 1.36, nur mit $A = \emptyset$. □

1.38 Bemerkung. In einem n-dimensionalen Vektorraum bilden je n linear unabhängig Vektoren eine Basis.

Beweis. Angenommen $(b_1, \ldots, b_n)$ sind linear unabhängig, aber keine Basis von V, dann könnte man sie nach Satz 1.36 zu einer Basis aus mindestens $n+1$ Elementen ergänzen. Das stünde im Widerspruch zu Satz 1.33. □

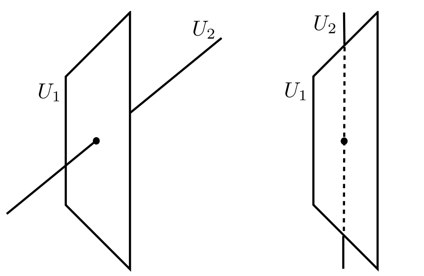
1.39 Bemerkung. Sei $W \subset V$ ein Teilraum, dann gilt $\dim W \leq \dim V$. Falls $\dim W < \infty$, so gilt sogar

$$\dim W = \dim V \quad \Leftrightarrow \quad W = V\,.$$

Beweis. Übungsaufgabe. □

1.40 Bemerkung. **Der Schnitt von Unterräumen**

Seien U_1 und U_2 Unterräume von V, dann ist auch $U_1 \cap U_2$ ein Unterraum von V. Die Dimension von $U_1 \cap U_2$ hängt aber nicht nur von $\dim U_1$ und $\dim U_2$ ab, sondern auch von der relativen Lage.



$\dim(U_1 \cap U_2) = 0$ $\quad$ $\dim(U_1 \cap U_2) = 1$

1.41 Definition. Sind U_1, U_2 Unterräume von V, so heißt

$$U_1 + U_2 := \{u + v \mid u \in U_1, v \in U_2\} \subset V$$

die **Summe** von U_1 und U_2.

Offensichtlich (mittlerweile ?) ist $U_1 + U_2 = \mathrm{Span}(U_1 \cup U_2)$ und somit ist $U_1 + U_2$ wieder ein Unterraum.

1.42 Satz. **Dimensionsformel für Unterräume**
Sind U_1 und U_2 endlichdimensionale Unterräume von V, so gilt

$$\dim(U_1 \cap U_2) + \dim(U_1 + U_2) = \dim U_1 + \dim U_2$$

Beweis. Anschaulich klar! Formal: Basisergänzungssatz.
Sei $(v_1, \ldots, v_r)$ eine Basis von $U_1 \cap U_2$. Wir ergänzen sie einmal zu einer Basis $(v_1, \ldots, v_r, w_1, \ldots, w_s)$ von U_1 und ein zweites Mal zu einer Basis von $(v_1, \ldots, v_r, z_1, \ldots, z_t)$ von U_2. Dann ist

$$(v_1, \ldots, v_r, w_1, \ldots, w_s, z_1, \ldots, z_t)$$

eine Basis von $U_1 + U_2$ (Übungsaufgabe). Damit ist $\dim U_1 \cap U_2 = r$, $\dim U_1 = r + s$, $\dim U_2 = r + t$ und $\dim(U_1 + U_2) = r + s + t$. Also gilt die zu beweisende Formel. $\square$

2 Lineare Abbildungen und Matrizen

2.1 Definition. <u>**Lineare Abbildung**</u>
Seien V und W Vektorräume über $\mathbb{K}$. Eine Abbildung $L : V \to W$ heißt **linear** oder Vektorraum-**Homomorphismus**, wenn

$$L(u_1 + u_2) = L(u_1) + L(u_2) \quad \text{und} \quad L(\alpha\, u_1) = \alpha\, L\,(u_1) \quad \text{für alle } u_1, u_2 \in V,\ \alpha \in \mathbb{K}\,.$$

Die Menge der Homomorphismen von V nach W wird mit $\mathrm{Hom}(V, W)$ oder $\mathcal{L}(V, W)$ bezeichnet.

2.2 Bemerkung. Es ist üblich bei linearen Abbildungen Lu statt $L(u)$ zu schreiben. Bei unendlichdimensionalen Räumen nennt man die linearen Abbildungen auch oft Operatoren. Falls $W = \mathbb{K}$ ist, so spricht man von Linearformen, 1-Formen oder linearen Funktionalen auf V.

2.3 Bemerkung. (a) Definiert man für $f, g \in \mathcal{L}\,(V, W)$ und $\alpha \in \mathbb{K}$

$$(f + g)(v) = f(v) + g(v) \quad \text{und} \quad (\alpha f)(v) = \alpha\, f(v)\,,$$

so wird $\mathcal{L}\,(V, W)$ selbst wieder zu einem Vektoraum über $\mathbb{K}$.

(b) Für zwei lineare Abbildungen zwischen $\mathbb{K}$-Vektorräumen $U \overset{T}{\to} V \overset{S}{\to} W$ ist die Hintereinanderausführung $S \circ T : U \to W$ eine lineare Abbildung. (Übungsaufgabe).

2.4 Beispiele. (a) Die einfachsten linearen Abbildungen sind der **Nulloperator**

$$0 : V \to W,\ u \mapsto 0_W$$

und **die Identität** auf V

$$\mathbb{1} : V \to V,\ u \mapsto u\,.$$

(b) **Der Ableitungsoperator:** Sei $I \subset \mathbb{R}$ ein Intervall, $V = C^1(I)$, $W = C(I)$ dann ist

$$D : V \to W,\ u \mapsto Du = u'$$

linear.

(c) **Das Integral:** Sei $I \subset \mathbb{R}$ ein kompaktes Intervall und $V = C(I)$ dann ist

$$L : V \to \mathbb{K},\ u \mapsto \int_I u(x)\mathrm{d}x$$

eine Linearform auf V.

2.5 Definition. <u>**Bild und Kern einer Abbildung**</u>
Für $f : V \to W$ seien

$$\begin{aligned} \mathrm{Bild}\, f &:= \{f(u) \mid u \in V\} \subset W \\ \mathrm{Kern}\, f &:= \{u \in V \mid f(u) = 0\} \subset V \end{aligned}$$

2.6 Satz. Seien V, W Vektorräume und $L \in \mathcal{L}(V, W)$, dann sind

$$\mathrm{Bild} L \subset W \quad \text{und} \quad \mathrm{Kern} L \subset V \quad \text{Untervektorräume.}$$

Beweis. Wegen $L0_V = 0_W$ ist $0_W \in \text{Bild}L$ und $0_V \in \text{Kern}L$. Also ist weder Bild$L$ noch Kern$L$ die leere Menge.
Sind $w_1 = Lv_1$ und $w_2 = Lv_2$ in Bild$L$, so ist auch $w_1 + w_2 = Lv_1 + Lv_2 = L(v_1 + v_2)$ und $\alpha w_1 = \alpha Lv_1 = L(\alpha v_1)$ in Bild$L$.
Sind $v_1, v_2 \in \text{Kern}L$, so sind auch $v_1 + v_1 \in \text{Kern}L$ und $\alpha v_1 \in \text{Kern}L$, da

$$L(v_1 + v_2) = Lv_1 + Lv_2 = 0 + 0 = 0 \quad \text{und} \quad L(\alpha v_1) = \alpha L(v_1) = \alpha \cdot 0 = 0\,.$$

□

2.7 Beispiel. Sei $a \in \mathbb{R}^3$. Die Abbildung $L : \mathbb{R}^3 \to \mathbb{R}$, $x \mapsto a \cdot x := \sum_{j=1}^3 a_j x_j$ hat Bild$L = \mathbb{R}$ falls $a \neq 0$ und Bild$L = \{0\}$ sonst. Der Kern von L, also $\{x \in \mathbb{R}^3 \,|\, a \cdot x = 0\}$ besteht aus allen Vektoren die senkrecht auf a stehen. Falls $a \neq 0$ ist, so ist das eine Ebene durch den Ursprung, andernfalls ganz $\mathbb{R}^3$.

2.8 Bemerkung. $L \in \mathcal{L}(V, W)$ ist genau dann injektiv, wenn Kern$L = \{0\}$ gilt.

Beweis. Sei Kern$L = \{0\}$. Dann gilt

$$Lv_1 = Lv_2 \;\Rightarrow\; Lv_1 - Lv_2 = 0 \;\Rightarrow\; L(v_1 - v_2) = 0 \;\Rightarrow\; v_1 - v_2 \in \text{Kern}L \;\Rightarrow\; v_1 - v_2 = 0\,,$$

also $v_1 = v_2$ und somit ist L injektiv. Ist andererseits Kern$L \neq \{0\}$, so bildet L neben dem Nullvektor noch weitere Vektoren auf die Null ab und ist somit nicht injektiv. □

2.9 Definition. $L \in \mathcal{L}(V, W)$ heißt

Monomorphismus,	falls L injektiv ist,
Epimorphismus,	falls L surjektiv ist,
Isomorphismus,	falls L bijektiv ist,
Endomorphismus,	falls $V = W$ gilt, und
Automorphismus,	falls L bijektiv ist und $V = W$ gilt.

2.10 Satz. **Dimensionsformel für lineare Abbildungen**
Sei dim$V = n < \infty$ und $L : V \to W$ linear. Dann ist

$$\dim(\text{Kern}\, L) + \underbrace{\dim(\text{Bild}\, L)}_{=:\,\text{Rang}L} = n$$

Beweis. Für $L = 0$ ist Kern$L = V$ und Bild$L = \{0\}$, also stimmt die Formel.

Sei $L \neq 0$ und dim(Kern$L$) $= m$. Wegen $L \neq 0$ ist $m < n$. Im Fall $m > 0$ wählen wir eine Basis $(b_1, \ldots, b_m)$ für Kern$L$ und ergänzen diese zu einer Basis $(b_1, \ldots, b_m, \ldots, b_n)$ für V. Falls $m = 0$ ist, so wählen wir irgendeine Basis $(b_1, \ldots, b_n)$ für V.

Wir zeigen, dass $(Lb_{m+1}, \ldots, Lb_n)$ eine Basis für Bild$L$ ist. Da diese aus $n - m$ Vektoren besteht, ist dann die Dimensionsformel bewiesen.

a) Bild$L = \text{Span}\{Lb_{m+1}, \ldots, Lb_n\}$:
„$\supset$“: $Lb_{m+1}, \ldots, Lb_n \in \text{Bild}L \Rightarrow \text{Span}\{Lb_{m+1} \ldots, Lb_n\} \subset \text{Bild}L$
„$\subset$“: Sei $u \in \text{Bild}L$, also $u = Lv$ für $v = \sum_{j=1}^n \alpha_j b_j \in V$, dann ist $u = Lv = \sum_{j=1}^n \alpha_j Lb_j = \sum_{j=m+1}^n \alpha_j Lb_j \in \text{Span}\{Lb_{m+1} \ldots, Lb_n\}$.

b) $Lb_{m+1}, \ldots, Lb_n$ sind linear unabhängig:
Sei $\sum_{j=m+1}^n \alpha_j Lb_j = 0$. Dann ist $u := \sum_{j=m+1}^n \alpha_j b_j \in \text{Kern}L$, da $Lu = 0$. Falls $m = 0$ folgt

$u = 0$ also $\alpha_1 = \ldots = \alpha_n = 0$.
Im Fall $m > 0$ gibt es Zahlen $\beta_1, \ldots, \beta_m$ mit

$$\sum_{j=m+1}^{n} \alpha_j b_j = u = \sum_{j=1}^{m} \beta_j b_j$$

Wegen der Eindeutigkeit der Basisdarstellung folgt aber $\alpha_{m+1} = \ldots = \alpha_n = 0$.

□

2.11 Satz. Sei $(v_1, \ldots, v_n)$ eine Basis von V. Zu vorgegebenen Vektoren $b_1, \ldots, b_n \in W$ gibt es genau eine lineare Abbildung $L : V \to W$ mit

$$Lv_1 = b_1, \ldots, Lv_n = b_n\,.$$

Insbesondere ist also eine lineare Abbildung $L : V \to W$ durch Kenntnis der Bildvektoren $Lv_1, \ldots, Lv_n$. vollständig beschrieben.

Beweis. Existenz: Seien $b_1, \ldots, b_n$ gegeben. Für $u = \sum_{j=1}^{n} \alpha_j v_j$ definieren wir $Lu := \sum_{j=1}^{n} \alpha_j b_j$. Diese Abbildung ist linear, denn sei $w = \sum_{j=1}^{n} \beta_j v_j$, dann ist

$$L(u + w) = \sum_{j=1}^{n} (\alpha_j + \beta_j) b_j = \sum_{j=1}^{n} \alpha_j b_j + \sum_{j=1}^{n} \beta_j b_j = Lu + Lw.$$

und

$$L(\lambda u) = \sum_{j=1}^{n} \lambda \alpha_j b_j = \lambda Lu.$$

Die Eindeutigkeit ist klar, denn die Linearität von L impliziert ja gerade

$$Lu = L(\sum_{j=1}^{n} \alpha_j v_j) \overset{L\,\text{linear}}{=} \sum_{j=1}^{n} \alpha_j Lv_j = \sum_{j=1}^{n} \alpha_j b_j\,.$$

□

2.12 Korollar. <u>**Basisisomorphismus**</u>
Sei V ein n-dimenaionaler $\mathbb{K}$-Vektorraum. Jede Basis $\mathcal{A} = (v_1, \ldots, v_n)$ von V definiert eine lineare Bijektion, also einen Isomorphismus, $(\,\cdot\,)_{\mathcal{A}} : V \to \mathbb{K}^n$ durch

$$u = \sum_{j=1}^{n} \alpha_j v_j \mapsto (u)_{\mathcal{A}} := (\alpha_1, \ldots, \alpha_n)\,.$$

Also ist jeder n-dimensionale $\mathbb{K}$-Vektorraum isomorph zum $\mathbb{K}^n$.

Beweis. $(\,\cdot\,)_{\mathcal{A}}$ ist die lineare Abbildung festgelegt durch $(v_j)_{\mathcal{A}} = e_j$, wobei $(e_1, \ldots, e_n)$ die kanonische Basis des $\mathbb{K}^n$ aus Beispiel 1.29 ist. Die Bijektivität folgt aus der Eindeutigkeit der Basisdarstellung, Satz 1.28. □

2.13 Definition. <u>**Matrix**</u>
Eine $m \times n$-**Matrix** über $\mathbb{K}$ ist eine Anordnung von $m \cdot n$ Elementen von $\mathbb{K}$ nach folgendem Schema:

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

Die $a_{ij} \in \mathbb{K}$ heißen **Koeffizienten, Komponenten** oder **Einträge** der Matrix. Die waagrecht geschriebenen n-Tupel

$$(a_{i1} \cdots a_{in})$$

heißen **Zeilen** und die senkrecht geschriebenen m-Tupel

$$\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

die Spalten der Matrix. Die Menge aller $m \times n$-Matrizen über $\mathbb{K}$ wird mit $M(m \times n, \mathbb{K})$ bezeichnet.

2.14 Definition. **Matrix einer linearen Abbildung**
Sei $\mathcal{A} = (v_1, \ldots, v_n)$ Basis von V, $\mathcal{B} = (w_1, \ldots, w_m)$ Basis von W und $L : V \to W$ eine lineare Abbildung. Die durch die Basisdarstellung der Bildvektoren

$$Lv_k = \sum_{j=1}^{m} a_{jk} w_j \qquad , \quad k = 1, \ldots, n\,,$$

bestimmten Koeffizienten a_{jk} **bilden die Matrix von L bezgl. der Basen $\mathcal{A}$ und $\mathcal{B}$**,

$$M_{\mathcal{A}}^{\mathcal{B}}(L) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

2.15 Bemerkung. Bei gegebenen Basen $\mathcal{A}$ von V und $\mathcal{B}$ von W, gibt es eine eins-zu-eins Korrespondenz zwischen linearen Abbildungen $L \in \mathcal{L}(V, W)$ und Matrizen $M_{\mathcal{A}}^{\mathcal{B}}(L) \in M(m \times n, \mathbb{K})$. Denn einerseits legt L gemäß Definition 2.14 $M_{\mathcal{A}}^{\mathcal{B}}(L)$ eindeutig fest. Andererseits wird durch eine Matrix

$$M = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

gemäß Satz 2.11 eindeutig die zugehörige lineare Abbildung $L_{\mathcal{A}}^{\mathcal{B}}(M) : V \to W$ durch

$$L_{\mathcal{A}}^{\mathcal{B}}(M) v_k := \sum_{j=1}^{m} a_{jk} w_j$$

definiert.

2.16 Merkregel. Die Spalten der Matrix $M_{\mathcal{A}}^{\mathcal{B}}(L)$ enthalten die Bilder der Basisvektoren aus $\mathcal{A}$ in der Basisdarstellung bzgl. der Basis $\mathcal{B}$:

$$\text{(Basisdarstellung von } Lv_k \text{ bzgl. } \mathcal{B} =) \qquad (Lv_k)_{\mathcal{B}} = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{pmatrix} \qquad (= k\text{-te Spalte von } M_{\mathcal{A}}^{\mathcal{B}}(L)).$$

2.17 Bemerkung. Im Fall $V = W$ verwendet man in der Regel im Bild- und Urbildraum dieselbe Basis $\mathcal{B}$. Statt $M_{\mathcal{B}}^{\mathcal{B}}(L)$ schreiben wir dann $M_{\mathcal{B}}(L)$.

2.18 Beispiele. (a) Die Identität $\mathbb{1}_V : V \to V, u \mapsto u$ besitzt als Matrix $M_{\mathcal{B}}(\mathbb{1}_V)$ bezüglich **jeder** Basis $\mathcal{B}$ von V die **Einheitsmatrix**

$$E_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} = (\delta_{ik})_{n\times n} \quad \text{mit} \quad \delta_{ik} = \begin{cases} 0 & \text{für}\, i \neq k \\ 1 & \text{für}\, i = k \end{cases} .$$

Wählt man verschiedenen Basen $\mathcal{A}$ und $\mathcal{B}$ im Urbild und im Bild, so gilt aber $M_{\mathcal{A}}^{\mathcal{B}}(\mathbb{1}_V) \neq E_n$.

(b) Zur Nullabbildung gehört die **Nullmatrix**

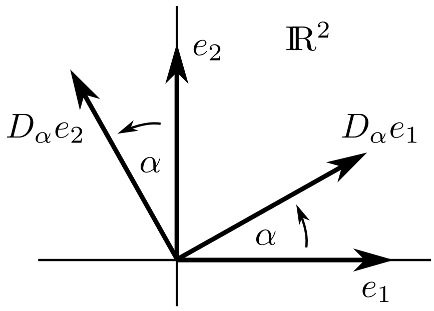
$$0 = \underbrace{\left.\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}\right\}}_{n} m$$

(c) **Drehungen im $\mathbb{R}^2$ bezüglich der kanonischen Basis $\mathcal{K} = (e_1, e_2)$.**
Die Drehung D_α um den Ursprung mit Drehwinkel α ist ein linearer Isomorphismus des $\mathbb{R}^2$. Die zugehörige Matrix ergibt sich aus den Bildern der Basisvektoren:

$$(D_\alpha e_1)_{\mathcal{K}} = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} \quad \text{und} \quad (D_\alpha e_2)_{\mathcal{K}} = \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix},$$

also

$$M_{\mathcal{K}}(D_\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$



Rechnen mit Matrizen

Ziel: Das Rechnen mit Vektoren und linearen Abbildungen zwischen Vektorräumen soll auf das Rechnen mit n-Tupeln und Matrizen reduziert werden.

2.19 Satz. Matrix-Vektor-Multiplikation
Sei $\mathcal{A} = (v_1, \ldots, v_n)$ Basis von V, $\mathcal{B} = (w_1, \ldots, w_m)$ Basis von W und $L : V \to W$ eine lineare Abbildung. Sei $M_{\mathcal{A}}^{\mathcal{B}}(L) = A$ die zugehörige Matrix und für $u \in V$ sei die Koordinatendarstellung

$$(u)_{\mathcal{A}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x \in \mathbb{K}^n .$$

Dann gelten für die Koordinatendarstellung des Bildvektors

$$(Lu)_{\mathcal{B}} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = y \in \mathbb{K}^m$$

die Gleichungen

$$y_j = \sum_{k=1}^{n} a_{jk}\, x_k \qquad (j = 1, \ldots, m).$$

Diese Gleichungen schreiben wir auch in Kurzform als

$$y = A\,x \qquad \text{(oder für später auch} \quad (Lu)_{\mathcal{B}} = M_{\mathcal{A}}^{\mathcal{B}}(L)(u)_{\mathcal{A}})$$

oder in Langform

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Die so definierte Matrix-Vektor-Multiplikation kann man sich wie folgt merken:
Zur Berechnung von $y_j = (Ax)_j = a_{j1}x_1 + \cdots + a_{jn}x_n$ denke man sich die j-te Zeile der Matrix über den Vektor x gelegt und bilde die Summe der Produkte übereinanderstehender Koeffizienten.

Beweis.

$$Lu = L\sum_{k=1}^{n} x_k v_k = \sum_{k=1}^{n} x_k L v_k = \sum_{k=1}^{n} x_k \sum_{j=1}^{m} a_{jk} w_j = \sum_{j=1}^{m} \underbrace{\left(\sum_{k=1}^{n} x_k a_{jk}\right)}_{y_j} w_j\,.$$

□

2.20 Satz. **Addition und skalare Multiplikation von Matrizen**

a) Haben die linearen Abbildungen $S, T : V \to W$ bezüglich fest gewählter Basen $\mathcal{A}$ für V und $\mathcal{B}$ für W die Matrixdarstellungen

$$M_{\mathcal{A}}^{\mathcal{B}}(S) = A = (a_{jk})\,, \quad M_{\mathcal{A}}^{\mathcal{B}}(T) = B = (b_{jk})\,,$$

so ist die Matrix von $S + T$ gegeben durch

$$M_{\mathcal{A}}^{\mathcal{B}}(S + T) = A + B := (a_{jk} + b_{jk})\,.$$

(koeffizientenweise Addition)

b) Die Matrix von λS mit $\lambda \in \mathbb{K}$ ist

$$M_{\mathcal{A}}^{\mathcal{B}}(\lambda S) = \lambda A := (\lambda a_{jk})\,.$$

(koeffizientenweise Multiplikation mit einem Skalar)

Die Matrix von $\lambda S + \mu T$ ist also $(\lambda a_{jk} + \mu b_{jk})$.

Beweis. Nachrechnen! □

2.21 Bemerkung. Damit ist $M(m \times n\,, \mathbb{K})$ selbst ein $\mathbb{K}$-Vektorraum.

2.22 Satz. **Matrixmultiplikation**

Gegeben seien zwei lineare Abbildungen $T : U \to V$ und $S : V \to W$ zwischen Vektorräumen U, V und W über $\mathbb{K}$ mit den Basen $\mathcal{A} = (u_1, \ldots, u_n)$ für U, $\mathcal{B} = (v_1, \ldots, v_m)$ für V und $\mathcal{C} = (w_1, \ldots, w_\ell)$ für W.
Die entsprechenden Matrizen seien $A = M_{\mathcal{B}}^{\mathcal{C}}(S)$ und $B = M_{\mathcal{A}}^{\mathcal{B}}(T)$. Dann definiert man das Matrixprodukt durch $A \cdot B := M_{\mathcal{A}}^{\mathcal{C}}(ST) =: C$ und es gilt

$$c_{ik} = \sum_{j=1}^{m} a_{ij} b_{jk} \qquad \text{für } i = 1, \ldots, \ell; \quad k = 1, \ldots, n\,.$$

Beweis. Nachrechnen: Nach Definition der c_{ik} gilt

$$(S \circ T)u_k = \sum_{i=1}^{\ell} c_{ik}\, w_i \,.$$

Andererseits ist

$$Tu_k = \sum_{j=1}^{m} b_{jk}\, v_j \qquad \text{und} \qquad Sv_j = \sum_{i=1}^{\ell} a_{ij}\, w_i \,.$$

Also

$$S(Tu_k) = S \sum_{j=1}^{m} b_{jk}\, v_j = \sum_{j=1}^{m} b_{jk}\, Sv_j = \sum_{j=1}^{m} b_{jk} \sum_{i=1}^{\ell} a_{ij}\, w_i = \sum_{i=1}^{\ell} \underbrace{\left(\sum_{j=1}^{m} a_{ij}\, b_{jk} \right)}_{c_{ik}} w_i \,.$$

□

2.23 Merkregel. Man merkt sich die Matrixmultiplikation graphisch folgendermaßen (zeichnen Sie sich das auf!):
Zur Berechnung von c_{ik} denke man sich die i-te Zeile der Matrix A über die k-te Spalte der Matrix B gelegt und bilde die Summe der Produkte übereinanderstehender Koeffizienten.
Die k-te Spalte von $A \cdot B$ entsteht also durch Matrix-Vektor-Multiplikation Ab_k, wobei b_k die k-te Spalte von B ist.

2.24 Bemerkung. Statt $A \cdot B$ schreibt man meistens AB.

2.25 Beispiele. (a) Seien

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & -1 \\ 0 & 0 \\ -1 & 1 \end{pmatrix} \,.$$

Dann ist

$$A \cdot B = \begin{pmatrix} 1\cdot 1 + 2\cdot 0 + 3\cdot(-1) & 1\cdot(-1) + 2\cdot 0 + 3\cdot 1 \\ 3\cdot 1 + 2\cdot 0 + 1\cdot(-1) & 3\cdot(-1) + 2\cdot 0 + 1\cdot 1 \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix}$$

und

$$B \cdot A = \begin{pmatrix} -2 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & -2 \end{pmatrix} \,.$$

(b) Sei

$$A = (\mathrm{i}, -\mathrm{i}, 1) \quad \text{und} \quad B = \begin{pmatrix} 1 & 1 \\ \mathrm{i} & -1 \\ -\mathrm{i} & \mathrm{i} \end{pmatrix} \,.$$

Dann ist $AB = (1, 3\mathrm{i})$ und BA macht keinen Sinn.

2.26 Satz. **Rechenregeln für die Matrixmultiplikation**
Für Matrizen A, B und C gilt, sofern die Mutliplikation aus Formatgründen definiert ist,

$$\begin{aligned} A(BC) &= (AB)C && \text{(Assoziativgesetz)} \\ A(B+C) &= AB + AC && \text{(Distributivgesetze)} \\ (A+B)C &= AC + BC && \end{aligned}$$

Beweis. Folgt sofort aus den entsprechenden Eigenschaften der assoziierten linearen Abbildungen. □

2.27 Warnung. Die Matrixmultiplikation ist weder kommutativ noch „nullteilerfrei“, d.h.

(1) Es gibt quadratische Matrizen A, B mit $AB \neq BA$.

(2) Es gibt Matrizen $A \neq 0, B \neq 0$ mit $AB = 0$.

Beweis. Seien $A = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$ und $B = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, so ist

$$AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0 \quad \text{und} \quad BA = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} \neq AB\,.$$

□

2.28 Bemerkung. Die Algebra der $n \times n$ -Matrizen

Multipiziert man zwei Matrizen $A, B \in M(n \times n, \mathbb{K})$, so ist das Produkt AB wieder eine $n \times n$-Matrix. $M(n, \mathbb{K}) := M(n \times n, \mathbb{K})$ ist also ein $\mathbb{K}$-Vektorraum, für dessen Elemente auch eine Multiplikation

$$M(n, \mathbb{K}) \times M(n, \mathbb{K}) \to M(n, \mathbb{K})$$

erklärt ist. Die Multiplikation ist assoziativ und distributiv und es gilt $\alpha(AB) = (\alpha A)B = A(\alpha B)$ für $\alpha \in \mathbb{K}$ und $A, B \in M(n, \mathbb{K})$.

Ein Vektorraum mit einer solchen Multiplikation heißt **Algebra.** Da es in $M(n, \mathbb{K})$ ein neutrales Element der Multiplikation gibt, nämlich die Einheitsmatrix E_n und da die Multiplikation für $n \geq 2$ nicht kommutativ ist, spricht man von einer nichtkommutativen Algebra mit Einselement.

2.29 Beispiele. Matrixdarstellung von linearen Abbildungen der Ebene

(a) **Drehungen im $\mathbb{R}^2$** (vgl. Bsp. 2.18)
Sei D_α die Drehung im $\mathbb{R}^2$ um den Ursprung mit Winkel α im positiven Sinne. Wegen $D_\varphi \circ D_\alpha = D_{\varphi+\alpha}$ muss auch $M(D_\varphi)M(D_\alpha) = M(D_{\varphi+\alpha})$ gelten: Nachrechnen für die Matrizen liefert

$$\begin{aligned} M(D_\varphi)M(D_\alpha) &= \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \\ &= \begin{pmatrix} \cos\varphi\cos\alpha - \sin\varphi\sin\alpha & -\cos\varphi\sin\alpha - \sin\varphi\cos\alpha \\ \sin\varphi\cos\alpha + \cos\varphi\sin\alpha & -\sin\varphi\sin\alpha + \cos\varphi\cos\alpha \end{pmatrix} \\ &= \begin{pmatrix} \cos(\varphi+\alpha) & -\sin(\varphi+\alpha) \\ \sin(\varphi+\alpha) & \cos(\varphi+\alpha) \end{pmatrix} = M(D_{\varphi+\alpha}) \end{aligned}$$

(b) Sei $S_y : \mathbb{R}^2 \to \mathbb{R}^2$ die **Spiegelung** an der y-Achse, d.h.

$$S_y e_1 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad S_y e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Die zugehörige Matrix bezüglich der kanonischen Basis ist also

$$M_{\mathcal{K}}(S_y) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Es gilt $S_y^2 = S_y \circ S_y = \mathbb{1}_{\mathbb{R}^2}$, also auch

$$M_{\mathcal{K}}(S_y)M_{\mathcal{K}}(S_y) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2 = M_{\mathcal{K}}(\mathbb{1}_{\mathbb{R}^2})\,.$$

(c) Sei P_y die **Projektion** auf die y-Achse, d.h.

$$P_y e_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad P_y e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Die zugehörige Matrix bezüglich der kanonischen Basis ist also

$$M_{\mathcal{K}}(P_y) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Es gilt $P_y^2 = P_y$ (das definiert eine Projektion), also auch

$$M_{\mathcal{K}}(P_y) \cdot M_{\mathcal{K}}(P_y) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = M_{\mathcal{K}}(P_y).$$

2.30 Bemerkung. Für lineare Abbildungen $A : \mathbb{K}^n \to \mathbb{K}^m$ bezeichnen wir im folgenden die zugehörige Matrix (bzgl. der kanonischen Basis) mit demselben Symbol, also mit A statt mit $M_{\mathcal{K}}(A)$.

Eine wichtige Anwendung der linearen Algebra ist das Lösen linearer Gleichungen der Form $Ax = y$ wobei y und A bekannt sind und x gesucht ist. Deshalb ist wichtig, die Umkehrabbildung A^{-1} zu verstehen, da $x = A^{-1}y$.

2.31 Satz. (a) Ist $L : V \to W$ ein Isomorphismus (also eine lineare Bijektion), so ist auch die Umkehrabbildung $L^{-1} : W \to V$ ein Isomorphismus.

(b) Eine lineare Abbildung $L : V \to W$ kann nur umkehrbar sein, wenn $\dim V = \dim W$ gilt.

(c) Haben V und W dieselbe endliche Dimension, so gilt für lineare Abbildung $L : V \to W$

$$L \text{ ist bijektiv} \quad \Leftrightarrow \quad L \text{ ist injektiv} \quad \Leftrightarrow \quad L \text{ ist surjektiv}.$$

Beweis. (a) Wir müssen nur zeigen, dass L^{-1} linear ist. Seien $w_1, w_2 \in W$ und $\alpha \in \mathbb{K}$, dann gibt es wegen der Bijektivität eindeutig bestimmte Vektoren $v_1, v_2 \in V$ mit

$$w_1 = L(v_1),\, w_2 = L(v_2) \quad \text{bzw.} \quad v_1 = L^{-1}(w_1),\, v_2 = L^{-1}(w_2).$$

Da L linear ist, gilt

$$L(\alpha v_1 + v_2) = \alpha L(v_1) + L(v_2) = \alpha w_1 + w_2,$$

also

$$L^{-1}(\alpha w_1 + w_2) = \alpha_1 v_1 + v_2 = \alpha L^{-1}(w_1) + L^{-1}(w_2).$$

(b) Sei $L : V \to W$ umkehrbar, also injektiv und sujektiv. Dann gilt

$$\begin{aligned} L \quad \text{surjektiv} &\Rightarrow \quad \operatorname{Bild} L = W \\ L \quad \text{injektiv} &\overset{2.8}{\Rightarrow} \quad \operatorname{Kern} L = \{0\}, \end{aligned}$$

und mit der Dimensionsformel Satz 2.10 folgt

$$\dim V = \underbrace{\dim \operatorname{Kern} L}_{=0} + \dim \underbrace{\operatorname{Bild} L}_{=W} = \dim W.$$

(c) Es ist nun zu zeigen, dass für $\dim V = \dim W = n < \infty$ und $L \in \mathcal{L}(V, W)$ gilt

$$L \quad \text{injektiv} \quad \Leftrightarrow \quad L \quad \text{surjektiv}\,.$$

Also,

$$\begin{aligned} L \quad \text{injektiv} &\overset{2.8}{\Leftrightarrow} \quad \operatorname{Kern} L = \{0\} \overset{\operatorname{Kern} L \text{ ist UR}}{\Leftrightarrow} \dim\, \operatorname{Kern} L \,=\, 0 \\ &\overset{2.10}{\Leftrightarrow} \quad \dim \operatorname{Bild} L \,=\, n \overset{\operatorname{Bild} L \text{ ist UR}}{\Leftrightarrow} \operatorname{Bild} L \,=\, W \\ &\Leftrightarrow \quad L \text{ ist surjektiv}\,. \end{aligned}$$

□

2.32 Bemerkung. In unendlich dimensionalen Räumen gilt (c) nicht mehr. Beispielsweise ist der Ableitungsoperator auf $P_{\mathbb{K}}$ surjektiv (jedes Polynom ist Ableitung eines anderen Polynoms) aber nicht injektiv (verschiedene Polynome können dieselbe Ableitung besitzen).

2.33 Definition. Reguläre Matrizen

Eine Matrix $A \in M(n, \mathbb{K})$ heißt **invertierbar** oder **regulär**, wenn es eine Matrix $B \in M(n, \mathbb{K})$ gibt mit

$$AB = BA = E$$

Die Matrix B wird dann mit A^{-1} bezeichnet.

2.34 Satz. (a) Die Inversie A^{-1} einer Matrix A ist eindeutig bestimmt.

(b) Der Vektorraum V besitze die endliche Basis $\mathcal{B}$ und es sei $A = M_{\mathcal{B}}(L)$ für ein $L \in \mathcal{L}(V)$. Dann gilt

$$A \text{ invertierbar} \Leftrightarrow L \text{ invertierbar}$$

und

$$A^{-1} = M_{\mathcal{B}}(L^{-1}) \quad \text{falls } A \text{ oder } L \text{ invertierbar ist}\,.$$

(c) Sind $A, B \in M(n, \mathbb{K})$ und gilt $AB = E$, so sind beide invertierbar und zueinander invers. (Es genügt also in 2.33 eine der beiden Bedingungen zu prüfen).

Beweis. (b) „ $\Leftarrow$ “: Sei $B = M_{\mathcal{B}}(L^{-1})$. Dann folgt aus $LL^{-1} = L^{-1}L = \mathbb{1}_V$ für die entsprechenden Matrizen $M(L)M(L^{-1}) = M(L^{-1})M(L) = E$, also $AB = BA = E$.

„ $\Rightarrow$ “ Sei $B \in M(n, \mathbb{K})$ so dass

$$AB = BA = E_n \quad \text{und} \quad T = L_{\mathcal{B}}(B)\,.$$

Dann folgt für die lineare Abbildung

$$L_{\mathcal{B}}(A)\, L_{\mathcal{B}}(B) = L_{\mathcal{B}}(B)\, L_B(A) = L_{\mathcal{B}}(E_n)$$

also

$$LT = TL = \mathbb{1}_V\,.$$

Damit ist L invertierbar und $T = L^{-1}$.

(a) Die Eindeutigkeit der inversen Matrix folgt aus der Eindeutigkeit der Umkehrabbildung.

(c) Für zugehörige lineare Abbildungen gilt $S \circ T = \mathbb{1}_V$. Damit ist S surjektiv und T injektiv (Übungsaufgabe) und somit sind beide umkehrbar (Satz 2.31 (c)).
Also $S = T^{-1}$ und $T = S^{-1}$ und somit nach (b) $AB = BA = E$.

□

2.35 Bemerkung. Sind A, B invertierbare Matrizen, so ist auch AB invertierbar und

$$(AB)^{-1} \,=\, B^{-1}A^{-1}$$

Beweis. Das gilt für die Komposition beliebiger invertierbarer Abbildungen, also insbesondere auch für Matrizen. □

Basiswechsel und Koordinatentransformation

Seien $\mathcal{A} = (w_1, \ldots, w_n)$ und $\mathcal{B} = (v_1, \ldots, v_n)$ Basen von V. Ein Vektor $u \in V$ besitze die Basisdarstellungen

$$u = \sum_{k=1}^{n} x_k w_k \quad \text{bzw.} \quad u = \sum_{k=1}^{n} y_k v_k \ .$$

Wie rechnet man die Koordinatenvektoren

$$(u)_{\mathcal{A}} = x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{und} \quad (u)_{\mathcal{B}} = y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

ineinander um?

Oder analog: $L \in \mathcal{L}(V)$ besitze die Matrixdarstellungen

$$M_{\mathcal{A}}(L) = A \quad \text{und} \quad M_{\mathcal{B}}(L) = B\,.$$

Wie rechnet man A und B ineinander um?

Sei $\mathcal{A} = (w_1, \ldots, w_n)$ die Ausgangsbasis und $\mathcal{B} = (v_1, \ldots, v_n)$ die neue Basis. Dann sind die Basisvektoren v_j der neuen Basis $\mathcal{B}$ typischerweise in der Basisdarstellung der alten Basis $\mathcal{A}$ gegeben, also $v_j = \sum_{k=1}^{n} s_{kj} w_k$. Damit können wir aber direkt die Matrix S der Identitätsabbildung $\mathbb{1}_V$ hinschreiben, wobei wir im Urbild die Basis $\mathcal{B}$ wählen und im Bild die Basis $\mathcal{A}$ (erinnere: die Spalten der Matrix sind die Bilder der Basisvektoren),

$$M_{\mathcal{B}}^{\mathcal{A}}(\mathbb{1}_V) = S = (s_{kj})\,.$$

Wir erwarten natürlich, dass $x = Sy$ und $y = S^{-1}x$ gilt.

2.36 Satz. Die Transformationsmatrix $S = M_{\mathcal{B}}^{\mathcal{A}}(\mathbb{1}_V)$ ist regulär und ihre Inverse ist

$$S^{-1} = M_{\mathcal{A}}^{\mathcal{B}}(\mathbb{1}_V)\,.$$

Beweis. Es gilt

$$E_n = M_{\mathcal{A}}^{\mathcal{A}}(\mathbb{1}_V) = M_{\mathcal{A}}^{\mathcal{A}}(\mathbb{1}_V \cdot \mathbb{1}_V) = M_{\mathcal{B}}^{\mathcal{A}}(\mathbb{1}_V) M_{\mathcal{A}}^{\mathcal{B}}(\mathbb{1}_V) = S \cdot M_{\mathcal{A}}^{\mathcal{B}}(\mathbb{1}_V)\,.$$

Also ist gemäß Satz 2.34 (c)

$$M_{\mathcal{A}}^{\mathcal{B}}(\mathbb{1}_V) = S^{-1} \ .$$

□

2.37 Satz. <u>Basiswechsel</u>

Seien $\mathcal{A}$ und $\mathcal{B}$ Basen von V und $S = M_{\mathcal{B}}^{\mathcal{A}}(\mathbb{1}_V)$,

(a) Für die Koordinatorvektoren $x = (u)_{\mathcal{A}}$ und $y = (u)_{\mathcal{B}}$ eines Vektors $u \in V$ gilt

$$x = Sy \quad \text{bzw.} \quad y = S^{-1}x\,.$$

(b) Für die Matrizen $A = M_{\mathcal{A}}(L)$ und $B = M_{\mathcal{B}}(L)$ einer linearen Abbildung $L \in \mathcal{L}(V)$ gilt:

$$B = S^{-1}\,A\,S \quad \text{bzw.} \quad A = S\,B\,S^{-1}\,.$$

Beweis. (a) Es gilt mit Satz 2.19

$$x = (u)_{\mathcal{A}} = (\mathbb{1}_V \cdot u)_{\mathcal{A}} = M_{\mathcal{B}}^{\mathcal{A}}(\mathbb{1}_V)\,(u)_{\mathcal{B}} = S\,y\,.$$

Durch Multiplikation von links mit S^{-1} folgt $S^{-1}x = y$.

(b) $B = M_{\mathcal{B}}(L) = M_{\mathcal{B}}(\mathbb{1}_V\, L\, \mathbb{1}_V) = M^{\mathcal{B}}_{\mathcal{A}}(\mathbb{1})\; M^{\mathcal{A}}_{\mathcal{A}}(L)\; M^{\mathcal{A}}_{\mathcal{B}}(\mathbb{1}) = S^{-1}\, A\, S\;.$

□

2.38 Definition. **Ähnliche Matrizen**

Zwei Matrizen $A, B \in M(n, \mathbb{K})$ heißen **ähnlich**, falls es eine reguläre Matrix $S \in M(n, \mathbb{K})$ gibt, so dass

$$B = S^{-1}\, A\, S$$

gilt. Gemäß Satz 2.37 gehören ähnliche Matrizen also zur selben linearen Abbildung bezüglich verschiedener Basen.

Für $L \in \mathcal{L}(V, W)$ hatten wir bereits Rang$L = \dim$ (Bild$L$) definiert. Eine Matrix $A \in M(m\times n, \mathbb{K})$ können wir immer als lineare Abbildung $L_{\mathcal{K}}(A) : \mathbb{K}^n \to \mathbb{K}^m$ auffassen.

2.39 Definition. **Rang, Spaltenrang und Zeilenrang einer Matrix**

Ist $A \in M(m \times n, \mathbb{K})$ so nennt man

$$\operatorname{Rang} A = \dim \operatorname{Bild}\,(A\, : \mathbb{K}^n \to \mathbb{K}^m)$$

den **Rang** von A.
Die Maximalzahl linear unabhängiger Spaltenvektoren von A heißt der **Spaltenrang** und die Maximalzahl linear unabhängiger Zeilenvektoren heißt der **Zeilenrang**.

Offensichtlich ist Rang $A =$ Spaltenrang A, da Bild$A$ von den Spalten von A aufgespannt wird.

2.40 Satz. **Spaltenrang gleich Zeilenrang**

Sei $A \in M(m \times n, \mathbb{K})$. Es gilt Spaltenrang $A =$ Zeilenrang A.

Beweis. Eine Spalte oder Zeile von A heiße **linear überflüssig**, falls sie aus den anderen Spalten bzw. Zeilen linearkombiniert werden kann.
Verkleinert man eine Matrix durch Weglassen einer linear überflüssigen Spalte, so ändert sich der Spaltenrang nicht. Wir zeigen zunächst, dass sich auch der Zeilenrang nicht ändert.

Angenommen in $A = (a_{ki})_{m\times n}$ sei die j-te Spalte linear überflüssig, also gibt es $\alpha_i \in \mathbb{K}$ mit

$$a_{kj} = \sum_{i=1,\, i\neq j}^{n} \alpha_i\, a_{ki} \quad \text{für } k = 1, \ldots, m\,.$$

Wir zeigen nun, dass eine Linearkombination von Zeilen genau dann Null ist, wenn die entsprechende Linearkombination der Zeilen jeweils ohne die j-te Komponente Null ist, also

$$\sum_{k=1}^{m} \beta_k\, a_{ki} = 0 \quad \forall \text{ i} \quad \Leftrightarrow \quad \sum_{k=1}^{m} \beta_k\, a_{ki} = 0 \quad \forall \text{ i} \neq j\,.$$

Nur "$\Leftarrow$" ist zu zeigen:

$$\sum_{k=1}^{m} \beta_k\, a_{kj} = \sum_{k=1}^{m} \beta_k \sum_{i=1,\, i\neq j}^{n} \alpha_i\, a_{ki} = \sum_{i=1,\, i\neq j}^{n} \alpha_i \sum_{k=1}^{m} \beta_k\, a_{ki} = 0\,.$$

Also hat A nach Weglassen der j-ten Spalte dieselbe Maximalzahl linear unabhängiger Zeilen, also denselben Zeilenrang.
Ebenso gilt natürlich, dass das Weglassen einer linear überflüssigen Zeile den Spaltenrang nicht ändert.

Nun verkleinern wir A durch sukzessives Weglassen linear überflüssiger Spalten und Zeilen solange, bis alle Zeilen und Spalten linear unabhängig sind. Spalten- und Zeilenrang haben sich dabei nicht verändert, und es gilt Zeilenrang = Zeilenzahl und Spaltenrang = Spaltenzahl. Dann muss aber die resultierende Matrix quadratisch sein: sei s die Spaltenzahl und z die Zeilenzahl. Dann bilden die Spalten ein s-Tupel von linear unabhängigen Vektoren im $\mathbb{R}^z$, was $s \leq z$ impliziert. Analog bilden die Zeilen ein z-Tupel von linear unabhängigen Vektoren im $\mathbb{R}^s$, was $s \geq z$ impliziert. Also $s = z$ und Zeilenrang = Spaltenrang. □

2.41 Definition. **Elementare Matrixumformungen**
Man unterscheidet drei Typen **elementarer Zeilenumformungen** einer Matrix $A \in M(m \times n, \mathbb{K})$, nämlich

Typ 1: Vertauschung zweier Zeilen.

Typ 2: Multiplikation einer Zeile mit einem Skalar $\lambda \neq 0$, $\lambda \in \mathbb{K}$.

Typ 3: Addition eines beliebigen Vielfachen einer Zeile zu einer anderen Zeile.

Analog sind **elementare Spaltenumformungen** definiert.

2.42 Satz. Elementare Umformungen ändern den Rang einer Matrix nicht.

Beweis. Zeilenumformungen verändern die lineare Hülle der Zeilenvektoren nicht, also bleibt auch Zeilenrang = dim(Span{Zeilen}) erhalten. Wegen Zeilenrang = Spaltenrang bleibt der Rang erhalten. Analog argumentiert man bei Spaltenumformungen. □

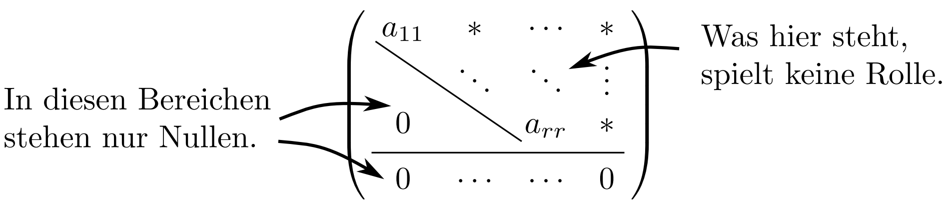
Zur Bestimmung des Rangs einer Matrix führt man also solange elementare Umformungen durch, bis man den Rang einfach ablesen kann, z.B. (Telefonmatrix)

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \quad \to \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 4 & 2 & 0 \end{pmatrix} \quad \to \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \Rightarrow \quad \mathrm{Rang} A = 2\,.$$

2.43 Definition. Die Elemente a_{ii} einer Matrix heißen die **(Haupt-)diagonalelemente**. Von den anderen Elementen sagt man, sie stünden „oberhalb“ bzw. „unterhalb“ der Diagonalen, je nachdem ob $i < j$ oder $i > j$ ist.

2.44 Definition. und Satz **Zeilenstufenform**
Eine Matrix $A \in M(m \times n, \mathbb{K})$ hat die **Zeilenstufenform**, falls die ersten r Hauptdiagonalelemente a_{ii} von Null verschieden sind, die letzten $m - r$ Zeilen sowie alle Elemente unterhalb der Hauptdiagonalen jedoch gleich Null sind. Eine Matrix A in Zeilenstufenform hat Rang$A = r$.

Beweis. Aus der Zeichnung ist Zeilenrang$A = r$ ist offensichtlich.



□

Verfahren zur Bestimmung des Ranges einer Matrix

Sei $A \in M(m \times n, \mathbb{K})$ gegeben und $A \neq 0$. (Sonst ist Rang $A = 0$ und wir sind fertig).

Durch Vertauschen von Zeilen und Spalten (falls notwendig) können wir $a_{11} \neq 0$ erreichen.

Dann zieht man für $k \geq 2$ von der k-ten Zeile das $\frac{a_{k1}}{a_{11}}$-fache der ersten Zeile ab und bekommt eine Matrix der Form

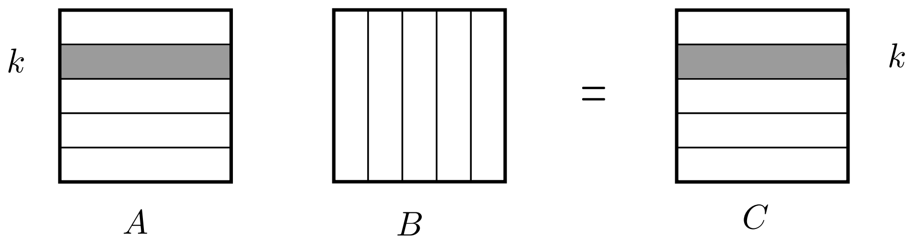
$$\left(\begin{array}{c|ccc} a_{11} & a_{12} & \cdots & a_{1n} \\ \hline 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{array}\right)$$

Nun verfährt man analog mit A' und so weiter, bis entweder $A' = 0$ oder $A' \in M(1 \times n - m + 1, \mathbb{K})$. In beiden Fällen liegt die Matrix in Zeilenstufenform vor und man kann den Rang gemäß Satz 2.44 ablesen.

Verfahren zur Bestimmung der inversen Matrix

Beobachtung: Gilt für drei Matrizen $A, B, C \in M(n, \mathbb{K})$ die Gleichung $AB = C$ und überführt man A und C durch die gleichen elementaren Zeilenumformungen in Matrizen A' und C' so gilt auch $A'B = C'$.

Beweis. Die k-te Zeile von A liefert durch Multiplikation mit den Spalten von B die k-te Zeile von C. Also führen alle elementaren Zeilenunformungen in A zu den gleichen Umformungen im Ergebnis C.



□

2.45 Satz. <u>**Bestimmung der inversen Matrix**</u>
Erhält man E durch elementare Zeilenumformungen aus A, so verwandeln dieselben Zeilenumformungen die Matrix E in A^{-1}.

Beweis. Obige Beobachtung für $AA^{-1} = E$. □

2.46 Beispiel. Sei

$$A = \begin{pmatrix} 0 & -1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 0 & 0 & 2 \end{pmatrix} \qquad \in M(4 \times 4, \mathbb{R})$$

Wir wenden nun auf A und E jeweils die selben Zeilenumformungen an:

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

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

$$\left(\begin{array}{cccc} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{array}\right) \qquad \left(\begin{array}{cccc} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{cccc} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \end{array}\right) \qquad \left(\begin{array}{cccc} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{cccc} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{array}\right) \qquad \left(\begin{array}{cccc} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & -2 & 1 & 1 \end{array}\right)$$

$$E = \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) \qquad \left(\begin{array}{cccc} -1 & 2 & -1 & 0 \\ -\frac{1}{2} & -1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -1 & \frac{1}{2} & \frac{1}{2} \end{array}\right) = A^{-1}$$

Ein Test kann nicht schaden:

$$\begin{aligned} AA^{-1} &= \frac{1}{2}\left(\begin{array}{cccc} 0 & -1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 0 & 0 & 2 \end{array}\right)\left(\begin{array}{cccc} -2 & 4 & -2 & 0 \\ -1 & -2 & 1 & 1 \\ 1 & 0 & 1 & -1 \\ 1 & -2 & 1 & 1 \end{array}\right) \\ &= \frac{1}{2}\left(\begin{array}{cccc} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{array}\right) = \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) = E \end{aligned}$$

Das allgemeine Verfahren sollte aus diesem Beispiel klar werden!

2.47 Bemerkung. Die Inverse einer 2×2-Matrix

$$A = \left(\begin{array}{cc} a & b \\ c & d \end{array}\right) \quad \text{ist} \quad A^{-1} = \frac{1}{ad - bc}\left(\begin{array}{cc} d & -b \\ -c & a \end{array}\right).$$

3 Lineare Gleichungen

3.1 Definition. **Lineare Gleichungssysteme**

Ein lineares Gleichungssystem aus m Gleichungen für n Unbekannte hat die Form

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

Gegeben sind die Koeffizienten $a_{ik} \in \mathbb{K}$ und die Zahlen $b_k \in \mathbb{K}$. Gesucht sind alle n-Tupel $x = (x_1, \ldots, x_n)$ welche diese Gleichungen erfüllen.

In Matrix-Vektor Notation kann man das System in kompakter Form schreiben als

$$A\,x = b\,,$$

wobei

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in M(m \times n, \mathbb{K})\,, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n \quad \text{und} \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{K}^m\,.$$

Ein lineares Gleichungssystem ist lösbar, falls $b \in \mathrm{Bild} A$, also falls

$$\mathrm{Rang}\, A = \mathrm{Rang}\,(\,A\,|\,b\,) \quad \text{(um die Spalte } b \text{ erweiterte Matrix)}\,.$$

3.2 Beispiele. (a) **Basisdarstellung**: Ist $\mathcal{A} = (v_1, \ldots, v_n)$ eine Basis des $\mathbb{K}^n$, so führt die Bestimmung der Basisdarstellung eines vorgegebenen Vektors $u \in \mathbb{K}^n$ auf ein $n \times n$- Gleichungssystems $u = x_1v_1 + \cdots x_nv_n = Ax$ mit $A = (v_1, \ldots, v_n) \in M(n, \mathbb{K})$.

(b) Das Nachprüfen der linearen Unabhängigkeit von n Vektoren $v_1, \ldots, v_n$ des $\mathbb{K}^m$ führt auf das homogene (d.h. $b = 0$) $m \times n$-Gleichungssystem $0 = x_1v_1 + \cdots + x_nv_n = Ax$ mit $A = (v_1, \ldots v_n) \in M(m \times n, \mathbb{K})$ und $x \in \mathbb{K}^n$.

3.3 Definition. **Allgemeine lineare Gleichungen**

Gegeben ist eine lineare Abbildung $L : V \to W$ und ein Vektor $b \in W$. Gesucht sind alle Lösungen $u \in V$ der **linearen Gleichung**

$$Lu = b\,.$$

Falls $b = 0$, so heißt die Gleichung **homogen**, falls $b \neq 0$, so heißt die Gleichung **inhomogen**.

Lineare Gleichungssysteme sind ein Spezialfall dieses Konzeptes mit $V = \mathbb{K}^n$, $W = \mathbb{K}^m$ und $L : x \mapsto Ax$ für $A \in M(m \times n, \mathbb{K})$.

3.4 Definition. (a) Ist $L : V \to W$ injektiv, so hat die Gleichung $Lu = b$ für jedes $b \in W$ höchstens eine Lösung $u \in V$. Man spricht dann von **eindeutiger Lösbarkeit** des Problems $Lu = b$.

Wegen L injektiv $\Leftrightarrow$ $\mathrm{Kern}L = \{0\}$ ist $Lu = b$ genau dann eindeutig lösbar, wenn $Lu = 0$ nur die "triviale Lösung" $u = 0$ besitzt.

(b) Ist $L : V \to W$ surjektiv, so hat die Gleichung $Lu = b$ für jedes $b \in W$ mindestens eine Lösung $u \in V$. Man sagt dann, das Problem $Lu = b$ ist **universell lösbar**.

Ist $L : V \to W$ bijektiv, so ist das Problem $Lu = b$ **universell und eindeutig lösbar**, d.h. die Gleichung $Lu = b$ besitzt für jedes $b \in W$ genau eine Lösung $u \in V$.

3.5 Satz. Sei $\mathcal{L}_0 = \operatorname{Kern} L$ die Lösungsmenge der homogenen Gleichung $Lu = 0$. Ist u_b eine spezielle Lösung der inhomogenen Gleichung $Lu = b$, dann ist die gesamte Lösungsmenge („die allgemeine Lösung") durch

$$\mathcal{L}_b = u_b + \mathcal{L}_0 := \{u_b + u_0 \,|\, u_0 \in \mathcal{L}_0\}$$

gegeben.

Beweis. Sei $Lu_b = b$ und $Lu_0 = 0$, dann ist $L(u_b + u_0) = b$, also $u_b + u_0 \in \mathcal{L}_b$. Sei andererseits auch $L\tilde{u}_b = b$, dann ist $u = \tilde{u}_b - u_b \in \mathcal{L}_0$, da $Lu = L\tilde{u}_b - Lu_b = b - b = 0$. □

3.6 Bemerkung. Geometrisch ist die Lösungsmenge $\mathcal{L}_b$ einer linearen Gleichung also ein **affiner Unterraum** von V, d.h. ein um einen Vektor verschobener Unterraum.

3.7 Merkregel. Addiert man zu einer Lösung der inhomogenen Gleichung eine Lösung der homogenen Gleichung, so erhält man wieder eine Lösung der inhomogenen Gleichung. Umgekehrt ist die Differenz von je zwei Lösungen der inhomogenen Gleichung eine Lösung der homogenen Gleichung.

Verfahren zur Lösung linearer Gleichungssysteme: Der Gauß'sche Algorithmus

Beobachtung: Verändert man die erweiterte Matrix

$$\left(\begin{array}{c|c} A & b \end{array} \right)$$

durch elementare Zeilenumformungen zu einer Matrix

$$\left(\begin{array}{c|c} A' & b' \end{array} \right)$$

so gilt

$$Ax = b \Leftrightarrow A'x = b'\,.$$

Beweis. Spezialfall von $AB = C \to A'B = C'$, wie beim Invertieren von Matrizen bewiesen. □

Wir erläutern den Algorithmus an einem Beispiel:

$$\begin{array}{rcrcrcl} x_1 & + & 2\,x_2 & + & 3\,x_3 & = & 1 \\ 4\,x_1 & + & 5\,x_2 & + & 6\,x_3 & = & 3 \\ 7\,x_1 & + & 8\,x_2 & + & 9\,x_3 & = & 5 \end{array}$$

Die zugehörige erweiterte Matrix ist

$$\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 3 \\ 7 & 8 & 9 & 5 \end{array}$$

Elementare Zeilenumformungen liefern

$$\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & -3 & -6 & -1 \\ 0 & -6 & -12 & -2 \end{array} \quad \longrightarrow \quad \begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 3 & 6 & 1 \\ 0 & 0 & 0 & 0 \end{array} \quad \swarrow \text{ Wenn hier keine Null stünde, dann gäbe es keine Lösung!}$$

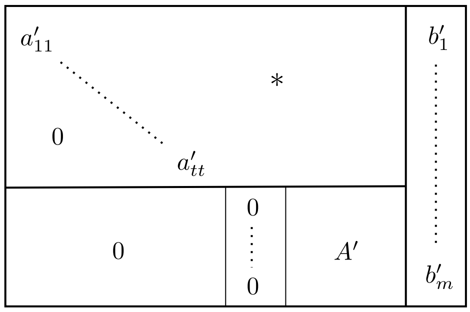
Rückübersetzung:

$$\begin{array}{rcrcrcl} x_1 & + & 2\,x_2 & + & 3\,x_3 & = & 1 \\ & & 3\,x_2 & + & 6\,x_3 & = & 1 \end{array}$$

Wähle $x_3 = s$, also $x_2 = \frac{1}{3} - 2x_3 = \frac{1}{3} - 2s$ und $x_1 = 1 - 2x_2 - 3x_3 = 1 - \frac{2}{3} + 4s - 3s = \frac{1}{3} + s$.
Die allgemeine Lösung lautet also

$$\mathcal{L}_b = \left\{ x = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ 0 \end{pmatrix} + s \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \,\middle|\, s \in \mathbb{R} \right\}$$

Das allgemeine Schema lautet also: bringe die erweiterte Matrix in Zeilenstufenform und löse das System dann von hinten her auf. Da man aber zunächst nur Zeilenumformungen verwendet, kann folgende Situation eintreten:

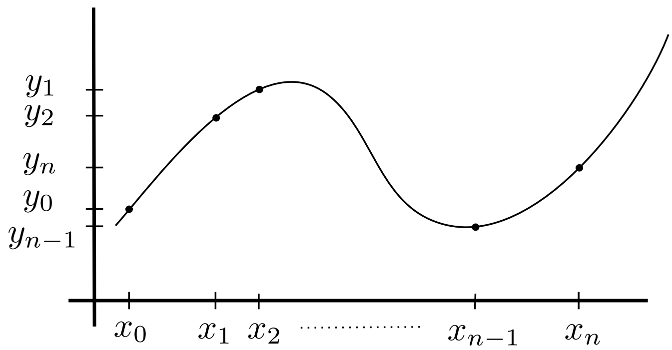


Dann muss man eben doch Spalten vertauschen und die Unbekannten dementsprechend umnummerieren.

Anwendung: Interpolation von Funktionen durch Polynome

Gegeben seien Stützstellen $x_0 < x_1 < \cdots < x_n$ und Zahlenwerte $y_0, y_1, \ldots, y_n \in \mathbb{R}$.
Gesucht ist ein Polynom p von Grad $\leq n$ mit

$$p(x_0) = y_0, \ \ldots, \ p(x_n) = y_n\,.$$



Um das als lineares Gleichungssystem zu formulieren, definieren wir auf $P_{\mathbb{R}}^{(n)}$ die lineare Abbildung

$$L : P_{\mathbb{R}}^{(n)} \to \mathbb{R}^{n+1},\ p \mapsto Lp = \begin{pmatrix} p(x_0) \\ \vdots \\ p(x_n) \end{pmatrix}.$$

Das Interpolationsproblem lautet dann:
Gegeben $y = (y_0, \ldots, y_n) \in \mathbb{R}^{n+1}$. Gesucht ist $p \in P_{\mathbb{R}}^{(n)}$ mit $Lp = y$.

Da Kern$L = \{0\}$ gilt (ein nichttriviales Polynom vom Grade $\leq n$ hat höchstens n Nullstellen), ist L injektiv. Da dim$P_{\mathbb{R}}^{(n)} = n + 1 =$ dim$\mathbb{R}^{n+1}$ ist, folgt die Surjektivität und somit die Bijektivität von L. Also ist $Lp = y$ universell und eindeutig lösbar.

Um nun das Interpolationspolynom zu berechnen, müssen wir L invertieren, $p = L^{-1}y$. Dazu genügt es wie immer, die Bilder der kanonischen Basisvektoren zu berechnen,

$$L^{-1}y = L^{-1}(y_0e_1 + \ldots + y_ne_{n+1}) = y_0\, L^{-1}e_1 + \ldots + y_n\, L^{-1}e_{n+1}\,.$$

Die Polynome $\ell_k = L^{-1}e_{k+1}$ heißen Lagrange-Polynome. Da $\ell_k(x_j) = 0$ für $j \neq k$ und $\ell_k(x_k) = 1$, ist

$$\ell_k(x) = \prod_{j\neq k} \frac{(x - x_j)}{(x_k - x_j)}\,.$$

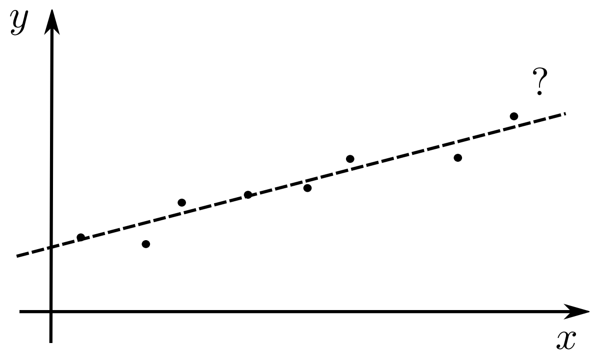
Insgesamt haben wir also $p(x) = \sum_{k=0}^{n} y_k\, \ell_k\,(x)$.

Anwendung: Die Methode der kleinsten Quadrate

Beispiel: Zwischen den Messgrößen X und Y wird ein linearer Zusammenhang der Form

$$Y = aX + b$$

mit unbekannten Parametern $a, b \in \mathbb{R}$ vermutet.



Ein Experiment ergibt für (x, y) Wertepaare

$$(x_1, y_1), \ldots, (x_m, y_m)\,.$$

Aufgrund von Messfehlern liegen aber die Punkte (x_k, y_k) nicht alle auf einer Geraden.

Mit Hilfe der **Methode der kleinsten Quadrate** bestimmt man eine Ausgleichsgerade durch die Punkte (x_k, y_k).

Für eine Gerade $y = ax + b$ ist

$$Q(a, b) = \sum_{k=1}^{m} (ax_k + b - y_k)^2$$

die Summe der vertikalen Abstandsquadrate der Punkte (x_k, y_k) von der Geraden $y = ax + b$. Man bestimme nun a und b so, dass $Q(a, b)$ minimal wird.

Wir verallgemeinern das Problem zunächst ein wenig: sei eine $m \times n$-Matrix A mit $m > n =$ Rang$A$ und ein Vektor $y \in \mathbb{R}^m$ gegeben. Dann ist das Gleichungssystem $Ax = y$ im allgemeinen unlösbar, wenn nicht gerade y in Bild A liegt. Man sagt, die Gleichung ist **überbestimmt**. Wie zuvor kann man aber einen Vektor u suchen, so dass $Au - y$ so klein wie möglich wird und zwar im Sinne der „kleinsten Quadrate“: Minimiere

$$\|Au - y\|^2_{\mathbb{R}^m} = \sum_{i=1}^{m} \left(\sum_{k=1}^{n} a_{ik} u_k - y_i \right)^2 .$$

Das Problem der Ausgleichsgerade ist als Spezialfall enthalten:

$$A = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_m \end{pmatrix} \quad \text{und} \quad u = \begin{pmatrix} b \\ a \end{pmatrix} .$$

Eine notwendige Bedingung für ein Minimum ist das Verschwinden der partiellen Ableitungen $\partial_{u_j} ||Au - y||^2 \; = \; 0$ für $\; j = 1, \ldots, n$, also

$$\begin{aligned} 0 \quad &= \quad \frac{\partial}{\partial u_j} \sum_{i=1}^{m} \left(\sum_{k=1}^{n} a_{ik}\, u_k - y_i \right)^2 \\ &= \quad 2 \sum_{i=1}^{m} a_{ij} \left(\sum_{k=1}^{n} a_{ik}\, u_k - y_i \right) \\ &= \quad 2 \sum_{k=1}^{n} u_k \sum_{i=1}^{m} a_{ij}\, a_{ik} - 2 \sum_{i=1}^{m} a_{ij}\, y_i \, . \end{aligned}$$

Mit der zu A **transponierten Matrix** (Zeilen und Spalten vertauscht!)

$$A^T := \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{mn} \end{pmatrix}$$

ergibt sich also die notwendige Bedingung

$$A^T A\, u = A^T y \, .$$

Diese Gleichung heißt die **Gaußsche Normalengleichung**.

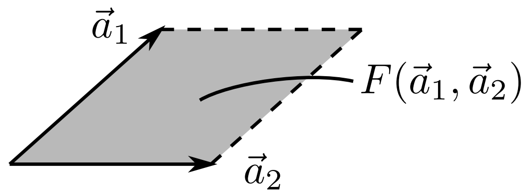
Wir werden später sehen, dass diese Gleichung universell und eindeutig lösbar ist und, dass die Lösung tatsächlich $\|Au - y\|^2$ minimiert.

4 Determinanten

Die 2×2-Determinante

Das lineare Gleichungssystem $Ax = b$ mit $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = (\vec{a}_1, \vec{a}_2)$ und $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ ist genau dann universell und eindeutig lösbar, wenn A bijektiv ist, also wenn $\mathrm{Span}\{\vec{a}_1,\, \vec{a}_2\} = \mathbb{R}^2$ ist, d.h. wenn $\vec{a}_1$ und $\vec{a}_2$ linear unabhängig sind.

Wann sind zwei Vektoren in $\mathbb{R}^2$ linear unabhängig ?
Genau dann, wenn das von ihnen aufgespannte Parallelogramm nicht entartet ist, also wenn die (orientierte) Fläche



$$F(\vec{a}_1, \vec{a}_2) = |\vec{a}_1||\vec{a}_2| \sin\varphi = |\vec{a}_1||\vec{a}_2| \cos(\varphi + \tfrac{\pi}{2}) = \vec{a}_1 \cdot \vec{a}_2^{\perp} = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} \cdot \begin{pmatrix} a_{22} \\ -a_{12} \end{pmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

ungleich Null ist.
Für zwei Vektoren $x, y \in \mathbb{R}^2$ nennt man

$$\det(x, y) = x_1 y_2 - x_2 y_1$$

die Determinante bzw. die Abbildung

$$\det : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}\,, (x, y) \mapsto \det(x, y)$$

die Determinantenform auf $\mathbb{R}^2$.
Die Determinante der Matrix $A = (\vec{a}_1, \vec{a}_2)$ ist $\det A := \det(\vec{a}_1, \vec{a}_2)$.
Die Determinantenform ist

(a) linear in jedem Argument, d.h. $x \mapsto \det(x, y)$ und $y \mapsto \det(x, y)$ sind Linearformen.
(b) antisymmetrisch bzw. alternierend, d.h. $\det(x, y) = -\det(y, x)$.
(c) normiert, d.h. $\det(e_1, e_2) = 1$ für die kanonischen Einheitsvektoren.

Die 3×3 Determinante

Auch in drei Dimensionen sind Vektoren genau dann linear unabhängig, wenn das von ihnen aufgespannte Volumen

$$\mathrm{Vol}\,(x, y, z) = x \cdot (y \times z)$$

von Null verschieden ist. Hier ist

$$\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3,\ (x, y) \mapsto x \times y := \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}$$

das Vektorprodukt im $\mathbb{R}^3$. Der Vektor $x \times y$ steht senkrecht auf der von x und y aufgespannten Ebene und sein Betrag $|x \times y|$ liefert den Flächeninhalt des von x und y aufgespannten Parallelogramms.

Wieder ist die Abbildung

$$\det : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}, \quad (x, y, z) \mapsto x \cdot (y \times z)$$

eine alternierende, normierte Multilinearform, d.h.

(a) det ist linear in jeder Variablen,

(b) Vertauschung zweier Argumente verändert das Vorzeichen, $\det(x, y, z) = -\det(y, x, z) = \det(y, z, x)$ usw.,

(c) $\det(e_1, e_2, e_3) = 1$.

Wieder kann man auch die Determinante einer 3×3-Matrix $A = (\vec{a}_1, \vec{a}_2, \vec{a}_3)$, $\vec{a}_j \in \mathbb{R}^3$, durch die Determinante der Spaltenvektoren definieren: $\det A := \det(\vec{a}_1, \vec{a}_2, \vec{a}_3)$.

Man könnte nun die Determinante in $\mathbb{R}^n$ mittels des von n Vektoren aufgespannten Volumens definieren. Das ist aber unpraktisch und im Moment ist ja auch noch gar nicht klar, wie dieses Volumen genau definiert ist. Stattdessen nehmen wir deshalb die Eigenschaften (a), (b) und (c) als Grundlage der Definition.

4.1 Definition. <u>Determinantenform</u>

Eine Abbildung

$$F : \underbrace{\mathbb{K}^n \times \cdots \times \mathbb{K}^n}_{m-\text{mal}} \to \mathbb{K} \qquad (n \geq 2)$$

heißt

(a) **Multilinearform auf $\mathbb{K}^n$**, kurz **m-Form**, wenn F in jedem der m Argumente linear ist:

$$F(x_1, \ldots, x_j + \alpha y, \ldots, x_m) = F(x_1, \ldots, x_j, \ldots, x_m) + \alpha F(x_1, \ldots, y, \ldots, x_m)$$

für alle $x_1, \ldots, x_m, y \in \mathbb{K}^n$, $\alpha \in \mathbb{K}$, $j = 1, \ldots, m$.

(b) **Alternierende m-Form** auf $\mathbb{K}^n$, wenn (a) gilt und wenn F beim Vertauschen zweier Argumente das Vorzeichen ändert:

$$F(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_m) = -F(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_m) \qquad \forall\, i \neq j \in \{1, \ldots, m\}\,.$$

(c) **Determinantenform** auf $\mathbb{K}^n$, wenn $m = n$ gilt, (a) und (b) erfüllt sind, und $F(e_1, \ldots, e_n) = 1$ für die kanonische Basis $(e_1, \ldots, e_n)$ des $\mathbb{K}^n$ gilt.

4.2 Satz. Für eine alternierende m-Form F auf $\mathbb{K}^n$ gilt: Sind $a_1, \ldots, a_m$ linear abhängig, so folgt $F(a_1, \ldots, a_m) = 0$. Insbesondere ist $F(a_1, \ldots, a_m) = 0$, falls zwei Einträge gleich sind.

Beweis. Für zwei gleiche Einträge ergibt das Vertauschen $F(\ldots, a, \ldots, a, \ldots) = -F(\ldots, a, \ldots, a, \ldots)$ also $F(\ldots, a, \ldots, a, \ldots) = 0$.

Seien $a_1, \ldots, a_m \in \mathbb{K}^n$ linear abhängig, etwa $a_1 = \sum_{k=2}^{m} \alpha_k a_k$. Dann gilt wegen der Lineariät im ersten Argument

$$F(a_1, \ldots, a_m) = \sum_{k=2}^{m} \alpha_n F(a_k, a_2, \ldots, a_k, \ldots, a_m) = 0\,.$$

□

Es folgt, dass jede alternierende m-Form auf $\mathbb{K}^n$ mit $m > n$ identisch Null sein muss.

4.3 Satz. <u>Kriterium für das Alternieren einer m-Form</u>

Verschwindet eine m-Form $F(a_1, \ldots, a_m)$ falls zwei benachbarte Argumente gleich sind, also

$$a_k = a_{k+1} \text{ für ein } k \in \{1, \ldots, m-1\} \quad \Rightarrow \quad F(a_1, \ldots, a_m) = 0\,,$$

so ist sie alternierend.

Beweis. Wegen

$$\begin{aligned} 0 &= F(\ldots, a+b, a+b, \ldots) \\ &= F(\ldots, a, a, \ldots) + F(\ldots, a, b, \ldots) + F(\ldots, b, a, \ldots) + F(\ldots, b, b, \ldots) \\ &= F(\ldots, a, b, \ldots) + F(\ldots, b, a, \ldots) \end{aligned}$$

ändert F sein Vorzeichen bei Vertauschungen benachbarter Argumente. Sind a_j und a_k für $j < k$ nicht benachbart, so erreicht man durch sukzessives Vertauschen benachbarter Argumente, dass a_j und a_k die Plätze tauschen, wobei man das mit $2(k-j)-1$ Vertauschungen erreicht,

$$\ldots, a_j, \underbrace{a_{j+1}, \ldots, a_{k-1}, a_k}_{k-j} \ldots .$$

Man benötigt beispielsweise $k-j$ Vertauschungen, um a_j an die k-te Stelle zu bringen. Dann steht a_k an der $k-1$-ten Stelle und man benötigt $k-j-1$ Vertauschungen um a_k and die j-te Stelle zu bringen.

Bei jeder Vertauschung ändert F das Vorzeichen, also

$$\begin{aligned} F(\ldots, a_j, \ldots, a_k, \ldots) &= (-1)^{2(k-j)-1} F(\ldots, a_k, \ldots, a_j, \ldots) \\ &= -F(\ldots, a_k, \ldots, a_j, \ldots) . \end{aligned}$$

□

4.4 Satz. **Existenz und Eindeutigkeit der Determinantenform auf $\mathbb{K}^n$**

(a) Für jedes $n \geq 2$ gibt es genau eine Determinantenform auf $\mathbb{K}^n$. Sie wird mit $\det_n$ oder einfach nur det bezeichnet.

(b) Hat die $n \times n$-Matrix $A = (a_{jk})$ die Spalten $a_1, \ldots, a_n \in \mathbb{K}^n$ so heißt $\det(a_1, \ldots, a_n)$ die Determinante von A, bezeichnet mit

$$|A| = \det A = \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \\ a_{m1} & \ldots & a_{mn} \end{vmatrix} .$$

(c) Die Determinante einer $n \times n$-Matrix $A = (a_{jk})$ läßt sich wie folgt durch $(n-1) \times (n-1)$-Determinanten ausdrücken:

$$\begin{aligned} |A| &= \sum_{j=1}^{n} (-1)^{i+j} a_{ij} |A_{ij}| \qquad \text{(Entwicklung nach der } i\text{-ten Zeile)} \\ &= \sum_{j=1}^{n} (-1)^{j+k} a_{jk} |A_{jk}| \qquad \text{(Entwicklung nach der } k\text{-ten Spalte).} \end{aligned}$$

Dabei sei A_{ij} diejenige $(n-1) \times (n-1)$-Matrix, welche aus A durch Streichen der i-ten Zeile und der j-ten Spalte hervorgeht.

4.5 Beispiele. (a) Durch Entwicklung nach der letzten Spalte erhalten wir

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} - a_{23} \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} + a_{33} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \\ &= a_{13}(a_{21}a_{32} - a_{31}a_{22}) - a_{23}(a_{11}a_{32} - a_{31}a_{12}) + a_{33}(a_{11}a_{22} - a_{21}a_{12}) \\ &= \vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3) . \end{aligned}$$

Merkschema für 3×3-Determinanten:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{matrix} \quad = \underline{a_{11}a_{22}a_{33}} + \underline{a_{12}a_{23}a_{31}} + \underline{a_{13}a_{21}a_{32}}$$
$$- \underline{a_{31}a_{22}a_{13}} - \underline{a_{32}a_{23}a_{11}} - \underline{a_{33}a_{21}a_{12}}$$

(b) In konkreten Beispielen entwickelt man zweckmäßigerweise nach denjenigen Spalten bzw. Zeilen, die möglichst viele Nullen enthalten, z.B. hier nach der 2-ten Spalte

$$\begin{vmatrix} 2 & 1 & 2 & 2 \\ 1 & 0 & 1 & 1 \\ -2 & 0 & 1 & -2 \\ 3 & 4 & 1 & 0 \end{vmatrix} = -1 \begin{vmatrix} 1 & 1 & 1 \\ -2 & 1 & -2 \\ 3 & 1 & 0 \end{vmatrix} + 4 \begin{vmatrix} 2 & 2 & 2 \\ 1 & 1 & 1 \\ -2 & 1 & -2 \end{vmatrix} = -(-6-2-3+2) = 9\,.$$

Wir beweisen Satz 4.4 in mehreren Schritten.

4.6 Satz. Stimmen zwei alternierende n-Formen F und G auf dem $\mathbb{K}^n$ auf der kanonischen Basis überein, so sind sie gleich. Insbesondere gibt es höchstens eine Determinantenform auf dem $\mathbb{K}^n$.

Beweis. Wir zeigen, dass für jede alternierende n-Form H auf $\mathbb{K}^n$ gilt:

$$H(e_1, \dots, e_n) = 0 \Rightarrow H = 0\,.$$

Daraus folgt dann

$$F(e_1, \dots, e_n) = G(e_1, \dots, e_n) \Rightarrow F = G\,,$$

weil $H = F - G$ selbst wieder eine alternierende n-Form ist.

Sei also $H(e_1, \dots, e_n) = 0$. Dann ist auch $H(e_{\pi(1)}, \dots, e_{\pi(n)}) = 0$ für jede Permutation

$$\pi : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}\,,$$

da jede Permutation Komposition von Vertauschungen ist.

Wegen der Multilinerarität ist aber dann für beliebige $a_1, \dots, a_n \in \mathbb{K}^n$

$$\begin{aligned} H(a_1, \dots, a_n) &= H(\sum_{j=1}^n a_{j1}e_j, a_2, \dots, a_n) = \sum_{j=1}^n a_{j1}\, H(e_j, a_2, \dots, a_n) = \ldots = \\ &= \sum_{j=1}^n \ldots \sum_{k=1}^n a_{j1} \ldots a_{kn}\, H(e_j, \dots, e_k) = 0\,. \end{aligned}$$

□

Es bleibt also die Existenz einer Determinantenform $\det_n$ auf $\mathbb{K}^n$ zu zeigen.

Dazu verwenden wir die Entwicklung nach Zeilen und zeigen per Induktion, dass diese tatsächlich eine Determinantenform liefert.

Induktionsanfang: Für $n = 2$ liefert $\det_2(A) = a_{11}a_{22} - a_{12}a_{21}$ eine Determinantenform in den Spalten von A.

Induktionsschritt: Sei $\det_{n-1}$ eine Determinantenform auf $\mathbb{R}^{n-1}$ und sei $i \in \{1, \dots, n\}$ beliebig aber fest. Für $A \in M(n \times n, \mathbb{K})$ sei nun

$$\det_n(A) := \sum_{j=1}^n (-1)^{i+j}\; a_{ij}\; \det_{n-1}(A_{ij})\,. \tag{4.1}$$

Es ist zu zeigen, dass $\det_n$ eine Determinantenform in den Spalten von A ist:

(a) **$\det_n$ ist multilinear,** denn $a_{ij}\det_{n-1}(A_{ij})$ ist linear in jeder Spalte von A:

in der j-ten Spalte, da $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} \mapsto a_{ij}$ linear ist,

in den anderen Spalten, da $\det_{n-1}A_{ij}$ nach Induktionsannahme eine Multilinearform ist.

(b) **$\det_n$ ist alternierend**: nach Satz 4.3 müssen wir nur zeigen, dass $\det_n(A) = 0$ wenn zwei benachbarte Spalten gleich sind.

Sei etwa $\vec{a}_k = \vec{a}_{k+1}$. Für $j \neq k$ und $j \neq k+1$ hat A_{ij} dann zwei gleiche Spalten, weswegen $\det_{n-1}A_{ij} = 0$ gilt. Damit bleibt in (4.1) nur noch

$$\det_n(A) = (-1)^{i+k}\, a_{ik} \det_{n-1}(A_{ik}) \,+\, (-1)^{i+k+1}\, a_{i,k+1} \det_{n-1}(A_{i,k+1})\,.$$

Wegen $\vec{a}_k = \vec{a}_{k+1}$ ist aber $a_{ik} = a_{i,k+1}$ und $A_{ik} = A_{i,k+1}$ und somit $\det_n|A| = 0\,$.

(c) **$\det_n$ ist normiert**, da

$$\begin{aligned} \det_n(E_n) &= \sum_{j=1}^{n} (-1)^{i+j}\ \delta_{ij}\ \det_{n-1}(E_{ij}) \\ &= (-1)^{2i}\det_{n-1}(E_{ii}) = \det_{n-1}(E_{n-1}) = 1\,. \end{aligned}$$

Also gibt es für jedes $n \geq 2$ eine Determinantenform auf $\mathbb{K}^n$, welche nach Satz 4.6 eindeutig ist.

Insbesondere liefert die Entwicklung nach der i-ten Zeile für jedes $i \in \{1, \ldots, n\}$ das gleiche Ergebnis.

Die Behauptung zur Entwicklung nach Spalten folgt aus $\det A = \det A^T$, der Aussage des nächsten Satzes.

4.7 Satz. Determinante der transponierten Matrix

Sei $A^T = (a_{ki})$ die zu $A = (a_{ik}) \in M(n, \mathbb{K})$ transponierte Matrix, also

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ a_{12} & a_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{1n} & \ldots & \ldots & a_{nn} \end{pmatrix} \quad \text{für} \quad A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{n1} & \ldots & \ldots & a_{nn} \end{pmatrix}.$$

Dann gilt $|A^T| = |A|$.

Beweis. Für $n = 2$ gilt $\begin{vmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$.

Sei also $n \geq 3$. Es gilt $|A^T| = \det(z_1, \ldots, z_n)$, wobei $z_1, \ldots, z_n$ die Spalten von A^T bzw. die Zeilenvektoren von A sind.

Sei andererseits $D(z_1, \ldots, z_n) := |A|$ aufgefasst als Funktion der Zeilen von A.

Wir werden zeigen, dass D eine Determinantenform ist. Wegen der Eindeutigkeit gilt dann aber $D = \det$ und somit $|A| = |A^T|$.

(a) **D ist linear im i-ten Argument:** Entwicklung von $|A|$ nach der i-ter Zeile liefert

$$D(z_1, \ldots, z_n) = |A| = \sum_{j=1}^{n} (-1)^{i+j}\, a_{ij}\, |A_{ij}|\,.$$

Jeder Summand ist linear in z_i, da z_i in $|A_{ij}|$ nicht vorkommt und $z_i = \begin{pmatrix} a_{i1} \\ \vdots \\ a_{in} \end{pmatrix} \mapsto a_{ij}$ linear ist.

(b) **D ist alternierend**: Sei $z_k = z_{k+1}$. In der Entwicklung nach der i-ten Zeile für $i \neq k$ und $i \neq k+1$ $(n \geq 3)$ ist jeder Term gleich Null, da A_{ij} für jedes j zwei gleiche Zeilen hat. Damit ist $|A_{ij}| = 0$, da A_{ij} nicht vollen Rang hat.

(c) $D(e_1, \ldots, e_n) = |E_n| = 1$.

□

Nun können wir noch den Teil von Satz 4.4 über die Entwicklung nach Spalten beweisen:

$$|A| = |A^T| \underset{\substack{\uparrow \\ \text{Entwicklung nach } k\text{-ter Zeile}}}{=} \sum_{j=1}^{n} (-1)^{k+j} \, a^T_{kj} \, |(A^T)_{kj}| = \sum_{j=1}^{n} (-1)^{k+j} \, a_{jk} \, |A_{jk}| \, .$$

4.8 Satz. **Multiplikationssatz für Determinanten**
Für $A, B \in M(n, \mathbb{K})$ gilt

$$|AB| = |A| \cdot |B| \, .$$

Beweis. Sind $b_1, \ldots, b_n$ die Spalten von B, so hat AB die Spalten $Ab_1, \ldots, Ab_n$.
Für festes A und beliebiges B betrachten wir

$$F(b_1, \ldots, b_n) = \det(Ab_1, \ldots, Ab_n) = |AB| \, .$$

Als Komposition der linearen Abbildungen $b_i \mapsto Ab_i$ und det ist F multilinear. F ist alternierend, da det alternierend ist und es gilt

$$F(e_1, \ldots, e_n) = \det(a_1, \ldots, a_n) = |A|$$

Auch durch $G(b_1, \ldots, b_n) = |A| \det(b_1, \ldots, b_n)$ ist eine alternierende n-Form mit $G(e_1, \ldots, e_n) = |A|$ gegeben.
Falls $|A| = 0$ ist, so gilt $G = F$. Falls $|A| \neq 0$ ist, so sind $\frac{G}{|A|}$ und $\frac{F}{|A|}$ Determinantenformen und wegen der Eindeutigkeit gleich. Also $G = F$, d.h. $|A| \cdot |B| = |AB|$. □

4.9 Korollar. Eine Matrix $A \in M(n, \mathbb{K})$ ist genau dann invertierbar, wenn $|A| \neq 0$ ist. Dann gilt

$$|A^{-1}| = \frac{1}{|A|} \, .$$

Beweis. Falls A invertierbar ist, so folgt aus dem Multiplikationssatz

$$1 = |E| = |A \cdot A^{-1}| = |A||A^{-1}| \, ,$$

also

$$|A| \neq 0 \quad \text{und} \quad |A^{-1}| = \frac{1}{|A|} \, .$$

Ist A nicht invertierbar, so ist A nach Satz 2.31 nicht surjektiv. Also sind die Spalten von A linear abhängig und es gilt nach Satz 4.2, dass $|A| = 0$ ist. □

4.10 Zusammenfassung. Für eine $n \times n$ Matrix A sind also äquivalent

(a) $|A| \neq 0$

(b) A ist invertierbar.

(c) Das Gleichungssystem $Ax = b$ ist universell und eindeutig lösbar.

(d) Die Spalten von A sind linear unabhängig.

(e) Die Zeilen von A sind linear unabhängig.

(f) $\mathrm{Rang} A = n$

(g) $A : \mathbb{K}^n \to \mathbb{K}^n$ ist injektiv.

(h) $A : \mathbb{K}^n \to \mathbb{K}^n$ ist surjektiv.

4.11 Satz. und Definition. **Die Determinante eines Endomorphismus**
Sei V ein Vektorraum der Dimension $n < \infty$, $L \in \mathcal{L}(V)$ und $\mathcal{B}$ eine Basis von V.

Die Zahl $|M_{\mathcal{B}}(L)|$ hängt nicht von der Basis $\mathcal{B}$ ab und wird mit $\det L$ bezeichnet.

Beweis. Ähnliche Matrizen haben gemäß des Multiplikationssatzes gleiche Determinanten,

$$|S^{-1}AS| = |S^{-1}|\,|A|\,|S| = |A|\,.$$

Da für zwei Basen $\mathcal{A}$ und $\mathcal{B}$ gemäß Satz 2.37 die Matrizen $M_{\mathcal{A}}(L)$ und $M_{\mathcal{B}}(L)$ ähnlich sind, folgt die Aussage. □

4.12 Satz. **Die Cramersche Regel**
Für eine invertierbare $n \times n$-Matrix $A = (a_1, \dots, a_n)$ über $\mathbb{K}$ und $b \in \mathbb{K}^n$ ist die eindeutige Lösung der Gleichung $Ax = b$ gegeben durch

$$x_i = \frac{1}{|A|} \det(a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n) \quad \text{für} \quad i = 1, \dots, n\,.$$

Beweis. Wir schreiben die Gleichung $Ax = b$ als

$$x_1 a_1 + \cdots + x_n a_n = b$$

bzw.

$$x_1 a_1 + \cdots + (x_i a_i - b) + \cdots + x_n a_n = 0\,.$$

Also sind die Spalten der Matrix

$$\begin{pmatrix} a_{11} & \cdots & (x_i a_{1i} - b_1) & \cdots & a_{1n} \\ \vdots & & & & \vdots \\ a_{n1} & \cdots & (x_i a_{in} - b_n) & \cdots & a_{nn} \end{pmatrix}$$

linear abhängig und ihre Determinante somit Null. Linearität in der i-ten Spalte liefert schließlich

$$x_i \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} - \begin{vmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ \vdots & & & & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix} = 0\,.$$

□

4.13 Bemerkung. Die Cramersche Regel ist nicht zur expliziten Berechnung von Lösungen geeignet, zeigt aber z.B. die stetige Abhängigkeit der Lösungen von A und b.

Determinante und Volumen im $\mathbb{R}^n$

Seien $\vec{a}_1, \dots, \vec{a}_n$ Vektoren im $\mathbb{R}^n$, dann ist das von ihnen aufgespannte **Parallelotop** die Menge

$$P(\vec{a}_1, \dots, \vec{a}_n) := \left\{ \sum_{j=1}^{n} \alpha_j \vec{a}_j \mid 0 \leq \alpha_j \leq 1 \text{ für } j = 1, \dots, n \right\}.$$

Für $n = 1$ ist $P(\vec{a}_1)$ eine Strecke, für $n = 2$ ist $P(\vec{a}_1, \vec{a}_2)$ ein Parallelogramm und für $n = 3$ ist $P(\vec{a}_1, \vec{a}_2, \vec{a}_3)$ ein Spat. Der n-dimensionale Einheitswürfel ist $P(\vec{e}_1, \dots, \vec{e}_n)$.
Wir haben uns bereits überlegt, dass

$$\text{Fläche}(P(\vec{a}_1, \vec{a}_2)) = |\det(\vec{a}_1, \vec{a}_2)|$$

und

$$\text{Volumen}(P(\vec{a}_1, \vec{a}_2, \vec{a}_3)) = |\det(\vec{a}_1, \vec{a}_2, \vec{a}_3)|$$

gilt. Dieser Zusammenhang besteht auch für $n \geq 4$. Da wir das "Volumen" von Teilmengen des $\mathbb{R}^n$ noch nicht definiert haben, erheben wir diese Setzung zur Definition. Das müssen wir natürlich begründen. Dazu stellen wir fest, dass wir folgende Eigenschaften a priori von jeder vernünftigen Definition des n-dimensionalen Volumens

$$\text{Vol}(P(a_1, \dots, a_n)) =: V(a_1, \dots, a_n)$$

fordern müssen:

(a) **Homogenität in jeder Richtung**

$$V(\dots, a_{i-1}, \alpha a_i, a_{i+1}, \dots) = |\alpha| V(a_1, \dots, a_n)$$

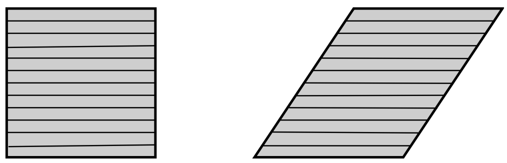
für jedes $i = 1, \dots, n$ und $\alpha \in \mathbb{R}$.

(b) **Cavalierisches Prinzip**

$$V(\dots, a_{i-1}, a_i + \alpha a_k, a_{i+1}, \dots) = V(a_1, \dots, a_n)$$

für $i = 1, \dots, n$ und $k \neq i$.

"Ein Papierstapel ändert sein Volumen nicht, wenn man die Blätter gegeneinander verschiebt":



(c) **Der Einheitswürfel hat Volumen** 1

$$V(e_1, \dots, e_n) = 1.$$

4.14 Satz. **Eindeutigkeit des "Volumens"**
Es gibt genau eine Funktion $V : \mathbb{R}^n \times \dots \times \mathbb{R}^n \to \mathbb{R}$ welche die drei Eigenschaften (a), (b) und (c) hat. Sie ist gegeben durch

$$V(a_1, \dots, a_n) = |\det(a_1, \dots, a_n)|$$

und heißt Volumen des Parallelotops $P(a_1, \dots, a_n)$.

Beweis. $|\det(a_1, \dots, a_n)|$ hat offensichtlich die Eigenschaften (a), (b) und (c). Erfülle nun auch V (a), (b) und (c) und sei

$$F(a_1, \dots, a_n) := \frac{V(a_1, \dots, a_n)}{|\det(a_1, \dots, a_n)|} \det(a_1, \dots, a_n)$$

falls $a_1, \ldots, a_n$ linear unabhängig und $F(a_1, \ldots, a_n) = 0$ sonst. Wir zeigen, dass F eine Determinantenform ist. Dann folgt aus der Eindeutigkeit $F = \det$ und somit die Aussage des Satzes.

Wir zeigen zunächst, dass F multilinear ist. Dazu setzen wir $a_1 = \alpha x + \beta y$ und weisen o.B.d.A. die Linearität im ersten Argument nach.

(i) Falls $a_2, \ldots, a_n$ linear abhängig sind, gilt

$$F(\alpha x + \beta y, a_2, \ldots, a_n) = 0 = 0 + 0 = \alpha F(x, a_2, \ldots, a_n) + \beta F(y, a_2, \ldots, a_n)\,.$$

(ii) Seien also $a_2, \ldots, a_n$ linear unabhängig. Falls $\beta y = \sum_{j=2}^n t_j a_j$, dann ist

$$\begin{aligned} F(\alpha x + \beta y, \ldots) &= \frac{V(\alpha x + \beta y, \ldots)}{|\det(\alpha x + \beta y, \ldots)|}\det(\alpha x + \beta y, \ldots) \\ &= \frac{V(\alpha x, \ldots)}{|\det(\alpha x, \ldots)|}\det(\alpha x, \ldots) \\ &= \frac{|\alpha| V(x, \ldots)}{|\alpha||\det(x, \ldots)|}\,\alpha\det(\alpha x, \ldots) \\ &= \alpha F(x, \ldots) = \alpha F(x, \ldots) + \beta \underbrace{F(y, \ldots)}= 0\,. \end{aligned}$$

Dabei folgt $V(\alpha x + \beta y, \ldots) = V(\alpha x, \ldots)$ aus dem Cavalierischen Prinzip.

(iii) Seien schließlich $\beta y, a_2, \ldots, a_n$ linear unabhängig und $\alpha x = t_1 \beta y + \sum_{j=2}^n t_j a_j$, dann ist

$$\begin{aligned} F(\alpha x + \beta y, \ldots) &= F(\,(1 + t_1)\beta y, \ldots) = (1 + t_1)\beta F(y, \ldots) = \beta F(y, \ldots) + F(t_1 \beta y, \ldots) \\ &= \beta F(y, \ldots) + F(\alpha x, \ldots) = \beta F(y, \ldots) + \alpha F(x, \ldots)\,, \end{aligned}$$

wobei wir im ersten und im vorletzten Schritt wieder das Cavalierische Prinzip verwendet haben.

Also ist F multilinear. Da $F(\ldots, a, a, \ldots) = 0$ per Definition gilt, ist F alternierend und wegen $V(e_1, \ldots, e_n) = 1$ ist F eine Determinantenform. $\square$

5 Eigenwerte und Eigenvektoren

5.1 Definition. **Eigenwerte und Eigenvektoren**

Sei V ein Vektorraum über $\mathbb{K}$ und $L : V \to V$ ein Endomorphismus. Eine Zahl $\lambda \in \mathbb{K}$ heißt **Eigenwert** von L, falls es einen Vektor $v \neq 0$ aus V mit

$$Lv = \lambda v$$

gibt. Ein solcher Vektor heißt dann **Eigenvektor** von L zum Eigenwert λ.

5.2 Lemma. Ist $A = M_{\mathcal{B}}(L)$ die Matrix von $L \in \mathcal{L}(V)$ bezüglich einer Basis $\mathcal{B} = (v_1, \ldots, v_n)$ von V, so hat A genau dann **Diagonalgestalt**,

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

wenn für alle $j = 1, \ldots, n$ der Basisvektor v_j Eigenvektor von L zum Eigenwert λ_j ist.

Beweis. Es gilt $Lv_j = \lambda_j v_j \Leftrightarrow M_{\mathcal{B}}(L)e_j = \lambda_j e_j \Leftrightarrow$ die j-te Spalte von $M_{\mathcal{B}}(L)$ ist gleich $\lambda_j e_j$. $\square$

5.3 Definition. **Diagonalisierbarkeit**

Ein Endomorphismus für den eine Basis aus Eigenvektoren existiert heißt diagonalisierbar.

5.4 Beispiel. Sei $V = \mathbb{R}^2$, also $\mathbb{K} = \mathbb{R}$.

(a) **Spiegelung L an der "y-Achse":** Es ist $Le_1 = -e_1$ und $Le_2 = e_2$. Somit hat L die Eigenwerte $\lambda_1 = -1$ und $\lambda_2 = 1$ zu den Eigenvektoren e_1 und e_2. In der Basis aus Eigenvektoren, also der kanonischen Basis, hat L die Diagonalgestalt

$$M_{\mathcal{K}}(L) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

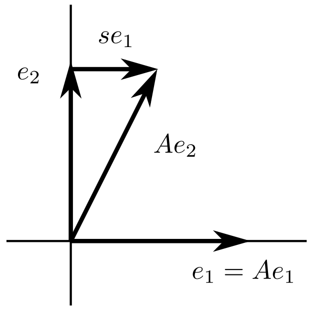
(b) **Drehung D_φ um den Winkel φ um den Ursprung:** Für $\varphi \notin \pi\mathbb{Z}$ hat D_φ hat keinen Eigenvektor, da jeder Vektor unter D_φ seine Richtung ändert. Also hat D_φ dann auch keinen (reellen) Eigenwert und ist nicht diagonalisierbar.

(c) **Scherung:**

Eine Scherung des $\mathbb{R}^2$ ist gegeben durch

$$A = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \quad \text{mit} \quad s \neq 0\,.$$

A hat nur den Eigenvektor e_1 zum Eigenwert $\lambda = 1$, also $Ae_1 = e_1$. Auch A ist somit nicht diagonalisierbar.



5.5 Bemerkung. Ein Vektor $v \in V \setminus \{0\}$ ist genau dann Eigenvektor von $L \in \mathcal{L}(V)$ zum Eigenwert $\lambda \in \mathbb{K}$, wenn

$$v \in \operatorname{Kern}(L - \lambda \mathbb{1}_V)$$

gilt. Denn $v \in \operatorname{Kern}(L - \lambda \mathbb{1}_V) \Leftrightarrow (L - \lambda \mathbb{1}_V)v = 0 \Leftrightarrow Lv = \lambda v$.

Also ist $\lambda \in \mathbb{K}$ genau dann Eigenwert von $L \in \mathcal{L}(V)$, wenn $\operatorname{Kern}(L - \lambda \mathbb{1}_V) \neq \{0\}$.

5.6 Definition. Eigenraum und geometrische Vielfachheit
Sei λ ein Eigenwert von $L \in \mathcal{L}(V)$, dann heißt der Unterraum

$$E_\lambda := \mathrm{Kern}(L - \lambda \mathbb{1}_V)$$

von V der **Eigenraum** von L zum Eigenwert λ. Die Dimension $\dim E_\lambda$ heißt **geometrische Vielfachheit** des Eigenwertes λ.

5.7 Bemerkung. Jeder Vektor $v \in E_\lambda$ mit $v \neq 0$ ist Eigenvektor zum Eigenwert λ. Bei der Wahl von Eigenvektoren besteht also immer eine Freiheit. Das geometrisch relevante Objekt ist der Eigenraum. Eigenräume zu verschiedenen Eigenwerten haben nur die Null gemeinsam, denn es kann ja nicht $Lv = \lambda v = \mu v$ für $\lambda \neq \mu$ gelten. Es gilt aber mehr:

5.8 Lemma. Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig
Sei $L \in \mathcal{L}(V)$ und seien $v_1, \ldots, v_\ell$ Eigenvektoren zu verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_\ell$, also $\lambda_i \neq \lambda_j$ für $i \neq j$, dann sind $v_1, \ldots, v_\ell$ linear unabhängig.

Beweis. **Induktionsanfang:** Für $\ell = 1$ ist v_1 linear unabhängig, da nach Definition $v_1 \neq 0$ gilt.
Induktionsschritt: Die Aussage sei richtig für ℓ. Seien nun $v_1, \ldots, v_{\ell+1}$ Eigenvektoren zu verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_{\ell+1}$ und sei

$$\alpha_1 v_1 + \cdots + \alpha_{\ell+1} v_{\ell+1} = 0\,.$$

Daraus erhält man durch Anwendung von L bzw. $\lambda_{\ell+1}\mathbb{1}_V$ die Gleichungen

$$\begin{aligned} \alpha_1 \lambda_1 v_1 + \cdots + \alpha_{\ell+1}\, \lambda_{\ell+1}\, v_{\ell+1} &= 0 \\ \alpha_1 \lambda_{\ell+1} v_1 + \cdots + \alpha_{\ell+1}\, \lambda_{\ell+1}\, v_{\ell+1} &= 0 \end{aligned}$$

und durch Subtraktion

$$\alpha_1(\lambda_1 - \lambda_{\ell+1})v_1 + \ldots + \alpha_\ell(\lambda_\ell - \lambda_{\ell+1})v_\ell = 0\,.$$

Nach Induktionsannahme sind $v_1, \ldots, v_\ell$ aber linear unabhängig, somit ist

$$\alpha_1(\lambda_1 - \lambda_{\ell+1}) = \cdots = \alpha_\ell(\lambda_\ell - \lambda_{\ell+1}) = 0\,.$$

Wegen $\lambda_i \neq \lambda_j$ für $i \neq j$ folgt $\alpha_1 = \cdots = \alpha_\ell = 0$, damit $\alpha_{\ell+1} v_{\ell+1} = 0$, also $\alpha_{\ell+1} = 0$ und schließlich die lineare Unabhängigkeit von $v_1, \ldots, v_{\ell+1}$. □

5.9 Bemerkung. Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig. Damit kann ein Endormophismus L auf einem n-dimensionalen Raum höchstens n verschiedene Eigenwerte haben. In diesem Fall bilden die Eigenvektoren dann eine Basis und L ist diagonalisierbar. Das Vorliegen von n verschiedenen Eigenwerten ist allerdings keine notwendige Bedingung für Diagonalisierbarkeit, beispielsweise hat die Einheitsmatrix nur den einen Eigenwert $\lambda = 1$ und die Nullmatrix nur den Eigenwert $\lambda = 0$. Beide sind aber diagonal.

5.10 Korollar. Sei $\dim V = n$ und $L \in \mathcal{L}(V)$. Seien $\lambda_1, \ldots, \lambda_\ell$ verschiedene Eigenwerte von L mit geometrischen Vielfachheiten $n_1, \ldots, n_\ell$ und $(v_1^{(i)}, \ldots, v_{n_i}^{(i)})$ jeweils eine Basis des Eigenraums E_{λ_i} zum Eigenwert λ_i.
Dann sind auch

$$v_1^{(1)}, \ldots, v_{n_1}^{(1)}, v_1^{(2)}, \ldots, v_{n_2}^{(2)}, \ldots, v_1^{(\ell)}, \ldots, v_{n_\ell}^{(\ell)}$$

linear unabhängig.

Insbesondere ist die Summe $n_1 + \cdots + n_\ell$ der geometrischen Vielfachheiten höchstens n und L ist genau dann diagonalisierbar, wenn $n_1 + \cdots + n_\ell = n$ ist.

Beweis. Ist $\sum_{i=1}^{\ell}\sum_{k=1}^{n_i}\alpha_k^{(i)}\,v_k^{(i)}=0$, so sind nach Lemma 5.8 die Vektoren $\sum_{k=1}^{n_i}\alpha_k^{(i)}\,v_k^{(i)}\in E_{\lambda_i}$ gleich Null und wegen der linearen Unabhängigkeit der $v_1^{(i)},\dots,v_{n_i}^{(i)}$ verschwinden alle Koeffizienten $\alpha_k^{(i)}$.
Also ist $(v_1^{(1)},\dots,v_{n_1}^{(1)},\dots,v_1^{(\ell)},\dots v_{n_\ell}^{(\ell)})$ ein linear unabhängiges $n_1+\ldots+n_\ell$-Tupel von Eigenvektoren.
Im Fall $n_1+\ldots+n_\ell=n$ ist es eine Basis aus Eigenvektoren und L somit diagonalisierbar.
Ist nun andererseits L als diagonalisierbar vorausgesetzt und sei m_i die Anzahl der Eigenvektoren zum Eigenwert λ_i in einer Basis aus Eigenvektoren, so ist $m_i\le \dim E_{\lambda_i}=n_i$, also

$$n=m_1+\ldots+m_\ell\le n_1+\ldots+n_\ell\le n\,,$$

woraus $n_1+\ldots+n_\ell=n$ und $m_j=n_j$ für alle $j=1,\dots,\ell$ folgt. □

Wie bereits bemerkt, ist $\lambda\in\mathbb{K}$ genau dann Eigenwert von $L\in\mathcal{L}(V)$, wenn $\mathrm{Kern}(L-\lambda 1\!\!1_V)\ne\{0\}$, also wenn $L-\lambda 1\!\!1_V$ nicht injektiv ist. Daraus ergibt sich das

5.11 Lemma. Sei $\dim V=n<\infty$ und $L\in\mathcal{L}(V)$. Dann ist $\lambda\in\mathbb{K}$ genau dann ein Eigenwert von L, wenn $\det(L-\lambda 1\!\!1_V)=0$ gilt.

Wird L bezüglich einer Basis von V durch die Matrix $A=(a_{ij})$ beschrieben, so ist

$$\det(L-\lambda 1\!\!1_V)=\det\begin{pmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22}-\lambda & \cdots & a_{2n}\\ \vdots & & \ddots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda\end{pmatrix}.$$

Diese von λ abhängige Determinante ist ein Polynom vom Grade n in λ und heißt das charakterische Polynom von L.

5.12 Lemma. Sei $\dim V=n<\infty$ und $L\in\mathcal{L}(V)$. Dann gibt es Zahlen $a_0,\dots,a_n\in\mathbb{K}$ mit

$$P_L(\lambda):=\det(L-\lambda 1\!\!1_V)=a_n\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_1\lambda+a_0$$

für alle $\lambda\in\mathbb{K}$. Es gilt

$$a_n=(-1)^n,\quad a_{n-1}=(-1)^{n-1}\mathrm{Spur}L:=(-1)^{n-1}\sum_{j=1}^{n}a_{jj}\quad\text{und}\quad a_0=\det L\,.$$

Man nennt P_L das **charakteristische Polynom** von L.

Beweis. Sei $A=M_{\mathcal{B}}(L)$ bezüglich einer Basis $\mathcal{B}$ von V. Wir bestimmen

$$\det(L-\lambda 1\!\!1_V)=\det(A-\lambda E_n)\,.$$

Aus

$$\det(A-\lambda B)=\sum_{j=1}^{n}(-1)^{i+j}(a_{ij}-\lambda b_{ij})\det\left((A-\lambda B)_{ij}\right)$$

sieht man sofort durch Induktion, dass für beliebige $n\times n$-Matrizen A und B $\det(A-\lambda B)=c_n\lambda^n+\cdots+c_1\lambda+c_0$ für geeignete $c_0,\dots,c_n\in\mathbb{K}$ ist.

Für $B=E_n$ ergibt Induktion und Entwicklung nach der ersten Zeile

$$\det(A-\lambda E_n)=(a_{11}-\lambda)\det\left((A-\lambda E_n)_{11}\right)\;+\;R(\lambda)\,.$$

Es ist $R(\lambda)$ ein Polynom vom Grad $n-2$, da $\det(A-\lambda E_n)_{1j}$ ein Polynom vom Grad $n-2$ ist. Nun ist nach Induktionsannahme

$$\det((A-\lambda E_n)_{11}) = (-1)^{n-1}\lambda^{n-1} + (-1)^{n-2}\sum_{j=2}^{n} a_{jj}\,\lambda^{n-2} + \cdots$$

und somit

$$\det(A-\lambda E_n) = (-1)^n\lambda^n + (-1)^{n-1}\sum_{j=1}^{n} a_{jj}\,\lambda^{n-1} + \cdots .$$

Der konstante Term ergibt sich aus

$$a_0 = P_L(0) = \det(A - 0\cdot E_n) = \det A\,.$$

□

5.13 Korollar. Die Eigenwerte sind die Nullstellen des charakteristischen Polynoms.

5.14 Beispiel. (a) **Spiegelung an der y-Achse:**

$$L = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow P_L(\lambda) = \det(L-\lambda E_2) = \det\begin{pmatrix} -1-\lambda & 0 \\ 0 & 1-\lambda \end{pmatrix} = (\lambda+1)(\lambda-1)\,.$$

Die Nullstellen und somit die Eigenwerte sind $\lambda_1 = 1,\ \lambda_2 = -1$.

(b) **Drehung um Winkel φ in der Ebene:**

$$D_\varphi = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & cos\varphi \end{pmatrix}$$

Also ist

$$P_{D_\varphi}(\lambda) = \det\begin{pmatrix} \cos\varphi-\lambda & -\sin\varphi \\ \sin\varphi & \cos\varphi-\lambda \end{pmatrix} = (\cos\varphi-\lambda)^2 + \sin^2\varphi > 0 \quad \text{falls} \quad \varphi \notin \pi\mathbb{Z}\,.$$

Für $\varphi = 0$ ist $D_\varphi = E_2$ und $P_{E_2}(\lambda) = (1-\lambda)^2$ hat die doppelte Nullstelle $\lambda_1 = 1$. Also hat D_0 den Eigenwert $\lambda_1 = 1$. Seine geometrische Vielfachheit ist 2, da der Eigenraum $E_{\lambda_1} = \mathbb{R}^2$ ist. Für $\varphi = \pi$ ist $D_\varphi = -E_2$, der einzige Eigenwert also $\lambda = -1$. Für alle anderen $\varphi \in (0, 2\pi)$ hat $P_{D_\varphi}(\lambda)$ keine Nullstellen, und somit hat D_φ dann auch keine Eigenwerte.

(c) **Scherung:**

$$A = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \quad\Rightarrow\quad P_A(\lambda) = \det\begin{pmatrix} 1-\lambda & s \\ 0 & 1-\lambda \end{pmatrix} = (1-\lambda)^2\,.$$

$P_A(\lambda)$ hat also die doppelte Nullstelle $\lambda_1 = 1$, der Eigenraum zum Eigenwert λ_1 ist aber nur eindimensional, $E_{\lambda_1} = \text{Span}\,(e_1)$. Die geometrische Vielfachheit ist hier also echt kleiner als die Ordnung der Nullstelle (=algebraische Vielfachheit).

Zur Erinnerung: Der Fundamentalsatz der Algebra

Jedes komplexe Polynom vom Grade $n \geq 1$, d.h. jede Abbildung $P : \mathbb{C} \to \mathbb{C}$ von der Form

$$P(z) = c_n z^n + \cdots + c_1 z + c_0\,,$$

wobei $c_0, \ldots, c_n \in \mathbb{C}$ und $c_n \neq 0$, hat mindestens eine Nullstelle.

5.15 Korollar. Für $n \geq 1$ hat jeder Endomorphismus eines n-dimensionalen komplexen Vektorraumes mindestens einen Eigenwert.

In der Vorlesung Mathe für Physiker 1 wurde aus dem Fundamentalsatz gefolgert:

5.16 Lemma. Jedes komplexe Polynom P zerfällt in Linearfaktoren: ist $P(z) = c_n\, z^n + \cdots + c_1 z + c_0$ mit $c_n \neq 0$, und sind $z_1, \cdots, z_r \in \mathbb{C}$ die paarweise verschiedenen Nullstellen von P, so gilt

$$P(z) = c_n \prod_{i=1}^{r} (z - z_i)^{m_i}$$

mit wohlbestimmten Exponenten $m_i \geq 1$, welche man die Vielfachheit der Nullstelle nennt.

5.17 Bemerkung. Für endliche Körper $\mathbb{K}$ gibt es verschiedene Polynome, d.h. Ausdrücke der Form $P(z) = a_n\, z^n + \cdots + a_0$, welche dieselbe (polynomiale) Abbildung $P : \mathbb{K} \to \mathbb{K}$, $z \mapsto P(z)$ definieren. Beispielsweise in $\mathbb{K} = \mathcal{F}_2 = \{0, 1\}$ definieren z und z^2 dieselbe Abbildung, da $0 \cdot 0 = 0$ und $1 \cdot 1 = 1$.

5.18 Definition. **Algebraische Vielfachheit**
Sei $\dim V = n < \infty\,, \mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ und $L \in \mathcal{L}(V)$. Falls λ_0 eine k-fache Nullstelle des charakteristischen Polynoms ist, d.h.

$$P_L(\lambda) = (\lambda - \lambda_0)^k\, Q(\lambda) \quad \text{mit} \quad Q(\lambda_0) \neq 0\,,$$

so heißt k die **algebraische Vielfachheit** des Eigenwertes λ_0 .

5.19 Lemma. Die geometrische Vielfachheit eines Eigenwertes λ ist immer kleiner oder gleich der algebraischen Vielfachheit.

Beweis. Ergänzt man eine Basis des n_i-dimensionalen Eigenraumes $E_{\lambda i}$ zu einer Basis von V, so hat die Matrix von L bezüglich dieser Basis die Form

$$\left(\begin{array}{c|c} \overbrace{\begin{matrix}\lambda_i & & 0\\ & \ddots & \\ 0 & & \lambda_i\end{matrix}}^{n_i} & * \\ \hline 0 & * \end{array}\right)$$

Aus der Leibnizformel (Aufgabe 28) liest man ab, dass in $P_L(\lambda) = \det\left(M(L) - \lambda\, E_n\right)$ der Faktor $(\lambda_i - \lambda)$ mindestens n_i-mal vorkommt. □

5.20 Korollar. Sei V ein komplexer n-dimensionaler Vektorraum und $L \in \mathcal{L}(V)$. Sind $\lambda_1, \ldots, \lambda_\ell$ die Eigenwerte von L mit algebraischen Vielfachheiten $m_1, \ldots, m_\ell$ so gilt $m_1 + \cdots + m_\ell = n$. Insbesondere ist L genau dann diagonalisierbar, wenn für jeden Eigenwert gilt $n_i = \dim E_{\lambda_i} = m_i$, also geometrische Vielfachheit = algebraische Vielfachheit.

5.21 Beispiele. (a) Sei $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$, dann ist

$$\begin{aligned} P_A(\lambda) &= \det\left(A - \lambda\, E_3\right) = \begin{vmatrix} 1-\lambda & 2 & 3 \\ 4 & 5-\lambda & 6 \\ 7 & 8 & 9-\lambda \end{vmatrix} \\ &= (1-\lambda)(5-\lambda)(9-\lambda) + 2 \cdot 6 \cdot 7 + 3 \cdot 4 \cdot 8 - 7 \cdot (5-\lambda) \cdot 3 - 8 \cdot 6 \cdot (1-\lambda) - (9-\lambda) \cdot 4 \cdot 2 \\ &= -\lambda^3 + 15\,\lambda^2 - 18\,\lambda + 0 = -\lambda\,(\lambda^2 - 15\,\lambda + 18) \end{aligned}$$

$$\Rightarrow \lambda_1 = 0\,, \lambda_2 = \frac{15 + 3\sqrt{17}}{2}\,, \lambda_3 = \frac{15 - 3\sqrt{17}}{2}$$

Um die Eigenräume zu bestimmen, müsste man nun die Gleichungssysteme

$$(A - \lambda_1 E_3)v_1 = 0\,, (A - \lambda_2 E_3)v_2 = 0\,, (A - \lambda_3 E_3)v_3 = 0$$

lösen. Da alle Eigenwerte einfach sind (d.h. einfache Nullstellen des charakteristischen Polynoms), bilden die zugehörigen Eigenvektoren eine Basis von $\mathbb{R}^3$ und A ist diagonalisierbar.

(b) Sei $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, dann ist

$$P_A(\lambda) = \det(A - \lambda E_2) = \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 1 = (\lambda+1)(\lambda-1)\,.$$

Die Eigenwerte sind also $\lambda_1 = 1$ und $\lambda_2 = -1$ und die zugehörigen Eigenräume liest man in diesem Fall einfach ab:

$$E_{\lambda_1} = \left\{\alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix} \mid \alpha \in \mathbb{R}\right\} \qquad \text{und} \qquad E_{\lambda_2} = \left\{\alpha \begin{pmatrix} 1 \\ -1 \end{pmatrix} \mid \alpha \in \mathbb{R}\right\}.$$

5.22 Satz. von Cayley-Hamilton

Sei $L \in \mathcal{L}(V)$ ein Endomorphismus des endlichdimensionalen Vektorraums V und P_L sein charakteristisches Polynom. Dann ist

$$P_L(L) = 0\,.$$

Hier steht $P_L(L) \in \mathcal{L}(V)$ für diejenige Abbildung, die man erhält, wenn man L in das charakteristische Polynom einsetzt, also

$$P_L(L) = a_n L^n + a_{n-1}L^{n-1} + \cdots + a_1 L + a_0 E_n\,.$$

Beweis. Für diagonalisierbare L ist die Aussage leicht zu sehen: In einer Basis $\mathcal{B}$ aus Eigenvektoren hat L^m für $m = 1, \ldots, n$ die Form

$$M_{\mathcal{B}}(L^m) = \begin{pmatrix} \lambda_1^m & & 0 \\ & \ddots & \\ 0 & & \lambda_n^m \end{pmatrix}$$

und somit

$$M_{\mathcal{B}}(P_L(L))_{ij} = \delta_{ij}\left(a_n\lambda_j^n + a_{n-1}\lambda_j^{n-1} + \cdots + a_1\lambda_j + a_0\right) = 0\,,$$

da die λ_j ja als Eigenwerte gerade die Nullstellen des charakteristischen Polynoms sind. Für nicht diagonalisierbare Endomorphismen ist der Beweis mit den uns im Moment zur Verfügung stehenden Mitteln nicht leicht und wir kommen erst später darauf zurück, vgl. Bemerkung 8.17. □

Anwendung: System linearer Differentialgleichungen 1. Ordnung mit konstanten Koeffizienten.

Die Differentialgleichungen

$$\begin{aligned} \dot{x}_1(t) &= a_{11}x_1(t) + \cdots + a_{1n}x_n(t) \\ &\vdots \\ \dot{x}_n(t) &= a_{n1}x_1(t) + \cdots + a_{nn}x_n(t) \end{aligned}$$

sind gekoppelt und in Kurzform schreiben wir

$$\dot{x}(t) = Ax(t)\,.$$

Es sei die Anfangsbedingung $x(0) = a \in \mathbb{K}^n$ gegeben und wir suchen die Lösung für alle anderen Zeiten, also eine Funktion $x : \mathbb{R} \to \mathbb{K}^n$ welche $\frac{\mathrm{d}}{\mathrm{d}t}x(t) = Ax(t)$ und $x(0) = a$ erfüllt.
Falls A diagonalisierbar ist, also

$$S^{-1}\,A\,S = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = D$$

für eine reguläre Matrix S gilt, dann ergibt sich für den Vektor $y(t) = S^{-1}x(t)$

$$\dot{y}(t) = S^{-1}\dot{x}(t) = S^{-1}Ax(t) = S^{-1}ASS^{-1}x(t) = S^{-1}ASy(t) = Dy(t)\,.$$

Wir erhalten also die entkoppelten Differentialgleichungen

$$\dot{y}_1(t) = \lambda_1 y_1(t),\ \ldots,\ \dot{y}_n(t) = \lambda_n y_n(t)\,.$$

Die Anfangsdaten für y ergeben sich aus denen für x durch

$$y(0) = S^{-1}x(0) = S^{-1}a =: b\,.$$

Die Lösungen für das entkoppelte System kann man direkt ablesen,

$$y_1(t) = b_1 \mathrm{e}^{\lambda_1 t},\ \ldots,\ y_n(t) = b_n \mathrm{e}^{\lambda_n t}$$

oder kurz

$$y\,(t) = \mathrm{e}^{Dt}\,b := \begin{pmatrix} \mathrm{e}^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & \mathrm{e}^{\lambda_n t} \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Durch Rücktransformation bekommt man auch wieder

$$x(t) = Sy(t) = S\mathrm{e}^{Dt}b = S\mathrm{e}^{Dt}S^{-1}a =: \mathrm{e}^{At}a\,.$$

5.23 Beispiel. <u>Gedämpfter harmonischer Oszillator</u>
Die Differentialgleichung des gedämpften harmonischen Oszillators lautet

$$\ddot{x}(t) = -\omega_0^2 x(t) - 2\gamma\dot{x}(t)\,,$$

wobei ω_0 die Frequenz des ungedämpften Oszillators ist und γ die Stärke der Dämpfung beschreibt. Diese lineare Differentialgleichung zweiter Ordnung ist äquivalent dem System linearer Differentialgleichungen erster Ordnung

$$\begin{pmatrix} \dot{x}(t) \\ \dot{v}(t) \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ -\omega_0^2 & -2\gamma \end{pmatrix}}_{=:A} \begin{pmatrix} x(t) \\ v(t) \end{pmatrix}.$$

Dieses System wollen wir nun durch Diagonalisieren lösen.

(a) **Bestimmen der Eigenwerte**

Das charakteristische Polynom lautet

$$P_A(\lambda) = \begin{vmatrix} -\lambda & 1 \\ -\omega_0^2 & -2\gamma - \lambda \end{vmatrix} = -\lambda(-2\gamma - \lambda) + w_0^2 = \lambda^2 + 2\gamma\lambda + w_0^2$$

und die Eigenwerte sind

$$\lambda_{1/2} = \frac{-2\gamma \pm \sqrt{4\gamma^2 - 4w_0^2}}{2}$$

Je nach Wahl der Parameter können die folgenden drei Fälle auftreten:

(i) $\gamma > \omega_0 \quad \Rightarrow \quad \lambda_{1/2} = -\gamma \pm \sqrt{\gamma^2 - \omega_0^2} \in \mathbb{R}$,
die Eigenwerte sind verschieden und reell.

(ii) $\gamma = \omega_0 \quad \Rightarrow \quad \lambda_{1/2} = -\gamma \in \mathbb{R}$,
es gibt nur einen Eigenwert.

(iii) $\gamma < \omega_0 \quad \Rightarrow \quad \lambda_{1/2} = -\gamma \pm \mathrm{i}\sqrt{\omega_0^2 - \gamma^2} \in \mathbb{C}$,
die Eigenwerte sind verschieden und komplex.

(b) **Bestimmen von Eigenvektoren**

Eigenvektoren zu λ_1 und λ_2 ergeben sich aus der jeweiligen linearen Gleichung

$$(A - \lambda_1 E_2)v^{(1)} = 0 \qquad \text{bzw.} \qquad (A - \lambda_2 E_2)v^{(2)} = 0\,,$$

also

$$\begin{pmatrix} -\lambda_{1/2} & 1 \\ -\omega_0^2 & -2\gamma - \lambda_{1/2} \end{pmatrix} \begin{pmatrix} v_1^{(1)/(2)} \\ v_2^{(1)/(2)} \end{pmatrix} = 0 \qquad \Rightarrow \qquad -\lambda_{1/2}\, v_1^{(1)/(2)} + v_2^{(1)/(2)} = 0\,.$$

Da diese Gleichungen sicherlich jeweils mindestens eine Lösung haben (wir wissen ja bereits, dass $\lambda_{1/2}$ Eigenwerte sind, also muss es auch zugehörige Eigenvektoren geben), können wir z.B. $v_1^{(1)/(2)} = 1$ setzen und erhalten

$$v^{(1)} = \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix} \quad \text{und} \quad v^{(2)} = \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix}.$$

Falls aber $\lambda_1 = \lambda_2 = -\gamma$ (also Fall (ii)), so haben wir nur einen Eigenvektor gefunden. Das ist dann aber auch der einzige, da der Kern von

$$(A - \lambda) = \begin{pmatrix} \gamma & 1 \\ -\gamma^2 & -\gamma \end{pmatrix}$$

nicht zweidimensional ist. Im Fall *(ii)* ist A somit nicht diagonalisierbar und wir müssen uns etwas anderes überlegen (vgl. Kapitel 8 über die Jordansche Normalform).

Sei nun $\lambda_1 \neq \lambda_2$.

(c) **Transformation auf Diagonalgestalt**

Es ist

$$D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = S^{-1}\, A\, S$$

wobei

$$S = \left(v^{(1)}, v^{(2)}\right) = \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix}.$$

Man erinnere sich: die Spalten der Marix sind die Bilder der Basisvektoren. Und S bildet hier von der Basisdarstellung bzgl. der Eigenvektoren in die Basisdarstellung bzgl. der kanonischen Basis ab. Die Inverse ist

$$S^{-1} = \frac{1}{\det S} \begin{pmatrix} \lambda_2 & -1 \\ -\lambda_1 & 1 \end{pmatrix} = \frac{1}{\lambda_2 - \lambda_1} \begin{pmatrix} \lambda_2 & -1 \\ -\lambda_1 & 1 \end{pmatrix}$$

(d) **Lösungsgenerator für die Differentialgleichung**

$$
\begin{aligned}
\mathrm{e}^{At} &= S\,\mathrm{e}^{Dt}\,S^{-1} = \frac{1}{\lambda_2-\lambda_1}\begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix}\begin{pmatrix} \mathrm{e}^{\lambda_1 t} & 0 \\ 0 & \mathrm{e}^{\lambda_2 t} \end{pmatrix}\begin{pmatrix} \lambda_2 & -1 \\ -\lambda_1 & 1 \end{pmatrix} \\
&= \frac{1}{\lambda_2-\lambda_1}\begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix}\begin{pmatrix} \lambda_2\mathrm{e}^{\lambda_1 t} & -\mathrm{e}^{\lambda_1 t} \\ -\lambda_1\mathrm{e}^{\lambda_2 t} & \mathrm{e}^{\lambda_2 t} \end{pmatrix} \\
&= \frac{1}{\lambda_2-\lambda_1}\begin{pmatrix} \lambda_2\mathrm{e}^{\lambda_1 t}-\lambda_1\mathrm{e}^{\lambda_2 t} & \mathrm{e}^{\lambda_2 t}-\mathrm{e}^{\lambda_1 t} \\ \lambda_1\lambda_2\left(\mathrm{e}^{\lambda_1 t}-\mathrm{e}^{\lambda_2 t}\right) & \lambda_2\mathrm{e}^{\lambda_2 t}-\lambda_1\mathrm{e}^{\lambda_1 t} \end{pmatrix}.
\end{aligned}
$$

Mit Hilfe des Lösungsoperators können wir nun für beliebige Anfangsdaten $x(0)$ und $v(0)$ die Lösung der Differentialgleichung angeben,

$$\begin{pmatrix} x(t) \\ v(t) \end{pmatrix} = \mathrm{e}^{At}\begin{pmatrix} x(0) \\ v(0) \end{pmatrix},$$

also

$$x(t) = \frac{1}{\lambda_2-\lambda_1}\left(\left(\lambda_2\mathrm{e}^{\lambda_1 t}-\lambda_1\mathrm{e}^{\lambda_2 t}\right)x(0) + \left(\mathrm{e}^{\lambda_2 t}-\mathrm{e}^{\lambda_1 t}\right)v(0)\right) \tag{5.1}$$

und analog für die Geschwindigkeit.

Fall (i): Beide Eigenwerte sind negativ, es gibt keine Oszillationen mehr (egal wie hoch die Anfangsgeschwindigkeit ist!). Man nennt diesen Fall deshalb den überdämpften Fall.

Fall (ii): Wir werden für diesen Fall (5.1) noch explizit auswerten. Dazu definieren wir

$$
\begin{aligned}
\lambda_1 = -\gamma + \mathrm{i}\sqrt{\omega_0^2-\gamma^2} \quad &=: \quad -\gamma + \mathrm{i}\omega =: \lambda \\
\lambda_2 = -\gamma - \mathrm{i}\sqrt{\omega_0^2-\gamma^2} \quad &= \quad -\gamma - \mathrm{i}\omega = \overline{\lambda}\,.
\end{aligned}
$$

Damit ist $\lambda_2 - \lambda_1 = \overline{\lambda} - \lambda = -2\mathrm{i}\omega$ und

$$
\begin{aligned}
\mathrm{e}^{At} &= -\frac{1}{2\mathrm{i}\omega}\begin{pmatrix} 2\mathrm{i}\,\mathrm{Im}(\overline{\lambda}\,\mathrm{e}^{\lambda t}) & 2\mathrm{i}\,\mathrm{Im}(\mathrm{e}^{\overline{\lambda}t}) \\ |\lambda|^2\,2\mathrm{i}\,\mathrm{Im}(\mathrm{e}^{\lambda t}) & 2\mathrm{i}\,\mathrm{Im}(\overline{\lambda}\,\mathrm{e}^{\overline{\lambda}t}) \end{pmatrix} \\
&= -\frac{\mathrm{e}^{-\gamma t}}{\omega}\begin{pmatrix} -\gamma\sin(\omega t)-\omega\cos(\omega t) & \sin(-\omega t) \\ \omega_0^2\sin(\omega t) & -\gamma\sin(\omega t)-\omega\cos(-\omega t) \end{pmatrix} \\
&= \mathrm{e}^{-\gamma t}\begin{pmatrix} \cos(\omega t)+\frac{\gamma}{\omega}\sin(\omega t) & \frac{1}{\omega}\sin(\omega t) \\ -\frac{\omega_0^2}{\omega}\sin(\omega t) & \cos(\omega t)-\frac{\gamma}{\omega}\sin(\omega t) \end{pmatrix}.
\end{aligned}
$$

Damit lautet die allgemeine Lösung zu den Anfangsdaten $x(0) = x_0$ und $v(0) = \dot{x}(0) = v_0$

$$x(t) = \mathrm{e}^{-\gamma t}\left(\left(\cos(\omega t)+\frac{\gamma}{\omega}\sin(\omega t)\right)x_0 + \frac{1}{\omega}\sin(\omega t)\,v_0\right). \tag{5.2}$$

Entsprechend findet man $v(t)$ aus der zweiten Zeile von e^{At} und man sollte sich zumindest einmal selbst überzeugen, dass das Ergebnis tatsächlich mit der Zeitableitung unserer Formel (5.2) für $x(t)$ übereinstimmt.

Wie schon gesagt, behandeln wir Fall (ii) später. Statt Diagonalform verwendet man dann die Jordansche Normalform, und in die kann man jede Matrix durch ein Ähnlichkeitstransformation bringen.

6 Vektorräume mit Skalarprodukt

6.1 Definition. Skalarprodukt
Sei V ein Vektorraum über $\mathbb{K}$. Eine Abbildung

$$\langle\cdot,\cdot\rangle : V \times V \to \mathbb{K}$$

heißt **Skalarprodukt** oder **inneres Produkt**, falls

(a) $\langle u, u\rangle \geq 0 \quad \forall\, u \in V \quad$ und $\quad \langle u, u\rangle = 0 \Leftrightarrow u = 0$. (positive Definitheit)

(b) $\langle\cdot,\cdot\rangle$ ist linear im zweiten Argument, also

$$\langle u, \alpha v + \beta w\rangle = \alpha\langle u, v\rangle + \beta\langle u, w\rangle$$

für alle $u, v, w \in V$ und $\alpha, \beta \in \mathbb{K}$.

(c) $\langle u, v\rangle = \overline{\langle v, u\rangle}$ für alle $u, v \in V$. (Symmetrie)

Einen Vektorraum V mit Skalarprodukt nennt man Skalarproduktraum oder Prä-Hilbertraum. Im Falle $\mathbb{K} = \mathbb{R}$ spricht man auch von euklidischen Räumen.

6.2 Bemerkung.
- Im Fall $\mathbb{K} = \mathbb{R}$ besagt (c), dass $\langle u, v\rangle = \langle v, u\rangle$. Zusammen mit (b) impliziert das die Linearität auch im ersten Argument. Für $\mathbb{K} = \mathbb{R}$ ist ein Skalarprodukt also eine Bilinearform.
- Im Fall $\mathbb{K} = \mathbb{C}$ ist $\langle\cdot,\cdot\rangle$ antilinear im ersten Argument, denn

$$\langle \alpha u + \beta w, v\rangle = \overline{\langle v, \alpha u + \beta w\rangle} = \overline{\alpha\langle v, u\rangle} + \overline{\beta\langle v, w\rangle} = \overline{\alpha}\langle u, v\rangle + \overline{\beta}\langle w, v\rangle\,.$$

Man spricht dann von einer Sesquilinearform (sesqui $= 1\frac{1}{2}$).

6.3 Beispiele. (a) $\mathbb{R}^n$ mit dem natürlichen Skalarprodukt

$$\langle x, y\rangle_{\mathbb{R}^n} = \sum_{k=1}^{n} x_k y_k$$

(b) $\mathbb{C}^n$ mit dem natürlichen Skalarprodukt

$$\langle x, y\rangle_{\mathbb{C}^n} = \sum_{k=1}^{n} \overline{x_k} y_k$$

6.4 Definition. und Satz. Die quadratsummierbaren Folgen
Sei ℓ^2 der Vektorraum der komplexen Folgen $a = (a_1, a_2, \dots)$ für welche die Reihe $\sum_{k=1}^{\infty} |a_k|^2$ konvergiert.
Für $a = (a_1, a_2, \dots), b = (b_1, b_2, \dots) \in \ell^2$ konvergiert die Reihe

$$\langle a, b\rangle_{\ell^2} := \sum_{k=1}^{\infty} \overline{a_k}\, b_k$$

und definiert ein Skalarprodukt auf ℓ^2.

Beweis. Für $a = (a_1, a_2, \ldots) \in \ell^2$ und $\lambda \in \mathbb{C}$ konvergiert auch $\sum_{k=1}^{\infty} |\lambda a_k|^2 = |\lambda|^2 \sum_{k=1}^{\infty} |a_k|^2$. Für $a, b \in \ell^2$ gilt

$$|a_k + b_k|^2 \leq 2(|a_k|^2 + |b_k|^2)$$

und

$$|\overline{a}_k b_k| \leq \frac{1}{2}(|a_k|^2 + |b_k|^2)$$

und somit (nach dem Majorantenkriterium)

$$\sum |a_k + b_k|^2 \leq 2\sum |a_k|^2 + 2\sum |b_k|^2 < \infty$$

und

$$\sum |\overline{a}_k b_k| \leq \frac{1}{2}\sum |a_k|^2 + \frac{1}{2}\sum |b_k|^2 < \infty \,.$$

Also sind auch λa und $a+b$ quadratsummierbar und der Ausdruck $\langle a, b\rangle_{\ell^2}$ ist endlich für $a, b \in \ell^2$.

Dass $\langle\cdot,\cdot\rangle$ ein Skalarprodukt ist, rechnet man einfach nach:

(a) $\langle a, a\rangle = \sum \overline{a}_k a_k = \sum |a_k|^2 \geq 0 \quad$ und $\quad \sum |a_k|^2 = 0 \;\Leftrightarrow\; a_k = 0 \;\forall\, k \;\Leftrightarrow\; a = 0\,.$

(b) $\langle a, \lambda b + c\rangle = \sum \overline{a}_k(\lambda b_k + c_k) = \lambda \sum \overline{a}_k b_k + \sum \overline{a}_k c_k = \lambda\langle a, b\rangle + \langle a, c\rangle$.

(c) $\langle a, b\rangle = \sum \overline{a}_k b_k = \overline{\sum a_k \overline{b}_k} = \overline{\langle b, a\rangle}\,.$

□

6.5 Definition. Orthogonalität

(a) Zwei Vektoren u, v eines Skalarproduktraumes V heißen **orthogonal**, falls

$$\langle u, v\rangle = 0\,.$$

Man schreibt dann auch $u \perp v$.

(b) Eine endliche oder abzählbare Menge von Vektoren $\{v_1, v_2, \ldots\}$ in V heißt Orthonormalsystem (ONS) oder Orthonormalfolge (ONF), falls

$$\langle v_i, v_j\rangle = \delta_{ij} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j\,. \end{cases}$$

Mit Hilfe eines Orthonormalsystems kann man einen Vektor orthogonal zerlegen. Sei $v_1, v_2, \ldots, v_N$ ein ONS in V, dann gilt für beliebiges $u \in V$ und $n \leq N$, dass

$$u = \underbrace{\sum_{j=1}^{n}\langle v_j, u\rangle v_j}_{=:u_n} + \underbrace{u - \sum_{j=1}^{n}\langle v_j, u\rangle v_j}_{=:u_n^{\perp}} = u_n + u_n^{\perp}\,.$$

Man rechnet leicht nach, dass $\langle u_n, u_n^{\perp}\rangle = 0$ gilt.

Diese simple Zerlegung hat zwei wichtige Konsequenzen:

6.6 Satz. Bessel und Cauchy-Schwarz

Sei V ein Skalarproduktraum und $v_1, \ldots, v_N$ ein ONS ($N = \infty$, d.h. eine Orthonormalfolge, ist erlaubt). Für $u \in V$ definieren wir

$$\|u\| := \sqrt{\langle u, u\rangle}\,.$$

Dann gilt für alle $u, v \in V$ und $n \leq N$ die

Bessel'sche Ungleichung $\qquad \|u\|^2 \geq \sum_{j=1}^{n} |\langle v_j, u\rangle|^2\,,$

und die

Cauchy-Schwarz'sche Ungleichung $\qquad |\langle u, v\rangle| \leq \|u\| \cdot \|v\|\,.$

Beweis. Wegen $\langle u_n, u_n^\perp\rangle = 0$ gilt der „Satz des Pythagoras“:

$$\begin{aligned}
\|u\|^2 &= \langle u,u\rangle = \langle u_n + u_n^\perp, u_n + u_n^\perp\rangle = \langle u_n,u_n\rangle + \langle u_n^\perp, u_n\rangle + \langle u_n, u_n^\perp\rangle + \langle u_n^\perp, u_n^\perp\rangle = \|u_n\|^2 + \|u_n^\perp\|^2 \\
&= \left\langle \sum_{j=1}^n \langle v_j, u\rangle v_j, \sum_{k=1}^n \langle v_k, u\rangle v_k \right\rangle + \|u_n^\perp\|^2 \\
&= \sum_{j=1}^n \sum_{k=1}^n \overline{\langle v_j, u\rangle} \langle v_k, u\rangle \underbrace{\langle v_j, v_k\rangle}_{=\delta_{jk}} + \|u_n^\perp\|^2 = \sum_{j=1}^n |\langle v_j, u\rangle|^2 + \|u_n^\perp\|^2 .
\end{aligned}$$

Daraus folgt mit der Positivität des Skalarprodukts sofort die Bessel'sche Ungleichung. Die Cauchy-Schwarz'sche Ungleichung ist trivial für $u = 0$, sei also $u \neq 0$. Dann ist $u_1 = \frac{u}{\|u\|}$ ein ONS (zugegeben ein sehr kleines) und Bessel für $n = 1$ liefert Cauchy-Schwarz:

$$\|v\|^2 \geq \left| \left\langle \frac{u}{\|u\|}, v \right\rangle \right|^2 = \frac{1}{\|u\|^2} |\langle u, v\rangle|^2 .$$

□

6.7 Definition. <u>**Normierter Raum**</u>

Eine Abbildung

$$\|\cdot\| : V \to [0,\infty)$$

auf einem $\mathbb{K}$-Vektorraum heißt Norm, falls

(a) $\|u\| = 0 \Leftrightarrow u = 0$ (Definitheit)

(b) $\|\lambda u\| = |\lambda|\,\|u\|$ für alle $\lambda \in \mathbb{K}, u \in V$ (Homogenität)

(c) $\|u+v\| \leq \|u\| + \|v\|$ für alle $u, v \in V$ (Dreiecksungleichung)

Eine Norm weist also jedem Vektor u seine Länge $\|u\|$ zu. Das Tupel $(V, \|\cdot\|)$ heißt normierter Raum.

6.8 Satz. In einem Skalarproduktraum $(V, \langle\cdot,\cdot\rangle)$ wird durch

$$\|u\| := \sqrt{\langle u, u\rangle} \qquad \text{für } u \in V$$

eine Norm auf V definiert.

Beweis. $\|u\| \geq 0$ und $\|u\| = 0 \Leftrightarrow u = 0$ folgen aus der positiven Definitheit (Definition 6.1 (a)) des Skalarproduktes. Die Homogenität folgt aus der Sesquilinearität des Skalarproduktes:

$$\|\lambda u\| = \sqrt{\langle \lambda u, \lambda u\rangle} = \sqrt{\lambda\overline{\lambda}\langle u, u\rangle} = |\lambda|\sqrt{\langle u,u\rangle} = |\lambda|\,\|u\| .$$

Die Dreiecksungleichung folgt aus Cauchy-Schwarz und $\mathrm{Re}(z) \leq |z|$ für $z \in \mathbb{C}$.

$$\begin{aligned}
\|u+v\|^2 &= \langle u+v, u+v\rangle = \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle = \langle u,u\rangle + \langle u,v\rangle + \overline{\langle u,v\rangle} + \langle v,v\rangle \\
&= \|u\|^2 + 2\,\mathrm{Re}\langle u,v\rangle + \|v\|^2 \\
&\leq \|u\|^2 + 2|\langle u,v\rangle| + \|v\|^2 \\
&\leq \|u\|^2 + 2\|u\|\,\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2 .
\end{aligned}$$

□

6.9 Lemma. Jedes Orthonormalsystem $v_1, v_2, \ldots$ ist linear unahängig.

Beweis. Aus $\sum_{j=1}^n \alpha_j v_j = 0$ folgt durch Projektion auf v_k, $k \in \{1,\ldots,n\}$,

$$0 = \left\langle v_k, \sum_{j=1}^n \alpha_j v_j \right\rangle = \sum_{j=1}^n \alpha_j \underbrace{\langle v_k, v_j \rangle}_{\delta_{kj}} = \alpha_k\,.$$

□

6.10 Korollar. Orthonormalbasen

(a) Ist $v_1,\ldots,v_n$ ein ONS in einem n-dimensionalen Skalarproduktraum V, so ist $(v_1,\ldots,v_n)$ eine Basis für V und heißt dann Orthonormalbasis (ONB).

(b) Sei $v_1,\ldots,v_n$ eine ONB von V, dann ist die Basisdarstellung von $u \in V$ durch

$$u = \langle v_1, u\rangle v_1 + \cdots + \langle v_n, u\rangle v_n = \sum_{k=1}^n \langle v_k, u\rangle v_k$$

gegeben.

(c) Für $u, w \in V$ und $v_1,\ldots,v_n$ eine ONB ist

$$\langle u, w\rangle = \sum_{k=1}^n \overline{\langle v_k, u\rangle}\langle v_k, w\rangle\,.$$

Insbesondere gilt die Parseval'sche Gleichung

$$\|u\|^2 = \sum_{k=1}^n |\langle v_k, u\rangle|^2\,.$$

Beweis. (a) In einem n-dimensionalen Raum bilden n linear unabhängige Vektoren eine Basis.

(b) Sei $u = \sum_{k=1}^n \alpha_j v_j$ die Basisdarstellung von u. Dann folgt durch Projektion auf v_k, $k \in \{1,\ldots,n\}$, dass

$$\langle v_k, u\rangle = \langle v_k, \sum_{k=1}^n \alpha_j v_j\rangle = \sum_{k=1}^n \alpha_j \langle v_k, v_j\rangle = \alpha_k\,.$$

(c) Für u, w gilt nach (b),

$$u = \sum_{j=1}^n \langle v_j, u\rangle v_j \quad \text{und} \quad w = \sum_{k=1}^n \langle v_k, w\rangle v_k\,,$$

also

$$\begin{aligned}\langle u, w\rangle &= \left\langle \sum_{j=1}^n \langle v_j, u\rangle v_j\,, \sum_{k=1}^n \langle v_k, w\rangle v_k \right\rangle = \sum_{j=1}^n \sum_{k=1}^n \overline{\langle v_j, u\rangle}\langle v_k, w\rangle\langle v_j, v_k\rangle \\ &= \sum_{j=1}^n \overline{\langle v_j, u\rangle}\langle v_j, w\rangle\,.\end{aligned}$$

□

6.11 Definition. Projektion

Sei V ein Vektorraum und $P \in \mathcal{L}(V)$. Dann heißt P **Projektion** falls,

(a) $P^2 = P$,

und **orthogonale Projektion** bezüglich eines Skalarprodukts $\langle\cdot,\cdot\rangle$, falls zusätzlich

(b) $\langle u, Pv\rangle = \langle Pu, v\rangle$ für alle $u, v \in V$.

6.12 Definition. **Orthogonales Komplement**
Für nichtleere Teilmengen $A \subset V$ eines Skalarproduktraumes heißt der Unterraum

$$A^{\perp} = \{u \in V \mid \langle u, v \rangle = 0 \;\; \forall\, v \in A\}$$

das **orthogonale Komplement** von A.

6.13 Satz. Sei $v_1, \ldots, v_n$ ein ONS in V und sei $P \in \mathcal{L}(V)$ definiert durch

$$Pu := \sum_{k=1}^{n} \langle v_k, u \rangle v_k \,. \quad (= u_n)$$

Dann ist P eine orthogonale Projektion, genannt Projektion auf $W = \operatorname{Span}\{v_1, \ldots, v_n\}$, und es gilt für alle $u \in V$

(a) $Pu \in W$ und $(1-P)u \in W^{\perp}$.

(b) $\|u - Pu\| = \min\{\|u - w\| \mid w \in W\}$

Schließlich hängt die Abbildung P nur von W ab und nicht von der Wahl der ONB $(v_1, \ldots, v_n)$ von W.

Beweis. (a) Sei $u \in V$, dann ist $Pu \in W$ klar. Sei $w \in W$, dann ist $w = \sum_{j=1}^{n} \lambda_j v_j$ für geeignete $\lambda_j \in \mathbb{K}$ und es gilt

$$\langle v, (1{-}P)u \rangle = \sum_{j=1}^{n} \bar{\lambda}_j \left(\langle v_j, u \rangle - \langle v_j \sum_{k=1}^{n} \langle v_k, u \rangle v_k \rangle \right) = \sum_{j=1}^{n} \bar{\lambda}_j \left(\langle v_j, u \rangle - \sum_{k=1}^{n} \langle v_k, u \rangle \langle v_j, v_k \rangle \right) = 0\,,$$

also $(1-P)u \in W^{\perp}$.

(b) Sei wieder $w \in W$ mit $w = \sum_{j=1}^{n} \lambda_j v_j$, dann gilt

$$\begin{aligned}
\|u - w\|^2 &= \|u - \sum_{j=1}^{n} \lambda_j v_j\|^2 = \langle u - \sum_{j=1}^{n} \lambda_j v_j \,, u - \sum_{k=1}^{n} \lambda_k v_k \rangle \\
&= \langle u, u \rangle - \sum_{j=1}^{n} \overline{\lambda_j} \langle v_j, u \rangle - \sum_{k=1}^{n} \lambda_k \langle u, v_k \rangle + \sum_{j=1}^{n} |\lambda_j|^2 \\
&= \|u\|^2 + \underbrace{\sum_{j=1}^{n} |\lambda_j - \langle v_j, u \rangle|^2}_{\geq 0} - \sum_{j=1}^{n} |\langle v_j, u \rangle|^2
\end{aligned}$$

Das eindeutige Minimum wird angenommen für $\lambda_j = \langle v_j, u \rangle$.

Wir zeigen nun, dass P eine orthogonale Projektion ist. Linearität folgt aus der Linearität des Skalarprodukts. Sei $w \in W$, dann gilt nach (b) $Pw = w$ und mit (a)

$$P^2 u = P\,\underbrace{(Pu)}_{\in W} = Pu \quad \forall\, u \in V\,.$$

P ist also eine Projektion. Wegen

$$\langle Pu, v \rangle - \langle u, Pv \rangle = \langle \underbrace{Pu}_{\in W} , \underbrace{v - Pv}_{\in W^{\perp}} \rangle - \langle \underbrace{u - Pu}_{\in W^{\perp}} , \underbrace{Pv}_{\in W} \rangle = 0$$

ist sie auch orthogonal. Die Eindeutigkeit folgt aus der Eindeutigkeit des Minimierers in (b) oder durch direktes Nachrechnen mit einer anderen ONB $(v'_1, \ldots, v'_n)$ von W.

□

Wie findet man Orthonormalsysteme bzw. Orthonormalbasen, und gibt es überhaupt immer welche?

Das Orthonormalisierungsverfahren von Gram-Schmidt

Sei V Skalarproduktraum und $u_1, u_2, \ldots$ eine endliche oder abzählbare linear unabhängige Menge von Vektoren in V. Wir konstruieren nun ein ONS $v_1, v_2, \ldots$ mit

$$\operatorname{Span}\{v_1, \ldots, v_n\} = \operatorname{Span}\{u_1, \ldots, u_n\} \quad \text{für alle } n \in \mathbb{N}\,.$$

1. Schritt: $v_1 = \frac{u_1}{\|u_1\|}$ erfüllt $\|v_1\| = \|\frac{u_1}{\|u_1\|}\| = 1$ und ist somit ein ONS mit $\operatorname{Span}\{v_1\}$=$\operatorname{Span}\{u_1\}$.

2. Schritt: Sei P_1 die orthogonale Projektion auf $W_1 = \operatorname{Span}\{u_1\} = \operatorname{Span}\{v_1\}$. Da $u_2 \notin W_1$, gilt $u_2 - P_1 u_2 \neq 0$ und wir können $v_2 = \frac{u_2 - P_1 u_2}{\|u_2 - P_1 u_2\|}$ definieren. Es gilt wieder $\|v_2\| = 1$ und $v_2 \in W_1^\perp$ da $u_2 - P_1 u_2 \in W_1^\perp$ und $W_1^\perp$ Unterraum ist. Da v_1, v_2 beide in $W_2 = \operatorname{Span}\{u_1, u_2\}$ liegen und linear unabhängig sind, folgt $\operatorname{Span}\{v_1, v_2\} = W_2$.

n-ter Schritt: Seien $v_1, \ldots, v_{n-1}$ schon konstruiert mit

$$\operatorname{Span}\{v_1, \ldots, v_{n-1}\} = \operatorname{Span}\{u_1, \ldots, u_{n-1}\} = W_{n-1}$$

und sei P_{n-1} die orthogonale Projektion auf W_{n-1}.
Dann folgt wieder aus $u_n \in W_{n-1}$, dass $u_n - P_{n-1} u_n \neq 0$ und

$$v_n = \frac{u_n - P_{n-1} u_n}{\|u_n - P_{n-1} u_n\|}$$

erfüllt $\|v_n\| = 1$ und $v_n \in W_{n-1}^\perp$. Wiederum folgt aus $v_1, \ldots, v_n \in \operatorname{Span}\{u_1, \ldots, u_n\}$ und $v_1, \ldots, v_n$ ONS, dass $\operatorname{Span}\{v_1, \ldots, v_n\} = \operatorname{Span}\{u_1, \ldots, u_n\}$. Bei abzählbar vielen u_1, u_2 folgt die Existenz des behaupteten ONS per Induktion.

6.14 Korollar. Für jeden n-dimensionalen Skalarproduktraum V gilt:

(a) V besitzt eine ONB.

(b) Jedes ONS $v_1, \ldots, v_m$ lässt sich zu einer ONB $(v_1, \ldots, v_n)$ für V ergänzen.

(c) Für U= $\operatorname{Span}\{v_1, \ldots, v_m\}$ gilt dann $U^\perp = \operatorname{Span}\,\{v_{m+1}, \ldots, v_n\}$ also $\dim U + \dim U^\perp = n$.

Die Norm in Skalarprodukträumen ist durch das Skalarprodukt gegeben, $\|u\| = \sqrt{\langle u, u\rangle}$. Man kann aber auch umgekehrt das Skalarprodukt aus der Norm rekonstruieren.

6.15 Satz. **Polarisation und Parallelogrammgleichung**

(a) In jedem Skalarproduktraum über $\mathbb{R}$ gilt

$$\langle u, v\rangle = \frac{1}{4}(\|u+v\|^2 - \|u-v\|^2)$$

(b) In jedem Skalarproduktraum über $\mathbb{C}$ gilt

$$\langle u, v\rangle = \frac{1}{4}(\|u+v\|^2 - \|u-v\|^2 + \mathrm{i}\|u - \mathrm{i}v\|^2 - \mathrm{i}\|u + \mathrm{i}v\|^2)$$

(c) In jedem Skalarproduktraum gilt die **Parallelogrammgleichung**.

$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2\,.$$

Umgekehrt ist ein Vektorraum mit Norm $\|\cdot\|$ genau dann ein Skalarproduktraum, wenn die Parallelogrammgleichung gilt. Das zugehörige Skalarprodukt ist dann durch die Polarisationsgleichung gegeben.

Beweis. (a) und (b) rechnet man einfach nach, indem man auf der rechten Seite jeweils $\|\cdot\|^2 = \langle\cdot,\cdot\rangle$ einsetzt und zusammenfasst. Ebenso zeigt man die Parallelogrammgleichung in Skalarprodukträumen. Es bleibt zu zeigen, dass durch die Polarisationsgleichung ein Skalarprodukt definiert wird, wann immer eine Norm die Parallelogrammgleichung erfüllt. Dazu prüft man direkt die Eigenschaften (a), (b) und (c) aus Definition 6.1. □

6.16 Definition. **Isometrie, orthogonale und unitäre Abbildung**

(a) Eine lineare Abbildung $L \in \mathcal{L}(V, W)$ zwischen normierten Räumen V und W heißt **Isometrie**, wenn $\|Lu\|_W = \|u\|_V$ für alle $u \in V$.

(b) Eine surjektive Isometrie L zwischen Skalarprodukträumen heißt **unitär**, bzw. im Fall $\mathbb{K} = \mathbb{R}$ und $V = W$ auch **orthogonal**.

6.17 Lemma. Sei $L \in \mathcal{L}(V, W)$ eine Isometrie. Dann gilt

(a) L ist injektiv.

(b) Falls $\dim W = \dim V < \infty$, so ist L surjektiv.

(c) Sind V und W Skalarprodukträume, so gilt auch $\langle Lu, Lv\rangle_W = \langle u, v\rangle_V \quad \forall\, u, v \in V$.

Für $W = V = \mathbb{R}^n$ mit dem natürlichen Skalarprodukt sind Isometrien also auch winkeltreu.

Beweis. (a) $Lu = 0 \;\Rightarrow\; \|u\|_V = \|Lu\|_W = 0 \;\Rightarrow\; u = 0\,.$

(b) Für lineare Abbildungen zwischen endlichdimensionalen Räumen gilt nach Satz 2.31, dass L injektiv $\Leftrightarrow$ L surjektiv.

(c) folgt aus Polarisierungsidentität.

□

6.18 Bemerkung. Für $\dim V = \infty$ gilt (b) im Allgemeinen nicht. Ein Beispiel ist der Rechtsshift R auf $\ell^2 := \{(a_n)_{n\in\mathbb{N}} | \sum_{n=1}^{\infty} |a_n|^2 < \infty\}$. Es ist

$$R : \ell^2 \to \ell^2, (a_1, a_2, \ldots) \mapsto (0, a_1, a_2, \ldots)$$

offensichtlich eine Isometrie, aber nicht surjektiv.

6.19 Lemma. Sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine ONB von V. Dann ist die Koordinatenabbildung

$$\Phi_{\mathcal{B}} : V \to \mathbb{K}^n\,, u \mapsto (u)_{\mathcal{B}} = \begin{pmatrix} \langle v_1, u\rangle \\ \vdots \\ \langle v_n, u\rangle \end{pmatrix}$$

unitär, wobei $\mathbb{K}^n$ mit dem natürlichen Skalarprodukt versehen ist.

Beweis. Parseval'sche Gleichung. □

Also ist jeder endlichdimensionale Skalarproduktraum V über $\mathbb{K}$ **unitär äquivalent** ($\equiv$ **isometrisch isomorph**) zu $\mathbb{K}^n$ mit dem natürlichen Skalarprodukt.

6.20 Bemerkung. In der Mathematik betrachtet man zwei Objekte vom gleichen Typ (z.B. Vektorräume, Skalarprodukträume, Gruppen usw.) als im wesentlichen identisch, wenn es strukturerhaltende Abbildungen zwischen ihnen gibt: für Vektorräume sind das die Isomorphismen (lineare Bijektionen), für Skalarprodukträume die isometrischen Isomorphismen (also die unitäre

Abbildungen), für Gruppen die Gruppen-Isomorphismen (also Bijektionen, welche die Gruppenstruktur erhalten).

Wir betrachten im folgenden endlichdimensionale Skalarprodukträume. Da diese alle isometrisch isomorph zu $\mathbb{K}^n$ sind, betrachten wir zunächst nur den $\mathbb{K}^n$. Man sollte aber im Hinterkopf behalten, dass alle Resultate und Definitionen sich direkt auf beliebige endlichdimensionale Skalarprodukträume mittels des isometrischen Isomorphismus aus Lemma 6.19 übertragen lassen.

Außerdem finden sich viele der folgenden Konzepte in sehr ähnlicher Form auch für ∞-dimensionale Skalarprodukträume und insbesondere für Hilberträume wieder. Sie sind für die Physik von großer Bedeutung.

6.21 Definition. Die adjungierte Matrix

Die zu einer $m \times n$-Matrix A adjungierte Matrix A^* ist $A^* = \overline{A^T}$, also $a^*_{ij} = \overline{a_{ji}}$. Für reelle Matrizen ist demnach $A^* = A^T$.

6.22 Lemma. Sei $A \in M(n \times m, \mathbb{K})$ dann gilt:

(a) $\langle x, Ay \rangle_{\mathbb{K}^n} = \langle A^*x, y \rangle_{\mathbb{K}^m}$ für alle $x \in \mathbb{K}^n, y \in \mathbb{K}^m$

(b) Rang A^* = Rang A.

Falls $A \in M(n \times n, \mathbb{K})$ dann gelten weiterhin

(c) $\det A^* = \overline{\det A}$.

(d) λ ist genau dann Eigenwert von A, wenn $\overline{\lambda}$ Eigenwert von A^* ist. Die geometrische und die algebraische Vielfachheit stimmen überein.

Beweis. (a)
$$\begin{aligned}\langle x, Ay \rangle_{\mathbb{K}^n} &= \sum_{j=1}^{n} \overline{x_j}\,(Ay)_j = \sum_{j=1}^{n} \overline{x_j} \sum_{k=1}^{n} a_{jk} y_k = \sum_{k=1}^{m}\sum_{j=1}^{n} a_{jk}\overline{x_j}\, y_k \\ &= \sum_{k=1}^{m}\sum_{j=1}^{n} \overline{a^*_{kj} x_j}\; y_k = \sum_{k=1}^{m} \overline{(A^*x)_k}\, y_k = \langle A^*x, y\rangle_{\mathbb{K}^m}\,.\end{aligned}$$

(b) Da Spaltenrang = Zeilenrang gilt, folgt die Aussage daraus, dass Vektoren $a_1, \dots, a_\ell \in \mathbb{K}^n$ genau dann linear unabhängig sind, wenn $\overline{a_1}, \dots, \overline{a_\ell} \in \mathbb{K}^n$ linear unabhängig sind.

(c) Wir wissen bereits, dass $\det A = \det A^T$. Aus der Leibnizformel ergibt sich $\det A^* = \overline{\det A^T}$.

(d) Übungsaufgabe.

$\square$

6.23 Lemma. (a) Sei $A \in M(m \times k, \mathbb{K})$ und $B \in M(k \times n, \mathbb{K})$, dann gilt

$$(AB)^* = B^* A^*\,.$$

(b) Sei $A \in M(m \times n, \mathbb{K})$, dann ist $A^*A \in M(n \times n, \mathbb{K})$ und es gilt für $x \in \mathbb{K}^n$

$$\langle x, A^*A\,x \rangle_{\mathbb{K}^n} = \|A\,x\|^2_{\mathbb{K}^m}\,.$$

Beweis. (a) $$(AB)^*_{ij} = \overline{\sum_{\ell=1}^{k} a_{j\ell} b_{\ell i}} = \overline{\sum_{\ell=1}^{k} \overline{a^*_{\ell j}}\; \overline{b^*_{i\ell}}} = \sum_{\ell=1}^{k} b^*_{i\ell}\, a^*_{\ell j} = (B^*A^*)_{ij}$$

(b) Mit Lemma 6.22 (a) und der Symmetrie des Skalarprodukts gilt

$$\langle x, A^*Ax \rangle_{\mathbb{K}^n} = \overline{\langle A^*Ax, x\rangle}_{\mathbb{K}^n} = \overline{\langle Ax, Ax\rangle}_{\mathbb{K}^m} = \langle Ax, Ax\rangle_{\mathbb{K}^m} = \|Ax\|^2_{\mathbb{K}^m}\,.$$

$\square$

6.24 Bemerkung. Wir haben bisher nur die adjungierte Matrix aber nicht die adjungierte Abbildung definiert. Nun könnte man einfach mittels des Basisisomorphismus

$$\Phi_{\mathcal{B}} : V \to \mathbb{K}^n, \;\; u \mapsto (u)_{\mathcal{B}}$$

die Definition der adjungierten Matrix auf lineare Abbildungen $L \in \mathcal{L}(V, W)$ hochziehen: Seien $\dim V = n$, $\dim W = m$ und $\mathcal{A}$ und $\mathcal{B}$ Orthonormalbasen, dann bildet $M^{\mathcal{B}}_{\mathcal{A}}(L)$ den $\mathbb{K}^n$ auf $\mathbb{K}^m$ ab und $M^{\mathcal{B}}_{\mathcal{A}}(L)^*$ den $\mathbb{K}^m$ auf den $\mathbb{K}^n$. Also wäre

$$L^* := \Phi_{\mathcal{A}}^{-1} \left(M^{\mathcal{B}}_{\mathcal{A}}(L)\right)^* \Phi_{\mathcal{B}}$$

eine lineare Abbildung von W nach V. Diese Definition ist zwar naheliegend und gewissermassen auch korrekt, man übersieht dabei aber einen wichtigen strukturellen Aspekt.

Exkurs: Der Dualraum

6.25 Definition. <u>**Dualraum**</u>

Sei V ein endlichdimensionaler Vektorraum. Dann heißt der Vektorraum $\mathcal{L}(V, \mathbb{K})$, also der Raum der linearen Abbildungen von V nach $\mathbb{K}$, der **Dualraum** von V und wird mit V' bezeichnet.

6.26 Bemerkung. In unendlichdimensionalen topologischen Vektorräumen ist V' der Raum der stetigen linearen Abbildungen von V nach $\mathbb{K}$. In endlichdimensionalen Räumen sind aber alle linearen Abbildungen stetig.

6.27 Beispiel. Der Dualraum $(\mathbb{K}^n)'$ von $\mathbb{K}^n$ (=Spaltenvektoren) ist der Raum der $1 \times n$-Matrizen, $\mathcal{L}(\mathbb{K}^n, \mathbb{K}) = M(1 \times n, \mathbb{K})$, also der Raum der Zeilenvektoren: Für $x \in (\mathbb{K}^n)'$ und $y \in \mathbb{K}^n$ ist

$$x(y) = (x_1, \ldots, x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + \cdots + x_n y_n \, .$$

Es sind also $\mathbb{K}^n$ und $(\mathbb{K}^n)'$ in natürlicher Weise isomorph.

6.28 Bemerkung. Für beliebige Skalarprodukträume V gibt es eine natürliche Identifizierung I zwischen V und V': Zu $u \in V$ ist $I(u) \in V' = \mathcal{L}(V, \mathbb{K})$ gegeben durch

$$I(u) : V \to \mathbb{K}, \;\; v \mapsto \langle u, v \rangle \, .$$

Tatsächlich läßt sich jede lineare Abbildung $L : V \to \mathbb{K}$ in der Form $L(v) = \langle u, v \rangle$ für ein geeignetes $u \in V$ schreiben. Für $\mathbb{K}^n$ ist das klar und mit Lemma 6.19 folgt das für beliebige endlichdimensionale Skalarprodukträume. Somit ist $I : V \to V'$ eine antilineare Bijektion. Für unendlichdimensionale Hilberträume gilt ein entsprechendes Resultat.

6.29 Definition. <u>**Adjungierte Abbildung**</u>

Für $L \in \mathcal{L}(V, W)$ ist die adjungierte Abbildung L' definiert durch

$$L' : W' \to V', \;\; (\underbrace{L'w'}_{\in V'})(v) := w'(Lv) \quad \forall\, w' \in W' \, , \; v \in V \, .$$

Diese Definition nimmt keinen Bezug auf ein Skalarprodukt. In Skalarprodukträumen kann man nun aber die Dualräume V' und W' mit den Räumen selbst identifizieren und bekommt somit eine Abbildung $L^* : W \to V$ gemäß

$$L^* : W \to V \, , \;\; L^* w = I_V^{-1} \, L' \, I_W \, w \, .$$

Dann folgt die Eigenschaft

$$\langle w, Lv\rangle_W = \langle L^* w, v\rangle_V \quad \text{für alle } w \in W\,,\ v \in V\,,$$

direkt aus der Definition:

$$\langle w, Lv\rangle_W = \underbrace{Iw}_{\in W'}(\underbrace{Lv}_{\in W}) = (\underbrace{L'(Iw)}_{\in V'})(\underbrace{v}_{\in V}) = \langle I^{-1}L'Iw, v\rangle_V\,.$$

Wir merken uns also: Der adjungierte Operator zu $L \in \mathcal{L}(V,W)$ ist eigentlich eine Abbildung zwischen den Dualräumen W' und V'. Für Skalarprodukträume kann man aber die Dualräume mit den Räumen selbst identifizieren und somit die adjungierte Abbildung als Abbildung von W nach V auffassen.

6.30 Bemerkung. In der Physik nennt man die Elemente eines Vektorraums V oft die kontravarianten Vektoren und die Elemente des Dualraums V' die kovarianten Vektoren. Hat man ein Skalarprodukt auf V, so kann man ko- und kontravariante Vektoren ineinander umwandeln. Die Bezeichnungen gehen auf das unterschiedliche Transformationsverhalten unter Basiswechseln zurück, siehe Übungen.

6.31 Satz. Eine Matrix $A \in M(n\times n, \mathbb{K})$ definiert genau dann eine unitäre Abbildung $A : \mathbb{K}^n \to \mathbb{K}^n$, wenn $A^* = A^{-1}$ gilt.

Beweis. Mit Lemma 6.17 ist A genau dann unitär, wenn A isometrisch ist. Sei also $A^* = A^{-1}$ und $x \in \mathbb{K}^n$, dann gilt

$$\|Ax\|^2 = \langle Ax, Ax\rangle = \langle A^*Ax, x\rangle = \langle A^{-1}Ax, x\rangle = \langle x, x\rangle = \|x\|^2\,.$$

Also ist A eine Isometrie. Sei umgekehrt A eine Isometrie, dann ist für jedes $x \in \mathbb{K}^n$

$$\langle A^*Ax, y\rangle = \langle Ax, Ay\rangle = \langle x, y\rangle$$

für alle $y \in \mathbb{K}^n$ und somit $A^*Ax = x$, also $A^* = A^{-1}$. □

6.32 Korollar. Seien V und W endlichdimensionale Skalarprodukträume und $L \in \mathcal{L}(V,W)$. Es ist L genau dann unitär, wenn $L^* = L^{-1}$ gilt.

6.33 Bemerkung. Eine $n\times n$-Matrix ist genau dann unitär bzw. orthogonal, wenn ihre Spalten (oder analog ihre Zeilen) eine ONB des $\mathbb{K}^n$ bilden, denn für die Spalten von $A = (a_1, \ldots, a_n)$ gilt dann

$$\langle a_j, a_i\rangle = (A^*A)_{ji} = (E_n)_{ji} = \delta_{ji}\,.$$

Die Transformationsmatrix S eines Basiswechsels zwischen ONBen des $\mathbb{K}^n$ ist also unitär, da die Spalten als Bilder der Basisvektoren wieder eine ONB bilden. Umgekehrt kann man jede unitäre Matrix als Transformationsmatrix eines Basiswechsels zwischen ONBen interpretieren.

6.34 Satz. Sei S eine unitäre $n\times n$-Matrix.

(a) Jeder Eigenwet λ von S liegt auf dem Einheitskreis, d.h. $|\lambda| = 1$.

(b) $|\det S| = 1$

Beweis. (a) Sei $\lambda \in \mathbb{C}$ Eigenwert von S und u ein zugehöriger Eigenvektor. Dann gilt

$$\|u\| = \|Su\| = \|\lambda u\| = |\lambda|\|u\|\,.$$

Da $u \neq 0$, muss $|\lambda| = 1$ gelten.
(b) $1 = \det E_n = \det S^*S = \det S^* \det S = \overline{\det S}\,\det S = |\det S|^2$. □

6.35 Beispiel. Die unitären Abbildungen auf dem $\mathbb{R}^2$, also in diesem Fall die orthogonalen, sind die Spiegelungen und Drehungen. Nur diese lassen alle Winkel und Längen intakt. Formal folgt das so: da die Spalten einer orthogonalen Matrix eine ONB bilden, kann man für die erste Spalte

$$s_1 = \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix} \quad \text{für ein } \varphi \in [0, 2\pi)$$

ansetzen. Es gibt nun genau zwei zu s_1 orthogonale Vektoren der Länge eins, nämlich

$$s_2 = \begin{pmatrix} -\sin\varphi \\ \cos\varphi \end{pmatrix} \quad \text{und} \quad s_3 = \begin{pmatrix} \sin\varphi \\ -\cos\varphi \end{pmatrix}.$$

Es ist $D_\varphi = (s_1, s_2)$ die Drehung um den Winkel φ und $S_{\varphi/2} = (s_1, s_3)$ die Spiegelung an der Ursprungsgeraden mit Winkel $\varphi/2$ zur x-Achse.
Es gilt $\det D_\varphi = 1$, da Drehungen die Orientierung einer Basis nicht ändern und $\det S_{\varphi/2} = -1$, da Spiegelungen die Orientierung einer Basis ändern.

Exkurs: Gruppen

6.36 Definition. Gruppe
Eine Gruppe ist ein Paar $(G, \circ)$ bestehend aus einer Menge G und einer Abbildung

$$\begin{aligned} \circ : G \times G &\to G \\ (a, b) &\mapsto a \circ b \end{aligned}$$

mit den folgenden Eigenschaften:

(a) $(a \circ b) \circ c = a \circ (b \circ c)$ für alle $a, b, c \in G$ (Assoziativität)

(b) Es gibt ein $e \in G$ mit $a \circ e = e \circ a = a$ für alle $a \in G$ (Existenz des neutralen Elements).

(c) Zu jedem $a \in G$ gibt es ein $a^{-1} \in G$ mit $a \circ a^{-1} = a^{-1} \circ a = e$. (Existenz des Inversen).

Eine Teilmenge $H \subset G$, welche selbst wieder eine Gruppe ist, heißt Untergruppe.

6.37 Bemerkung. Es genügt in (b) und (c) jeweils nur $e \circ a = a$ und $a^{-1} \circ a = e$ zu fordern. Die jeweils andere Gleichung folgt dann:

- Jedes linksinverse Element ist auch rechtsinvers: sei $a^{-1} \circ a = e$, dann ist aufgrund der Assoziativität

$$\begin{aligned} a \circ a^{-1} &= e \circ (a \circ a^{-1}) = ((a^{-1})^{-1} \circ a^{-1}) \circ (a \circ a^{-1}) = \\ &= (a^{-1})^{-1} \circ (a^{-1} \circ a) \circ a^{-1} = (a^{-1})^{-1} \circ a^{-1} = e\,. \end{aligned}$$

- Jedes linksneutrale Element ist auch rechtsneutral: sei $e \circ a = a$, dann ist

$$a \circ e = a \circ (a^{-1} \circ a) = (a \circ a^{-1}) \circ a = a\,.$$

6.38 Beispiel. Matrixgruppen:
Die Verknüpfung ist im Folgenden durch das Matrixprodukt gegeben, das neutrale Element ist jeweils die Einheitsmatrix.

(a) Alle invertierbaren $n \times n$-Matrizen bilden die lineare Gruppe $GL(n, \mathbb{K})$:
für $A, B \in GL(n, \mathbb{K})$ ist auch AB wieder invertierbar, da $(A\,B)^{-1} = B^{-1}A^{-1}$.

(b) Die unitären komplexen Matrizen bilden die unitäre Gruppe $U(n)$:
für $A, B \in U(n)$ ist auch AB wieder unitär, da $(AB)^* = B^*A^* = B^{-1}A^{-1} = (AB)^{-1}$.
$U(n)$ ist Untergruppe von $GL(n)$.

(c) Die orthogonalen $n \times n$-Matrizen bilden die orthogonale Gruppe $O(n)$. $O(n)$ ist Untergruppe von $U(n)$.

(d) Die spezielle lineare Gruppe ist $SL(n, \mathbb{K}) = \{A \in GL(n, \mathbb{K}) \,|\, \det A = 1\}$. Wegen $\det AB = \det A \cdot \det B$ ist $SL(n, \mathbb{K})$ wieder eine Gruppe.

(e) Die spezielle unitäre Gruppe $SU(n) = \{A \in U(n) \,|\, \det A = 1\}$.

(f) Die spezielle orthogonale Gruppe $SO(n) = \{A \in O(n) \,|\, \det A = 1\}$ besteht genau aus den Drehmatrizen.

Analog zu (a)–(f) definiert man für Vektorräume V die Gruppen $GL(V), SL(V)$ und für Skalarprodukträume zusätzlich $U(V)$ und $SU(V)$.

(g) Sei Allgemein M eine Menge und $\mathrm{Bij}(M)$ die Menge der bijektiven Abbildung $f : M \to M$. Dann ist $(\mathrm{Bij}(M), \circ)$ eine Gruppe. Hier steht $\circ$ für die Komposition von Abbildungen und das neutrale Element ist die Identität.

6.39 Definition. **Abelsche Gruppe**

Ist eine Gruppe $(G, \circ)$ kommutativ, d.h.

$$a \circ b = b \circ a \quad \text{für alle } a, b \in G\,,$$

dann heißt G eine **abelsche** Gruppe.

6.40 Beispiel.

- $(\mathbb{Z}, +)$ ist eine abelsche Gruppe
- $(\mathbb{R}, +)$ ist eine abelsche Gruppe
- Ist $(V, +, \cdot)$ ein Vektorraum über $\mathbb{K}$, so ist $(V, +)$ eine abelsche Gruppe
- $SO(2)$ ist eine abelsche Gruppe
- $SO(3)$ ist nicht abelsch

6.41 Beispiel. **Transformation des Raumes und der Raumzeit**

- Die **Bewegungsgruppe** B ist die Gruppe der längentreuen Abbildungen $f : \mathbb{R}^3 \to \mathbb{R}^3$ mit $\|f(x) - f(y)\| = \|x - y\|$ für alle $x, y \in \mathbb{R}^3$. Man kann zeigen, dass sich jedes $f \in B$ schreiben läßt als Verknüpfung einer Translation und einer Drehung, also für $f \in B$ gibt es $a \in \mathbb{R}^3$ und $D \in O(3)$ so, dass $f(x) = Dx + a$.
- **Die Galilei-Gruppe** (nichtrelativistische Raumzeitsymmetrie)

  Die nichtrelativistischen Gesetze der Physik (Newton'sche Mechanik, Quantenmechanik) sind invariant unter

  – Drehungen und Tanslationen des Raumes

  $$x \mapsto x' = Dx + a\,, \quad D \in SO(3)\,, \;\; a \in \mathbb{R}^3\,.$$

  – Verschiebung des Zeitnullpunktes

  $$t \mapsto t' = t + \tau\,, \quad \tau \in \mathbb{R}\,.$$

  – gleichförmige Bewegung

  $$x \mapsto x' = x + tv\,, \quad v \in \mathbb{R}^3\,. \qquad \text{(Galilei-Boosts)}$$

  Zusammen bilden diese Transformationen die Galilei-Gruppe. Fassen wir Zeit und Ort zu einem Vektor $\begin{pmatrix} t \\ x \end{pmatrix} \in \mathbb{R}^4$ (=Raumzeit) zusammen, so sind alle Galilei-Transformationen von der Form

  $$\begin{pmatrix} t \\ x \end{pmatrix} \mapsto T_{(D,a,\tau,v)} \begin{pmatrix} t \\ x \end{pmatrix} := \begin{pmatrix} t + \tau \\ Dx + a + tv \end{pmatrix}.$$

Zusammen bilden die Abbildungen $T_{(D,a,\tau,v)} : \mathbb{R}^4 \to \mathbb{R}^4$ die inhomogene Galilei-Gruppe. Die Gruppeneigenschaft rechnet man direkt nach:

$$\begin{aligned} T_{(D_2,a_2,\tau_2,v_2)} \circ T_{(D_1,a_1,\tau_1,v_1)} \begin{pmatrix} t \\ x \end{pmatrix} &= T_{(D_2,a_2,\tau_2,v_2)} \begin{pmatrix} t+\tau_1 \\ D_1 x + a_1 + t v_1 \end{pmatrix} = \\ &= \begin{pmatrix} t+\tau_1+\tau_2 \\ D_2(D_1 x + a_1 + t v_1) + a_2 + (t+\tau_1) v_2 \end{pmatrix} \\ &= T_{(D_2 D_1, D_2\, a_1 + a_2 + \tau_1 v_2, \tau_1+\tau_2, D_2 v_1 + v_2)} \begin{pmatrix} t \\ x \end{pmatrix} . \end{aligned}$$

Die Galilei Transformationen erhalten die Gleichzeitigkeit. Für $\begin{pmatrix} t \\ x \end{pmatrix}$ und $\begin{pmatrix} s \\ y \end{pmatrix}$ gilt: Falls $t = s$, dann stimmen auch nach jeder Galilei Transformation die ersten Komponenten wieder überein. Setzt man $a = 0$ und $\tau = 0$, so bilden die $T_{(D,v)}$ die homogene Galileigruppe, eine Untergruppe von $GL(\mathbb{R}^4)$:

$$T_{(D,v)} \begin{pmatrix} t \\ x \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ v & D \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix} = \begin{pmatrix} t \\ Dx + vt \end{pmatrix} .$$

Was bedeutet es nun, dass ein physikalisches Gesetz oder eine Theorie Galilei-invariant ist? Dazu sollten die Lösungen durch Objekte auf der Raumzeit $\mathbb{R}^4$ gegeben sein, z. B. Weltlinien von Teilchen, d.h. Kurven

$$\gamma_j : \mathbb{R} \to \mathbb{R}^4 , \quad \xi \mapsto \gamma_j(\xi) = \begin{pmatrix} t_j(\xi) \\ x_j(\xi) \end{pmatrix} , \quad j = 1, \ldots, N ,$$

wobei γ_j die Weltlinie des j-ten Teilchens ist und normalerweise angenommen wird, dass $t_j : \mathbb{R} \to \mathbb{R}$, $\xi \mapsto t_j(\xi)$ eine Bijektion ist. Eine Theorie deren Lösungen solche Weltlinien sind, z.B. Newton'sche Mechanik, heißt nun Galilei-invariant, wenn mit jeder Lösung $(\gamma_j)_{j=1,\ldots,N}$ und für jede Galileitransformation $T_{(D,a,\tau,v)}$ auch $(T_{(D,a,\tau,v)} \circ \gamma_j)_{j=1,\ldots,N}$ eine Lösung ist. Die Parametrisierung der Kurven γ_j spielt hier übrigens keine Rolle.

- **Die Lorentz-Gruppe** (relativistische Raumzeitsymmetrie)
  Ersetzt man die Galilei-Boosts $T_{(\mathbb{1},0,0,v)}$ durch Lorentz-Boosts,

$$\begin{pmatrix} t \\ x \end{pmatrix} \mapsto \begin{pmatrix} t' \\ x' \end{pmatrix} = L_v \begin{pmatrix} t \\ x \end{pmatrix} := \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \begin{pmatrix} t + \frac{v \cdot x}{c^2} \\ x + vt \end{pmatrix} = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \begin{pmatrix} 1 & \frac{v^T}{c^2} \\ v & \mathbb{1} \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix}$$

erhält man die Lorentzgruppe (oder auch Poincarégruppe). Die Lorentztransformationen erhalten die Gleichzeitigkeit nicht: für $\begin{pmatrix} t \\ x \end{pmatrix}$ und $\begin{pmatrix} s \\ y \end{pmatrix}$ mit $t = s$ stimmen die ersten Komponenten von $L_v \begin{pmatrix} t \\ x \end{pmatrix}$ und $L_v \begin{pmatrix} t \\ y \end{pmatrix}$ im Allgemeinen nicht überein. Andererseits ist L_v auf den letzten drei Komponenenten auch nicht mehr längentreu:

$$L_v \begin{pmatrix} t \\ x \end{pmatrix} - L_v \begin{pmatrix} t \\ y \end{pmatrix} = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \begin{pmatrix} \frac{v \cdot (x-y)}{c^2} \\ (x-y) \end{pmatrix} .$$

Was bedeutet das für die Lösungen Lorentz-invarianter Theorien? Nehmen wir an, die Endpunkte eines ruhenden Stabes seien durch die Weltlinien

$$\gamma_1(\xi) = \begin{pmatrix} \xi \\ 0 \end{pmatrix} \quad \text{und} \quad \gamma_2(\zeta) = \begin{pmatrix} \zeta \\ 1 \end{pmatrix}$$

gegeben. (Der einfacheren Notation wegen nehmen wir einen eindimensionalen Raum also eine zweidimensionale Raumzeit an). In einer Lorentz-invarianten Theorie ist dann auch

$$\gamma_1'(\xi) = (L_v \circ \gamma_1)(\xi) = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \begin{pmatrix} \xi \\ v\,\xi \end{pmatrix}, \quad \gamma_2'(\zeta) = (L_v \circ \gamma_2)(\zeta) = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \begin{pmatrix} \zeta + \frac{v}{c^2} \\ 1 + v\,\zeta \end{pmatrix}$$

eine Lösung. Wie lang ist nun aber dieser bewegte Stab? Die Zeitkomponenten der Lösungen stimmen überein, wenn wir $\xi = \zeta + \frac{v}{c^2}$ wählen. Die Differenz der Ortskomponenten ist dann

$$\left\| \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \Big(v(\zeta + \frac{v}{c^2}) - 1 - v\,\zeta \Big) \right\| = \sqrt{1-\frac{v^2}{c^2}}\,,$$

also um den Faktor $\sqrt{1-\frac{v^2}{c^2}}$ kleiner als die ursprüngliche Differenz. Das ist die sogenannte Längen- oder Lorentzkontraktion. Bewegte Körper erscheinen uns (in unserem ruhenden Koordinatensystem) in der Bewegungsrichtung verkürzt. Natürlich könnten wir die Lorentztransformation auch als einen Koordinatenwechsel interpretieren, d.h. $\gamma_1'(\xi)$ und $\gamma_2'(\zeta)$ wären die Koordinaten des Stabes in einem gleichförmig mit der Geschwindigkeit $-v$ bewegten Koordinatensytems. Jemand der dieses Koordinatensystem verwendet (man spricht dann von einem bewegten Beobachter), würde dann ebenfalls die Länge des Stabes mit $\sqrt{1-\frac{v^2}{c^2}}$ angeben.

7 Symmetrische Operatoren

7.1 Definition. <u>**Symmetrie und Selbstadjungiertheit**</u>

(a) Eine $n \times n$-Matrix A heißt **symmetrisch** oder **selbstadjungiert**, wenn $A = A^*$ gilt.

(b) Eine lineare Abbildung $T \in \mathcal{L}(V)$ auf einem Skalarproduktraum V heißt **symmetrisch** oder **selbstadjungiert**, wenn

$$\langle Tu, v\rangle = \langle u, Tv\rangle \quad \text{für alle} \quad u, v \in V\,.$$

7.2 Bemerkung. Zumindest auf einem endlichdimensionalen Raum ist also T symmetrisch genau dann wenn $T = T^*$ gilt. (Vgl. Lemma 6.22 und Exkurs Dualraum). Dies ist wiederum genau dann der Fall, wenn die Matrix $M_{\mathcal{B}}(T)$ bzgl. einer (und damit bzgl. jeder) ONB symmetrisch ist.

7.3 Satz. Seien S und T Endomorphismen eines Skalarproduktraumes V. Dann gilt

(a) Sind S und T selbstadjungiert, so auch $\alpha S + \beta T$ für $\alpha, \beta \in \mathbb{R}$.

(b) Ist T selbstadjungiert und invertierbar, so ist auch T^{-1} selbstadjungiert.

(c) Ist V ein komplexer Skalarproduktraum, so ist T genau dann selbstadjungiert, wenn

$$\langle u, Tu\rangle \in \mathbb{R} \quad \text{für alle} \quad u \in V\,.$$

Beweis. (a) Klar, Sesquilinearität des Skalarprodukts.

(b) Seien $u, v \in V$, dann ist

$$\langle T^{-1}u, v\rangle = \langle T^{-1}u, TT^{-1}v\rangle = \langle TT^{-1}u, T^{-1}v\rangle = \langle u, T^{-1}v\rangle\,.$$

(c) Übungsaufgabe.

□

Wir werden zeigen, dass selbstadjungierte Operatoren auf endlichdimensionalen Skalarprodukträumen immer diagonalisierbar sind. Das folgt im Wesentlichen schon aus den folgenden sehr simplen Beobachtungen:

7.4 Satz. Sei V ein Skalarproduktraum und $T \in \mathcal{L}(V)$ selbstadjungiert. Dann gilt:

(a) Alle Eigenwerte von T sind reell.

(b) Eigenvektoren von T zu verschiedenen Eigenwerten sind orthogonal.

(c) Ist u ein Eigenvektor von T, so ist der zu u orthogonale Teilraum

$$W = \{u\}^\perp = \{v \in V \mid \langle u, v\rangle = 0\}$$

T-invariant, d.h. $T(W) \subset W$ (also $Tw \in W \ \forall w \in W$).

Beweis. (a) Sei $Tu = \lambda u$ mit $u \neq 0$, dann ist

$$\lambda\langle u, u\rangle = \langle u, \lambda u\rangle = \langle u, Tu\rangle = \langle Tu, u\rangle = \langle \lambda u, u\rangle = \overline{\lambda}\langle u, u\rangle\,,$$

also folgt mit $\langle u, u\rangle \neq 0$, dass $\lambda = \overline{\lambda}$.

(b) Sei $Tu = \lambda u$ und $Tv = \mu v$ für $\lambda \neq \mu$ und $u, v \neq 0$, dann ist

$$\lambda\langle u, v\rangle = \langle \lambda u, v\rangle = \langle Tu, v\rangle = \langle u, Tv\rangle = \langle u, \mu v\rangle = \mu\langle u, v\rangle ,$$

also folgt mit $\lambda \neq \mu$, dass $\langle u, v\rangle = 0$.

(c) Sei $Tu = \lambda u$ und $w \in W$, also $\langle w, u\rangle = 0$. Dann ist auch

$$\langle Tw, u\rangle = \langle w, Tu\rangle = \langle w, \lambda u\rangle = \lambda\langle w, u\rangle = 0 ,$$

also $Tw \in W$.

□

7.5 Satz. Jeder selbstadjungierte Endomorphismus T auf einem endlichdimensionalen Skalarproduktraum über $\mathbb{R}$ oder $\mathbb{C}$ besitzt wenigstens einen Eigenwert (welcher dann gemäß Satz 7.4 reell ist.)

Beweis. (a) Im Fall $\mathbb{K} = \mathbb{C}$ besitzt das charakteristische Polynom P_T von T nach dem Fundamentalsatz der Algebra eine Nullstelle $\lambda \in \mathbb{C}$. Diese ist ein Eigenwert von T.

(b) Auch im Fall $\mathbb{K} = \mathbb{R}$ hat das charakteristische Polynom P_T von T sicherlich eine Nullstelle $\lambda \in \mathbb{C}$. Diese ist Eigenwert von $A = M_{\mathcal{B}}(T)$ als Element von $\mathcal{L}(\mathbb{C}^n)$, wobei $\mathcal{B}$ eine ONB ist. Nach Satz 7.4 ist aber λ reell und somit auch Eigenwert von T.

□

7.6 Satz. **Diagonalisierbarkeit selbstadjungierter Endomorphismen**
Jeder selbstadjungierte Endomorphismus auf einem endlichdimensionalen Skalarproduktraum ist diagonalisierbar, besitzt also eine Orthonormalbasis aus Eigenvektoren.

Beweis. **Induktion über dim** $V = n$:
Im Fall $\dim V = 1$ ist jede lineare Abbildung $T : V \to V$ diagonalisierbar, denn ist $V = \mathrm{Span}\{v\}$ mit $v \neq 0$, so gibt es wegen $Tv \in V$ ein $\lambda \in \mathbb{K}$ mit $Tv = \lambda v$.
Sei also die Behauptung für alle Skalarprodukträume der Dimension n richtig, V ein $(n+1)$-dimensionaler Skalarproduktraum und $T \in \mathcal{L}(V)$ selbstadjungiert.
Gemäß Satz 7.5 hat T einen reellen Eigenwert λ_1. Sei $Tv_1 = \lambda_1 v_1$ mit $\|v_1\| = 1$ und $W_1 = \{v_1\}^\perp$. Nach Korollar 6.14 ist $\dim W_1 = \dim V - 1 = n$. Gemäß Satz 7.4 ist $T_1 = T|_{W_1}$ eine selbstadjungierte Abbildung auf dem Skalarproduktraum W_1. Nach Induktionsannahme gibt es eine ONB $v_2, \dots, v_{n+1}$ von W_1 und Zahlen $\lambda_2, \dots, \lambda_{n+1} \in \mathbb{K}$ mit

$$T_1 v_k = T v_k = \lambda_k v_k \quad \text{für} \quad k = 2, \dots, n+1 .$$

Also bilden $(v_1, \dots, v_{n+1})$ eine ONB für V mit

$$T v_k = \lambda_k v_k \quad \text{für} \quad k = 1, \dots, n+1 .$$

□

7.7 Korollar. **Hauptachsentransformation selbstadjungierter Operatoren**
Zu einem selbstadjungierten Endomorphismus $T : V \to V$ eines n-dimensionalen Skalaproduktraumes V über $\mathbb{K}$ läßt sich stets eine unitäre Abbildung

$$S : \mathbb{K}^n \to V$$

finden, welche T diagonalisiert, also

$$S^{-1}TS = \begin{pmatrix} \lambda_1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda_1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \lambda_m & 0 \\ 0 & \cdots & \cdots & 0 & \lambda_m \end{pmatrix}$$

erfüllt. Hier sind $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte von T, wobei jeder Eigenwert entsprechend seiner geometrischen Vielfachheit oft wiederholt wird.

7.8 Korollar. Spektraldarstellung selbstadjungierter Operatoren
Ist $T : V \to V$ selbstadjungiert auf dem endlichdimensionalen Skalraproduktraum V, dann gilt

$$T = \sum_{k=1}^{m} \lambda_k P_k,$$

wobei $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte sind und $P_k : V \to V$ jeweils die orthogonale Projektion auf den Eigenraum E_{λ_k} bezeichnet.

Beweis. Es reicht aus zu zeigen, dass beide Seiten auf Eigenvektoren von T diesselbe Wirkung haben, denn es gibt ja eine Basis aus Eigenvektoren. Sei also $Tv = \lambda_j v$, dann gilt

$$P_k v = \delta_{kj} v\,,$$

da Eigenvektoren zu verschiedenen Eigenwerten orthogonal sind. Also ist

$$\sum_{k=1}^{m} \lambda_k P_k v = \lambda_j v\,.$$

□

7.9 Bemerkung. Diese Spektraldarstellung läßt sich auf geeignete selbstadjungierte Operatoren auf ∞-dimensionalen Räumen verallgemeinern.

Zusammenfassend geben wir an dieser Stelle noch das Rezept zur Hauptachsentransformation von selbstadjungierten Matrizen an (bzw. Hauptachsentransformation von Operatoren: dann bestimme im nullten Schritt die Matriz bzgl. irgendeiner ONB).

Rezept zur Hauptachsentransformation von selbstadjungierten Matrizen

Sei $A \in M(n \times n, \mathbb{K})$ selbstadjungiert. Zur Bestimmung einer Hauptachsentransformation $S \in U(n)$ für A führe man die folgenden Schritte durch.

1. Schritt. Man bestimme die verschiedenen Eigenwerte $\lambda_1, \ldots, \lambda_m$ von A als die verschiedenen Nullstellen des charakteristischen Polynoms

$$P_A(\lambda) = \det(A - \lambda E_n)\,.$$

2. Schritt. Für jedes $k = 1, \ldots, m$ bestimme man eine Basis $(w_1^{(k)}, \ldots, w_{n_k}^{(k)})$ des Eigenraums E_{λ_k}, indem man das Gauß'sche Verfahren zur Lösung des Gleichungssystems $(A - \lambda_k E_n)x = 0$ anwendet.

3. Schritt. Man überführe die Basis $w_1^{(k)}, \ldots, w_{n_k}^{(k)}$ mittels des Gram-Schmidt'schen Verfahrens in eine Orthonormalbasis $(v_1^{(k)}, \dots, v_{n_k}^{(k)})$.

4. Schritt. Durch Aneinnanderreihung dieser Basen entsteht dann die ONB

$$(v_1, \dots, v_n) = (v_1^{(1)}, \dots, v_{n_1}^{(1)}, \dots, v_1^{(m)}, \dots v_{n_m}^{(m)})$$

von V aus Eigenvektoren und die Transformationsmatrix S enthält die Basisvektoren $(v_1, \dots, v_n)$ als Spalten.

7.10 Definition. **Quadratische Form**

Eine symmetrische Sesquilinearform auf einem $\mathbb{K}$-Vektorraum heißt quadratische Form. D.h.

$$Q : V \times V \to \mathbb{K}$$

ist eine quadratische Form, falls

(a) $\quad Q(u, \alpha v + w) = \alpha Q(u, v) + Q(u, w)$ für alle $u, v, w \in V$ und $\alpha \in \mathbb{K}$.

(b) $\quad Q(u, v) = \overline{Q(v, u)}$ für alle $u, v, \in V$.

Für eine quadratische Form Q gilt zwar aufgrund der Symmetrie $Q(v, v) \in \mathbb{R}$, aber nicht wie beim Skalarprodukt $Q(v, v) \geq 0$ bzw. $Q(v, v) = 0 \Leftrightarrow v = 0$.

7.11 Satz. Sei V ein Skalarproduktraum.

(a) Zu jedem symmetrischen Operator T auf V wird durch

$$Q_T(u, v) := \langle u, Tv \rangle$$

eine quadratische Form definiert.

(b) Falls $\dim V < \infty$, dann gibt es zu jeder quadratischen Form Q auf V einen eindeutigen symmetrischen Operator $T \in \mathcal{L}(V)$ mit

$$Q(u, v) = \langle u, Tv \rangle \qquad \forall\, u, v \in V\,.$$

Beweis. (a) Linearität von Q_T im zweiten Argument ist offensichtlich und die Symmetrie ebenfalls:

$$Q_T(u, v) = \langle Tu, v \rangle = \overline{\langle v, Tu \rangle} = \overline{Q_T(v, u)}\,.$$

(b) Sei eine ONB $\mathcal{A} = (v_1, \dots, v_n)$ von V gegeben. Dann hat der Operator definiert durch

$$Tv_k = \sum_{j=1}^{n} Q(v_j, v_k) v_j =: \sum_{j=1}^{n} a_{jk} v_j$$

die gewünschten Eigenschaften: T ist symmetrisch, denn die Matrix $M_{\mathcal{A}}(T) = (a_{jk})$ ist selbstadjungiert,

$$a_{jk} = Q(v_j, v_k) = \overline{Q(v_k, v_j)} = \overline{a_{kj}}\,,$$

und es gilt

$$\langle v_m, Tv_k \rangle = Q(v_m, v_k)$$

nach Definition. Wegen Sesqulinearität auf beiden Seiten folgt

$$\langle v, Tu \rangle = Q(v, u) \qquad \forall\, u, v \in V\,.$$

Die Eindeutigkeit folgt wieder aus der Definitheit des Skalarprodukts:

$$\langle u, Tv \rangle = \langle u, \tilde{T}v \rangle \quad \forall\, u \in V \quad \Rightarrow \quad Tv = \tilde{T}v\,.$$

□

7.12 Satz. und Definition. **Hauptachsentransformation quadratischer Formen**
Sei V ein n-dimensionaler Skalarproduktraum und Q eine quadratische Form , also

$$Q(u,v) = \langle u, Tv\rangle$$

mit einem symmetrischen Operator T. Nach Satz 7.5 gibt es eine ONB $\mathcal{A} = (v_1, \ldots, v_n)$ aus Eigenvektoren von T:

$$Tv_1 = \lambda_1 v_1, \ldots, Tv_n = \lambda_n v_n\,.$$

Dann hat Q bezüglich $\mathcal{A}$ die Hauptachsendarstellung

$$Q(u) := Q(u,u) = \sum_{j=1}^{n} \lambda_j |\langle v_j, u\rangle|^2\,.$$

Die Geraden Span$\{v_j\}$ heißen die Hauptachsen von Q.

7.13 Korollar. **Rayleigh-Ritz-Prinzip**
Seien die Eigenwerte $\lambda_1 < \lambda_2 < \cdots < \lambda_m$ des symmetrischen Operators T der Größe nach geordnet, so gilt

$$\lambda_1 = \min\{Q_T(u) \,|\, \|u\| = 1\} \quad \text{und} \quad \lambda_m = \max\{Q_T(u) \,|\, \|u\| = 1\}\,.$$

Beweis. Übungsaufgabe. □

7.14 Definition. **Positive Operatoren und Matrizen**
Ein symmetrischer Operator T heißt

- **positiv** und wir schreiben $T \geq 0$, falls $\langle u, Tu\rangle \geq 0 \quad \forall u \in V$.
- **positiv definit** und wir schreiben $T > 0$, falls $\langle u, Tu\rangle > 0 \quad \forall u \neq 0$.

Entsprechend definiert man positiv (definite) quadratische Formen und Matrizen, bzw. negative und negativ definite Operatoren.

7.15 Satz. In einem endlichdimensionalen Skalarproduktraum gilt

- $T \geq 0 \Leftrightarrow$ alle Eigenwerte von T sind nichtnegativ.
- $T > 0 \Leftrightarrow$ alle Eigenwerte von T sind positiv.

Beweis. Folgt sofort aus der Darstellung

$$\langle u, Tu\rangle = \sum_{k=1}^{n} \lambda_k \, |\langle v_k, u\rangle|^2$$

aus Satz 7.12. □

7.16 Beispiel. **Die Matrix der Gaußschen Normalengleichung (vgl. Ende Kapitel 3)**
Zu lösen ist

$$Ax = y$$

mit $A \in M(m \times n, \mathbb{R})$ und $y \in \mathbb{R}^m$ bekannt und $x \in \mathbb{R}^n$ gesucht, wobei Rang$A = n < m$. Im Allgemeinen ist $y \notin$ Bild$A$ und somit hat die Gleichung keine Lösung. Deshalb such man die beste Approximation, d.h. ein u das $\|Au - y\|$ minimiert. Am Ende von Kapitel 3 haben wir gezeigt, dass ein solches U die Gleichung $A^T Au = A^T y$ lösen muss. D.h., $B = A^T A$ ist zu invertieren. Es ist $B \in M(n \times n, \mathbb{R})$ positiv definit, denn

$$\langle x, Bx\rangle_{\mathbb{R}^n} = \langle x, A^T Ax\rangle_{\mathbb{R}^n} = \langle x, A^* Ax\rangle_{\mathbb{R}^n} = \langle Ax, Ax\rangle_{\mathbb{R}^m} = \|Ax\|^2_{\mathbb{R}^m} \geq 0\,.$$

Da A vollen Rang hat, also Rang $A = n$, ist Kern$A = \{0\}$ und somit

$$\langle x, Bx\rangle_{\mathbb{R}^n} = \|Ax\|^2_{\mathbb{R}^m} > 0 \quad \forall\, x \neq 0\,.$$

Nun sind aber positiv definite Matrizen insbesonder injektiv und somit invertierbar. Also haben die Gaußschen Normalengleichungen immer eine eindeutige Lösung. Dieselbe Lösung erhält man, wenn man $Au = P_{\mathrm{Bild}A}y$ löst, wobei $P_{\mathrm{Bild}A}$ die orthogonale Projektion auf das Bild von A ist.

7.17 Bemerkung. **Quadratische Formen und Kegelschnitte**
Die Lösungsmenge der Gleichung

$$\langle x, Ax\rangle + \langle b, x\rangle = c \qquad\qquad x, b \in \mathbb{R}^2,\ c \in \mathbb{R},\ A \in M(2\times 2, \mathbb{R})$$

beschreibt einen Kegelschnitt, denn nach Hauptachsentransformation $S \in O(2)$ von A (Drehung des Koordinatensystems) gilt mit

$$S^{-1}AS = D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \qquad y = S^{-1}x \qquad \text{und} \qquad \tilde{b} = S^{-1}b\,,$$

dass

$$\langle x, Ax\rangle + \langle b, x\rangle = c \quad \Leftrightarrow \quad \langle y, Dy\rangle + \langle \tilde{b}, y\rangle = c = \lambda_1 {y_1}^2 + \lambda_2 {y_2}^2 + \tilde{b}_1 y_1 + \tilde{b}_2 y_2\,.$$

Als Lösungsmenge hat man beispielsweise für $c > 0$ und

(a) $A > 0$ und $b = 0$, also $\lambda_1, \lambda_2 > 0$, eine Ellipse mit Hauptachsen $\lambda_1^{-1/2} Se_1$ und $\lambda_2^{-1/2} Se_2$,
(b) $\lambda_1 \cdot \lambda_2 < 0$ und $b = 0$ eine Hyperbel,
(c) $\lambda_1 > 0$, $\lambda_2 = 0$ und $b \neq 0$ eine Parabel, bzw. für $b = 0$ ein Geradenpaar,
(d) $A < 0$ und $b = 0$ die leere Menge.

7.18 Definition. **Funktionalkalkül für selbstadjungierte Operatoren**
Sei $T \in \mathcal{L}(V)$ selbstadjungiert, dim$V = n < \infty$ mit Eigenwerten $\lambda_1, \ldots, \lambda_n$. Für jede Funktion $f : \mathbb{R} \supset U \to \mathbb{K}$ mit $\{\lambda_1, \ldots, \lambda_n\} \subset U$ definiert man

$$f(T) : V \to V\ ,\ v \mapsto f(T)v := S \begin{pmatrix} f(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & f(\lambda_n) \end{pmatrix} S^{-1}v\,.$$

7.19 Beispiel. **Wurzel eines positiven Operators**
Sei $A \in \mathcal{L}(V)$ positiv und selbstadjungiert. Dann ist

$$\sqrt{A} := S \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} S^{-1}$$

selbst wieder positiv und es gilt $\sqrt{A} \cdot \sqrt{A} = A$.

8 Klassifikation von Matrizen

Wir führen nun die Begriffe der Äquivalenzrelation und der zugehörigen Äquivalenzklassen ein. Ziel des Folgenden ist eine mathematische Präzisierung der Idee, dass verschiedenen Objekte im Bezug auf bestimmte Eigenschaften gleich sind. Beispielsweise sind alle Vektoräume der endlichen Dimension n zueinander isomorph.

8.1 Definition. **Äquivalenzrelation**

Sei M eine Menge. Unter einer **Äquivalenzrelation** auf M versteht man eine Relation (d.h. eine Teilmenge $R \subset M \times M$, wobei man statt $(x, y) \in R$ auch $x \sim y$ schreibt), welche die folgenden Eigenschaften hat.

(a) Reflexivität: $x \sim x$ für alle $x \in M$

(b) Symmetrie: $x \sim y \Leftrightarrow y \sim x$

(c) Transitivität: $x \sim y$ und $y \sim z \Rightarrow x \sim z$

8.2 Beispiel. Sei V ein Vektorraum über $\mathbb{K}$ und $U \subset V$ ein Unterraum. Wir definieren $\sim_U$ auf V durch

$$x \sim_U y \quad \Leftrightarrow \quad x - y \in U\,.$$

Dann ist $\sim_U$ eine Äquivalenzrelation.

Beweis. (a) $x \sim_U x$, denn $x - x = 0 \in U$ für alle $x \in V$.

(b) $x \sim_U y \Leftrightarrow x - y \in U \Leftrightarrow y - x \in U \Leftrightarrow y \sim_U x$

(c) $x \sim_U y$ und $y \sim_U z \Rightarrow x - y \in U$ und $y - z \in U \Rightarrow x - z = (x - y) + (y - z) \in U \Rightarrow x - z \in U$.

□

8.3 Definition. **Äquivalenzklassen**

Sei $\sim$ eine Äquivalenzrelation auf M. Für $x \in M$ heißt die Teilmenge

$$[x] := \{y \in M \,|\, x \sim y\} \subset M$$

die **Äquivalenzklasse** von x bezüglich $\sim$.

8.4 Satz. Sei $\sim$ eine Äquivalenzrelation auf M. Dann gilt

(a) $M = \cup_{x \in M}[x]$

(b) $[x] \cap [y] \neq \emptyset \Leftrightarrow x \sim y \Leftrightarrow [x] = [y]$.

Beweis. (a) ist trivial, da $x \in [x]$. (b) ist es auch, aber das muss man sich zumindest einmal klar machen. Dazu findet man

$$[x] \cap [y] \neq \emptyset \quad \Leftrightarrow \quad \exists z \in [x] \cap [y] \quad \Leftrightarrow \quad \exists z \in M \text{ mit } z \sim x \text{ und } z \sim y \quad \Leftrightarrow \quad x \sim y\,.$$

Sei nun $x \sim y$ und $z \in [x]$, also $z \sim x$. Dann ist auch $z \sim y$ und somit $z \in [y]$, also $[x] \subset [y]$, und analog findet man $[y] \subset [x]$. Ist umgekehrt $[x] = [y]$, so ist insbesondere $y \in [x]$, also $x \sim y$. □

8.5 Definition. Eine Menge von nichtleeren Teilmengen von M die paarweise diskjunkt sind und deren Vereinigung ganz M ergibt, nennt man Zerlegung von M.

Die Menge der Äquivalenzklassen $\{[x] \,|\, x \in M\}$ ist also eine Zerlegung von M.

8.6 Definition. Quotient bzgl. einer Äquivalenzrelation
Ist $\sim$ eine Äquivalenzrelation auf M, so nennt man die Menge

$$M/\!\sim \;:=\; \{[x] \,|\, x \in M\}$$

der Äquivalenzklassen den **Quotienten** von M nach $\sim$. Die Abbildung

$$\pi_\sim : M \to M/\!\sim\,, \quad x \mapsto [x]$$

heißt die kanonische Projektion der Äquivalenzrelation.

Man kann also die Elemente der Menge M dadurch klassifizieren, dass man eine Äquivalenzrelation $\sim$ auf M angibt. Geht man nun zum Quotienten $M/\!\sim$ über, so betrachtet man diejenigen Elemente der Menge M als im wesentlichen gleich, welche unter $\sim$ äquivalent sind.

8.7 Beispiel. Auf dem $\mathbb{R}^3$ wird durch

$$x \sim y \quad \Leftrightarrow \quad x_3 = y_3$$

eine Äquivalenzrelation definiert. Es ist $\mathbb{R}^3/\!\sim$ isomorph zu $\mathrm{Span}\{e_3\}$ wenn wir zu jedem $[x] \in \mathbb{R}^3/\!\sim$ den Vertreter $y \in [x]$ mit $y_1 = y_2 = 0$ wählen. Die kanonische Projektion $\pi_\sim$ ist dann einfach die orthogonale Projektion auf $\mathrm{Span}\{e_3\}$. Diese Charakterisierung von $\mathbb{R}^3/\!\sim$ ist eine Klassifikation durch Repräsentanten.

Sei $\sim$ eine Äquivalenzrlation auf einer Menge M. Die Menge M "nach $\sim$" oder "bis auf $\sim$" zu klassifizieren heißt, $M/\!\sim$ und möglichst auch $\pi_\sim$ zu "verstehen". Zwei häufige Varianten dieses zunächst etwas vagen Konzepts sind

(a) **Klassifikation durch Repräsentanten.** Die Klassifikation durch Repräsentanten besteht darin, eine "überschaubare" Teilmenge $M_0 \subset M$ anzugeben, so dass

$$\pi|_{M_0} : M_0 \to M/\!\sim$$

bijektiv ist, es also zu jedem $x \in M$ genau ein $x_0 \in M_0$ mit $x \sim x_0$ gibt.
Besipiel 8.7: $M = \mathbb{R}^3$, $M_0 = \mathrm{Span}\{e_3\}$.

(b) **Klassifikation durch charakteristische Daten.** Die Klassifikation durch charakteristische Daten besteht darin, eine "wohlbekannte" Menge D und eine Abbildung $c : M \to D$ zu finden, so dass

$$x \sim y \quad \Leftrightarrow \quad c(x) = c(y)$$

gilt. Die Abbildung $c : M/\!\sim \to D$, $[x] \mapsto c(x)$ ist dann eine Bijektion. Typische Mengen D die als charakteristische Daten Verwendung finden sind $\mathbb{Z}$, $\mathbb{N}$ oder Polynome.

8.8 Beispiel. Sei V eine unendliche Menge und M die Menge aller endlichen Teilmengen von V. Für $X, Y \in M$ sei

$$X \sim Y \quad \Leftrightarrow \quad \text{es gibt eine Bijektion } f : X \to Y\,.$$

Dann erhält man eine Klassifikation von M nach $\sim$ durch charakteristische Daten, indem man setzt: $D := \mathbb{N}$ und $c(X) := |X| =$ Anzahl der Elemente von X.

8.9 Definition. Zwei Matrizen $A, B \in M(m \times n, \mathbb{K})$ heißen äquivalent, geschrieben $A \sim B$, wenn es invertierbare Matrizen P und Q gibt, sodass

$$B = Q^{-1}AP$$

gilt. Offensichtlich ist hierdurch eine Äquivalenzrelation auf $M(m \times n, \mathbb{K})$ erklärt:

(a) $A = E_m A E_n$

(b) $B = Q^{-1}AP \Leftrightarrow A = QBP^{-1}$

(c) $B = Q^{-1}AP$, $C = R^{-1}BS \Rightarrow C = R^{-1}Q^{-1}APS = (QR)^{-1}A(PS)$.

8.10 Satz. Zwei $m \times n$-Matrizen sind genau dann äquivalent, wenn sie denselben Rang haben.

Beweis. "$\Rightarrow$": Sei $A \sim B$, dann ist dim(Bild$B$)=dim(Bild$Q^{-1}AP$)=dim(Bild$A$), da Q^{-1} und P Isomorphismen sind.
"$\Leftarrow$": Sei Rang$A$ =Rang$B = r$. Sei $(v_1, \dots, v_{n-r})$ Basis von Kern$B$ und $(v_1, \dots, v_n)$ Basis von ganz $\mathbb{K}^n$. Dann ist $(w_1, \dots, w_r) := (Bv_{n-r+1}, \dots, Bv_n)$ eine Basis von Bild$B$, die wir zu einer Basis $(w_1, \dots, w_m)$ von ganz $\mathbb{K}^m$ ergänzen. Analog verfahren wir für A und erhalten Basen $(v'_1, \dots, v'_n)$ für $\mathbb{K}^n$ und $(w'_1, \dots, w'_m)$ für $\mathbb{K}^m$. Seien nun P und Q die Isomorphismen, welche die ungestrichenen in die gestrichenen Basen überführen. Dann gilt

$$QB = AP$$

für die Basisvektoren $v_1, \dots, v_n$ und somit für all $v \in \mathbb{K}^n$. $\square$

Der Rang ist also ein charakteristisches Datum für die Klassifikation der $m \times n$-Matrizen bis auf Äquivalenz, d.h.

$$M(m \times n, \mathbb{K})/\!\sim\ \cong \{0, 1, 2, \dots, \min(m, n)\}\,.$$

Man erhält aber auch leicht eine Klassifikation durch Repräsentanten:

$$E_r^{m\times n} = \left.\underbrace{\begin{pmatrix} \overbrace{\begin{matrix}1 & & 0\\ & \ddots & \\ 0 & & 1\end{matrix}}^{r} & 0 \\ 0 & 0 \end{pmatrix}}_{n}\right\}m$$

Man nennt die einfachsten Repräsentanten einer Äquivalenzklasse oft auch Normalformen.

Äquivalenz von Matrizen ist offenbar eine sehr grobe Klassifikation. Eine feinere ist Ähnlichkeit.

8.11 Definition. Zwei Matrizen $A, B \in M(n \times n, \mathbb{K})$ heißen **ähnlich**, wenn es eine reguläre $n \times n$-Matrix P gibt so, dass

$$B = P^{-1}AP$$

gilt.

Auch Ähnlichkeit ist wieder eine Äquivalenzrelation und ähnliche Matrizen sind sicherlich äquivalent im Sinne von Definition 8.9. Umgekehrt sind sicherlich nicht alle äquivalenten Matrizen ähnlich, da z.B. ähnliche Matrizen jeweils das gleiche charakteristische Polynom haben. Wir führen nun die Normalformenklassifikation ähnlicher Matrizen ein, die so genannte Jordansche Normalform.

8.12 Definition. Sei $z \in \mathbb{C}$ und $m \geq 1$. Die $m \times m$-Matrix

$$J_m(z) := \begin{pmatrix} z & 1 & 0 & \cdots & 0 \\ 0 & z & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & z & 1 \\ 0 & \cdots & \cdots & 0 & z \end{pmatrix}$$

heißt das Jordankästchen der Größe m zum Eigenwert z.

$J_m(z)$ hat den einzigen Eigenwert z mit algebraischer Vielfachheit m und geometrischer Vielfachheit 1: Das charakteristische Polynom von $J_m(z)$ ist

$$P_{J_m(z)}(\lambda) = (z - \lambda)^m$$

und der einzige Eigenvektor ist $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, denn

$$J_m(z) - zE_m = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & 1 \\ 0 & \cdots & \cdots & 0 & 0 \end{pmatrix}.$$

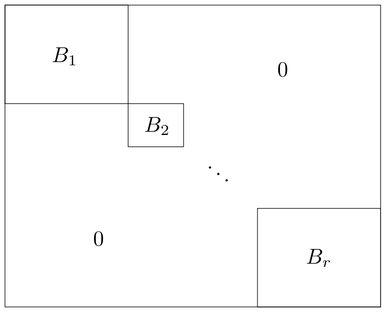
Für $m \geq 2$ ist $J_m(z)$ also so nichtdiagonalisierbar, wie eine komplexe $m \times m$-Matrix nur sein kann.

Die Jordansche Normalform für eine komplexe $n \times n$-Matrix A hat nun folgende Form. Zu jedem Eingenwert λ_k gehört eine Blockmatrix B_k aus n_k einzelnen Jordankästchen:

$$B_k = \begin{array}{|c|c|c|c|}\hline \begin{array}{cccc} \lambda_k & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_k \end{array} & & 0 & \\ \hline & \ddots & & \\ \hline & & \ddots & \\ \hline & 0 & & \begin{array}{cccc} \lambda_k & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_k \end{array} \\ \hline \end{array} \begin{array}{l} \left.\vphantom{\begin{array}{c}1\\1\\1\\1\end{array}}\right\} m_1^{(k)} \\ \left.\vphantom{1}\right\} m_2^{(k)} \\ \\ \left.\vphantom{\begin{array}{c}1\\1\\1\\1\end{array}}\right\} m_{n_k}^{(k)} \end{array}$$

Die algebraische Vielfachheit von λ_k ist dabei $\sum_{j=1}^{n_k} m_j^{(k)}$ und die geometrische Vielfachheit ist n_k.

Die gesamte Jordansche Normalform von A ist dann



8.13 Satz. **<u>Jordansche Normalform</u>**
Sei $A \in M(n \times n, \mathbb{C})$ mit verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$ mit geometrischen Vielfachheiten $n_1, \ldots, n_r$ und algebraischen Vielfachheiten $\ell_1, \ldots, \ell_r$.
Dann gibt es zu jedem $k = 1, \ldots, r$ eindeutig bestimmte Zahlen

$$m_1^{(k)} \leq m_2^{(k)} \leq \cdots \leq m_{n_k}^{(k)}, \quad \sum_{j=1}^{n_k} m_j^{(k)} = \ell_k$$

mit der Eigenschaft, dass es eine invertierbare Matrix $P \in M(n \times n, \mathbb{C})$ gibt, für die $P^{-1}AP$ die Blockmatrix ist, welche durch Aneinanderreihung der Jordan Kästchen

$$J_{m_1^{(1)}}(\lambda_1), \ldots, J_{m_{n_1}^{(1)}}(\lambda_1), J_{m_1^{(2)}}(\lambda_2), \ldots, J_{m_{n_2}^{(2)}}(\lambda_2), \ldots, J_{m_1^{(r)}}(\lambda_r), \ldots, J_{m_{n_r}^{(r)}}(\lambda_r)$$

längs der Diagonalen entsteht.

Bis auf die Reihenfolge der Eigenwerte liefert die Jordansche Normalform also eine Klassifikation ähnlicher Matrizen durch Repräsentanten.

8.14 Bemerkung. Eine Matrix ist also genau dann diagonalisierbar, wenn alle Jordankästchen die Größe 1 haben.

8.15 Beispiel. Bei der Lösung des gedämpften harmonischen Oszillators, vgl. Beispiel 5.23, standen wir vor dem Problem eine nichtdiagonalisierbare 2×2-Matrix zu exponentieren. Gemäß Satz 8.13 können wir diese auf die Form

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

bringen. Dann gilt aber wegen

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix},$$

dass

$$\exp\left[\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} t\right] := \sum_{n=0}^{\infty} \frac{t^n}{n!} \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix} = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix}.$$

8.16 Bemerkung. Für Diagonalmatrizen

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

und Funktionen $f : \mathbb{K} \to \mathbb{K}$ hatten wir

$$f(D) := \begin{pmatrix} f(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & f(\lambda_n) \end{pmatrix}$$

definiert. In Beispiel 8.15 hatten wir gesehen, dass man auch Jordankästchen exponentieren kann. Wir wollen nochetwas allgemeiner untersuchen, inwiefern man Funktionen von Jordankästchen bilden kann. Dazu setzen wir

$$J_m(\lambda) =: \lambda E_m + N_m$$

mit der **nilpotenten** Matrix

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

Eine Matrix oder ein Operator A heißt nilpotent, falls $A^n = 0$ für ein $n \in \mathbb{N}$. Im Fall von N_m gilt

$$((N_m)^n)_{ij} = \begin{cases} 1 & \text{falls } j - i = n \\ 0 & \text{sonst.} \end{cases} \qquad \text{für } n \in \mathbb{N}_0,$$

also insbesondere $(N_m)^n = 0$ für $n \geq m$. Damit ist

$$J_m(\lambda)^n = (\lambda E_m + N_m)^n = \sum_{k=0}^{n} \binom{n}{k} \lambda^{n-k} N_m^k = \sum_{k=0}^{m-1} \binom{n}{k} \lambda^{n-k} N_m^k,$$

also beispielsweise

$$\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}^{23} = \begin{pmatrix} \lambda^{23} & 23 \cdot \lambda^{22} & \frac{23 \cdot 22}{2!} \lambda^{21} \\ 0 & \lambda^{23} & 23 \cdot \lambda^{22} \\ 0 & 0 & \lambda^{23} \end{pmatrix}.$$

Für Polynome bzw. bei 0 analytische Funktionen f (d.h. Funktionen die sich als konvergente Potenzreihe um 0 schreiben lassen) findet man allgemeiner

$$f(J_m(\lambda)) = \begin{pmatrix} f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2!} & \cdots & \frac{f^{(m-1)}(\lambda)}{(m-1)!} \\ 0 & f(\lambda) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \frac{f''(\lambda)}{2!} \\ \vdots & & \ddots & f(\lambda) & f'(\lambda) \\ 0 & \cdots & \cdots & 0 & f(\lambda) \end{pmatrix}.$$

Schließlich können wir nun auch noch den Beweis des Satzes von Cayley-Hamilton nachliefern:

8.17 Bemerkung. Beweis des Satzes von Cayley-Hamilton (Satz 5.22)

Sei $\mathcal{B}$ eine Basis in der $M_{\mathcal{B}}(L)$ Jordansche Normalform hat. Es genügt nun, jeden Jordanblock für sich zu betrachten. Ist $J_m(\lambda_j)$ ein Block in der Jordanschen Normalform, so ist λ_j sicherlich Eigenwert der algebraischen Vielfachheit mindestens m. Das charakteristische Polynom enthält also einen Faktor $(\lambda - \lambda_j)^m$. Da

$$(J_m(\lambda_j) - \lambda_j E_m)^m = N_m^m = 0,$$

folgt $P_L(L) = 0$.

Auch die Hauptachsentransformation selbstadjungierter Matrizen liefert eine Klassifikation durch Repräsentanten.

8.18 Definition. Es bezeichne $\mathrm{Sym}(n, \mathbb{K})$ den Vektorraum der selbstadjungierten $n \times n$-Matrizen. Zwei Matrizen heißen **orthogonal ähnlich**, wenn es eine orthogonale Matrix $P \in U(n)$ gibt, so dass

$$B = P^{-1}AP$$

gilt.

8.19 Satz. Jede selbstadjungierte $n \times n$-Matrix ist bis auf die Reihenfolge der Eigenwerte zu genau einer Diagonalmatrix

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

orthogonal ähnlich.

Der Weg zur Jordanschen Normalform: eine Beweisskizze von Satz 8.13

1. Schritt: Jede komplexe $n \times n$-Matrix A läßt sich durch eine unitäre Ähnlichkeitstransformation Q auf obere Dreiecksform

$$R = \begin{pmatrix} \lambda_1 & r_{12} & \cdots & r_{1n} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & r_{n-1,n} \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} = Q^{-1}AQ$$

bringen, wobei $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von A, wiederholt gemäß ihrer algebraischen Vielfachheit, sind.

Beweis. Induktion über n, wobei $n = 1$ klar ist. Es gelte die Aussage also für $n-1$ und es sei $A \in M(n, \mathbb{C})$. Dann hat A mindestens einen Eigenvektor v und es gilt $Av = \lambda v$. Sei T ein Basiswechsel zu einer ONB $(v, w_1, \ldots, w_{n-1})$, dann ist

$$\tilde{A} = T^{-1}AT = \begin{pmatrix} \lambda & x^T \\ 0 & B \end{pmatrix}$$

mit $x \in \mathbb{C}^{n-1}$. Nach Induktionsannahme gibt es für $B \in M(n-1, \mathbb{C})$ eine unitäre Abbildung $\tilde{Q}$ mit

$$\tilde{Q}^{-1}B\tilde{Q} = R_{n-1}\,.$$

Sei $Q := T\begin{pmatrix} 1 & 0 \\ 0 & \tilde{Q} \end{pmatrix}$, dann ist

$$\begin{aligned} Q^{-1}AQ &= \begin{pmatrix} 1 & 0 \\ 0 & \tilde{Q}^{-1} \end{pmatrix} T^{-1}AT \begin{pmatrix} 1 & 0 \\ 0 & \tilde{Q} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{Q}^{-1} \end{pmatrix} \begin{pmatrix} \lambda & x^T\tilde{Q} \\ 0 & B\tilde{Q} \end{pmatrix} \\ &= \begin{pmatrix} \lambda & x^T\tilde{Q} \\ 0 & \tilde{Q}^{-1}B\tilde{Q} \end{pmatrix} = \begin{pmatrix} \lambda & x^T\tilde{Q} \\ 0 & R_{n-1} \end{pmatrix} = R_n\,. \end{aligned}$$

□

2. Schritt: Hauptvektoren

8.20 Definition. Sei λ Eigenwert von $A \in M(n, \mathbb{C})$ mit algebraischer Vielfachheit m, dann heißt jeder Vektor v mit

$$(A - \lambda E_n)^m v = 0\,, \quad v \neq 0\,,$$

Hauptvektor zum Eigenwert λ.

8.21 Satz. Alle Hauptvektoren zu einem Eigenwert λ von $A \in M(n, \mathbb{C})$ mit algebraischer Vielfachheit m bilden einen m-dimensionalen Unterraum H_λ, den so genannten Hauptraum des Eigenwertes. Dieser ist invariant unter A. Der Gesamtraum ist die direkte Summe der Haupträume,

$$\mathbb{C}^n = \bigoplus_\lambda H_\lambda \, .$$

8.22 Definition. Man sagt eine Summe $\sum_\lambda H_\lambda = \mathrm{Span}\{\cup_\lambda H_\lambda\}$ ist direkt und schreibt $\bigoplus_\lambda H_\lambda$, falls sich jeder Vektor $v \in \sum_\lambda H_\lambda$ in eindeutiger Weise als Linearkombination $v = \sum_\lambda v_\lambda$ mit $v_\lambda \in H_\lambda$ schreiben läßt.

Beweis. (von Satz 8.21) Sei $v \in H_\lambda$ also $(A-\lambda)^m v = 0$. Dann ist auch $Av \in H_\lambda$, denn

$$(A_\lambda)^m Av = (A_\lambda)^m (A-\lambda)v = (A_\lambda)(A-\lambda)^m v = 0 \, .$$

Um $\mathrm{dim} H_\lambda$ zu bestimmen, transformiert man A auf obere Dreiecksform:

$$A \to B = Q^{-1}AQ = \begin{pmatrix} \lambda & & & * \\ & \ddots & & \\ & & \lambda & \\ 0 & & & R \end{pmatrix}$$

Es ist klar, dass $(A-\lambda)^m v = 0$ genau dann, wenn $(B-\lambda)^m Q^{-1} v = 0$, die Haupträume von A und B zu λ also die gleiche Dimension haben. Nun sind die Vektoren e_i, $i = 1, \ldots, m$, genau die Hauptvektoren von B, also $\mathrm{dim} H_\lambda = m$. Da $\sum_\lambda \mathrm{dim} H_\lambda = \sum_\lambda m_\lambda = n$ ist, genügt es, die lineare Unabhängigkeit der Haupträume zu zeigen. Das ist eine Übungsaufgabe. □

3. Schritt: Zyklische Teilräume
Jeder Hauptraum H_λ einer Matrix $A \in M(n, \mathbb{C})$ läßt sich in eine direkte Summe zyklischer Teilräume zerlegen:

$$H_\lambda = V_1 \oplus V_2 \oplus \cdots \oplus V_\ell \, . \tag{8.1}$$

Ein Teilraum V heißt hier zyklisch, falls es einen Vektor $v \in V$ gibt (einen zyklischen Vektor), so dass die Vektoren

$$v,\ (A-\lambda E)v,\ (A-\lambda E)^2 v, \ldots,\ (A-\lambda E)^k v,$$

eine Basis von V bilden und $w := (A-\lambda E)^k v$ Eigenvektor von A zum Eigenwert λ ist. D.h., jeder der Teilräume V_i in (8.1) besitzt eine Basis der Form

$$B^{k_i} v_i, B^{k_i - 1} v_i, \ldots, v_i \, ,$$

wobei wir $B = (A - \lambda E)$ abkürzen und $w_i := B^{k_i} v_i$ Eigenvektor von A zum Eigenwert λ ist. Es ist dann $\mathrm{dim} V_i = k_i$ und $\sum_{i=1}^{\ell} k_i = \mathrm{dim} H_\lambda$.
Es läßt A die zyklischen Teilräume V_i invariant und hat dort die Matrixdarstellung

$$J_{k_i}(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix} .$$

Beweis. (Skizze) Um die geeigneten Eigenvektoren $w_i \in \mathrm{Kern} B$ zu finden, betrachtet man die Teilräume

$$U_j := \mathrm{Kern} B \cap \mathrm{Bild} B^j \subset \mathrm{Kern} B = E_\lambda .$$

Offensichtlich ist

$$\{0\} \subset U_m \subset U_{m-1} \subset \cdots \subset U_0 = \mathrm{Kern} B .$$

Man bestimmt nun die w_i induktiv, ausgehend von der leeren Menge. Sind $w_1, \ldots, w_{i-1}$ festgelegt, so nimmt man das größte k_i, für welches U_{k_i} nicht in $\mathrm{Span}\{w_1, \ldots, w_{i-1}\}$ enthalten ist, es also ein $v_i \neq 0$ gibt mit

$$B^{k_i} v_i = w_i \notin \mathrm{Span}\{w_1, \ldots, w_{i-1}\}, \; B w_i = 0 .$$

So entsteht eine Basis $\{w_1, \ldots, w_\ell\}$ von E_λ und zyklische Basen der zugehörigen V_i. Jetzt muss man noch

$$H_\lambda = V_1 \oplus \cdots \oplus V_\ell$$

zeigen.

(i) Die lineare Unabhängigkeit zeigen wir nur am Beispiel zweier zyklischer Ketten

$$B v_1, v_1, B v_2, v_2 .$$

Sei

$$\alpha_1 B v_1 + \alpha_2 v_1 + \beta_1 B v_2 + \beta_2 v_2 = 0 ,$$

dann liefert Multiplikation mit B

$$\alpha_2 \underbrace{B v_1}_{w_1} + \beta_2 \underbrace{B v_2}_{w_2} = 0 .$$

Es sind aber w_1 und w_2 nach Konstruktion linear unabhängig, also $\alpha_2 = \beta_2 = 0$ und somit auch $\alpha_1 = \beta_1 = 0$.

(ii) Es bleibt zu zeigen, dass jeder Vektor $u \in H_\lambda$ durch die Vektoren $B^j v_i$ darstellbar ist. Sei also $B^k u = 0$ für ein $k \leq m$. Für $k = 1$ bilden die w_i nach Konstruktion eine Basis für den Eigenraum $\mathrm{Kern} B$. Seien also die Vektoren in $\mathrm{Kern} B^{k-1}$ darstellbar und $u \in \mathrm{Kern} B^k$. Es ist

$$\underbrace{B^{k-1} u}_{\in \mathrm{Kern} B} = \sum_i \alpha_i w_i = \sum_i \alpha_i B^{k_i} v_i ,$$

wobei $\alpha_i = 0$ falls $k_i < k - 1$. Damit ergibt sich

$$B^{k-1} \underbrace{\left(u - \sum_i \alpha_i B^{k_i - k + 1} v_i \right)}_{u'} = 0 .$$

Also ist u' darstellbar und somit auch u.

□

8.23 Beispiel. Für eine komplexe 3×3-Matrix gibt es die folgenden qualitativ verschiedenen Jordanschen Normalformen $J = Q^{-1} A Q$.

(i) Paarweise verschiedenen Eigenwerte λ, ρ, σ:

$$J = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \rho & 0 \\ 0 & 0 & \sigma \end{pmatrix} .$$

(ii) Ein doppelter Eigenwert λ:

$$J = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \rho \end{pmatrix} \quad \text{oder} \quad J = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \rho \end{pmatrix}.$$

Im zweiten Fall ist $Q = (w_\lambda, v_\lambda, w_\rho)$ mit Eigenvektoren w_λ, w_ρ welche man aus

$$(A-\lambda)w_\lambda \quad \text{und} \quad (A-\rho)w_\rho = 0$$

erhält. Den Hauptvektor v_λ bekommt man als Lösung von

$$(A-\lambda)v_\lambda = w_\lambda\,.$$

(iii) Dreifacher Eigenwert λ:

$$J = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}, \quad \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}, \quad \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

Im zweiten Fall bestimmt man zunächst einen Hauptvektor v durch Lösen von

$$(A-\lambda)v \neq 0\,.$$

Der zugehörige Eigenvektor ist $w = (A-\lambda)v$. Schließlich bestimmt man noch einen weiteren linear unabhängigen Eigenvektor aus

$$(A-\lambda)\tilde{w} = 0 \quad \text{mit} \quad \langle w, \tilde{w}\rangle = 0\,.$$

Dann ist $Q = (w, v, \tilde{w})$.
Im dritten Fall ist $Q = (w, v, u)$ mit

$$(A-\lambda)w = 0\,, \quad (A-\lambda)v = w\,, \quad (A-\lambda)u = v\,.$$

8.24 Beispiel. Schließlich können wir nun den Fall (ii) aus Beispiel 5.23 behandeln. Ziel war es, die lineare Differentialgleichung

$$\begin{pmatrix} \dot{x}(t) \\ \dot{v}(t) \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ -\omega_0^2 & -2\gamma \end{pmatrix}}_{=:A} \begin{pmatrix} x(t) \\ v(t) \end{pmatrix}$$

für den Fall $\gamma = \omega_0$ zu lösen, in welchem A nicht diagonalisierbar ist. Es hat dann

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

den einzigen Eigenwert $\lambda = -\gamma$ mit dem eindimensionalen Eigenraum E_λ aufgespannt von

$$w = \begin{pmatrix} 1 \\ -\gamma \end{pmatrix}.$$

Den fehlenden Hauptvektor erhält man aus $(A-\lambda)v = w$, also

$$\begin{pmatrix} \gamma & 1 \\ -\gamma^2 & -\gamma \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -\gamma \end{pmatrix} \quad \Rightarrow \quad v = \begin{pmatrix} 1 \\ 1-\gamma \end{pmatrix}.$$

Die Transformationsmatrix auf Jordansche Normalform ist also

$$Q = (w, v) = \begin{pmatrix} 1 & 1 \\ -\gamma & 1-\gamma \end{pmatrix}, \quad Q^{-1} = \frac{1}{\det Q} \begin{pmatrix} 1-\gamma & -1 \\ \gamma & 1 \end{pmatrix} = \begin{pmatrix} 1-\gamma & -1 \\ \gamma & 1 \end{pmatrix}.$$

Damit finden wir

$$\begin{aligned} \mathrm{e}^{At} &= Q\mathrm{e}^{Jt}Q^{-1} = Q \begin{pmatrix} \mathrm{e}^{-\gamma t} & t\mathrm{e}^{-\gamma t} \\ 0 & \mathrm{e}^{-\gamma t} \end{pmatrix} Q^{-1} \\ &= \begin{pmatrix} 1 & 1 \\ -\gamma & 1-\gamma \end{pmatrix} \begin{pmatrix} \mathrm{e}^{-\gamma t} & t\mathrm{e}^{-\gamma t} \\ 0 & \mathrm{e}^{-\gamma t} \end{pmatrix} \begin{pmatrix} 1-\gamma & -1 \\ \gamma & 1 \end{pmatrix} \\ &= \begin{pmatrix} \mathrm{e}^{-\gamma t} + \gamma t\,\mathrm{e}^{-\gamma t} & t\mathrm{e}^{-\gamma t} \\ \dots & \dots \end{pmatrix}. \end{aligned}$$

Also ist

$$x(t) = \left(\mathrm{e}^{-\gamma t} + \gamma t\,\mathrm{e}^{-\gamma t}\right) x(0) + t\mathrm{e}^{-\gamma t} v(0).$$

Zum Abschluss betrachten wir nochmals die Äquivalenzrelation aus Beispiel 8.2 etwas näher. Für einen $\mathbb{K}$-Vektorraum V und einen Untervektorraum $U \subset V$ hatten wir für $x, y \in V$ definiert, dass

$$x \sim_U y \quad \Leftrightarrow \quad x - y \in U.$$

8.25 Satz. **Der Quotientenraum**

Sein V ein $\mathbb{K}$-Vektorraum und $U \subset V$ ein Unterraum. Der Quotientenraum $V/U := V/\sim_U$ hat die Struktur eines $\mathbb{K}$-Vektorraums, falls man die Addition und die skalare Multiplikation vertreterweise definiert, also

$$[x] + [y] = [x + y] \qquad \text{und} \qquad \lambda[x] := [\lambda x].$$

Sei $W \subset V$ ein Komplement von U in V, also $V = U \oplus W$. Dann ist

$$\pi|_W : W \to V/U, \ w \mapsto [w],$$

ein Isomorphismus. Falls $\dim V = n < \infty$, gilt somit die Dimensionsformel

$$\dim(U) + \dim(V/\sim_U) = \dim(V).$$

Beweis. Es ist zu zeigen, dass die Addition und die skalare Multiplikation wohldefiniert sind, also unabhängig von der Wahl des Vertreters ist. Das ist aber leicht zu sehen. Die Vektorraumeigenschaft folgt dann direkt.

Für die zweite Aussage müssen wir nur zeigen, dass π linear und injektiv ist. Linearität ist klar und Injektivität folgt aus

$$\pi|_W v = 0 \quad \Leftrightarrow \quad v \in [0] \cap W \quad \Leftrightarrow \quad v \in U \cap W \quad \Leftrightarrow \quad v = 0.$$

Da $\mathrm{Kern}(\pi) = U$ und $\mathrm{Bild}(\pi) = \mathrm{Bild}(\pi|_W)$ ist, folgt die Dimensionsformel aus der Dimensionsformel für lineare Abbildungen, Satz 2.10. $\square$

Mit Hilfe des Quotientenraums kann man nun auch die Idee präzisieren, dass jede lineare Abbildung auf dem "Komplement" ihres Kerns ein Isomorphismus auf ihr Bild ist.

8.26 Satz. **Der Homomorphiesatz für lineare Abbildungen**

Sei $L : V \to W$ linear. Dann ist

$$\tilde{L} : V/\mathrm{Kern}(L) \to \mathrm{Bild}(L), \ [v] \mapsto \tilde{L}[v] := Lv$$

ein Isomorphismus und L lässt sich schreiben als $L = \tilde{L} \circ \pi$.

Beweis. $\tilde{L}$ ist wohldefiniert, denn für $w \in [v]$ gilt $v - w \in \mathrm{Kern}(L)$ und somit $Lw = L(w) + L(v - w) = L(w + v - w) = Lv$. $\tilde{L}$ ist linear, denn für $[v], [w] \in V/\mathrm{Kern}(L)$ ist

$$\tilde{L}([v] + [w]) = \tilde{L}([v + w]) = L(v + w) = Lv + Lw = \tilde{L}[v] + \tilde{L}[w]$$

und analog auch $\tilde{L}(\lambda[v]) = \lambda\tilde{L}([v])$. Schließlich ist $\tilde{L}$ per Definition surjektiv und wegen

$$\tilde{L}[v] = 0 \quad \Leftrightarrow \quad v \in \mathrm{Kern}(L) \quad \Leftrightarrow \quad v \in [0] \quad \Leftrightarrow \quad [v] = [0]$$

auch injektiv. □

8.27 Bemerkung. Ein entsprechender Satz gilt auch für Gruppenhomomorphsimen. Seien G und H Gruppen und $f : G \to H$ ein Gruppenhomomorphismus, also eine Abbildung die mit den Gruppenoperationen vertauscht. Man definiert nun $\mathrm{Kern}(f) := \{g \in G \,|\, f(g) = 1_H\}$, wobei 1_H die Identität in H bezeichne. Dann ist $\mathrm{Kern}(f) \subset G$ eine Unterguppe (sogar ein sog. Normalteiler) und $G/\mathrm{Kern}(f)$ ist ebenfalls eine Gruppe, die über

$$\tilde{f} : G/\mathrm{Kern}(f) \to \mathrm{Bild}(f)\,,\ [g] \mapsto \tilde{f}([g]) := f(g)$$

isomorph zum Bild von f ist.