

LINEARE ALGEBRA FÜR PHYSIK BSc.

Skript zur Vorlesung in WS09/10, SS10, SS12

Bálint Farkas

Besonderen Dank an die Personen, die bei der Umsetzung des Manuskriptes in L^AT_EX mitgeholfen haben:

*Murat Bayram, Michaela Hamm, Philipp Isserstedt
Andreas Kleiner, Martin Ziegler, Parvaneh Zourchang*

Einige Tippfehler wurden von Hörer der Vorlesung entdeckt, besten Dank dafür! Weitere Kommentare und Vorschläge sind willkommen und erwünscht. Diese senden Sie bitte an: farkas@mathematik.tu-darmstadt.de. Danke!

INHALTSVERZEICHNIS

1. Grundlegende Begriffe, Wiederholung	1	4. Rechnen mit Matrizen	42
Was sind Vektoren?	1	Elementare Umformungen	42
Gruppen, Körper, Vektorräume	3	Lineare Gleichungssysteme	48
Kartesische Koordinaten, Geraden, Ebenen .	6	Inhomogene lineare Gleichungssysteme	51
Norm, Skalarprodukt Winkel, Orientierung .	7	Die Determinante	53
Die Hesse-Normalform	11	Permutationen	60
Linearformen, multilineare Abbildungen . . .	13	Determinanten und Gleichungssysteme	63
Lineare Gleichungssysteme	17	5. Spektraltheorie	67
Das Gauß'sche Eliminationsverfahren	18	Eigenwerte und Eigenvektoren	67
2. Lineare Unabhängigkeit, Basen	20	Das charakteristische Polynom	70
Lineare Unabhängigkeit und Hülle, Basen ..	20	Diagonalisierbarkeit	72
Existenz von Basen	23	Die Jordan-Normalform	74
Orthonormalbasen	24	6. Normale Abbildungen	76
Orthogonale Projektion	28	Adjungierte Abbildung	77
Gram-Schmidt-Verfahren	29	Normale Matrizen und Abbildungen	79
3. Lineare Abbildungen und Matrizen	30	Unitäre Matrizen und Abbildungen	80
Der Vektorraum der linearen Abbildungen .	30	Selbstadjungierte Matrizen und Abbildungen	82
Lineare Abbildungen auf Basen	31	Positive Matrizen und Abbildungen	85
Die Matrix einer linearen Abbildung	31	Index	87
Operationen mit Matrizen	34		
Basiswechsel	38		

1. GRUNDLEGENDE BEGRIFFE, WIEDERHOLUNG

Die Gegenstände der linearen Algebra sind Vektoren, Mengen von Vektoren (Vektorräume) und lineare Abbildungen.

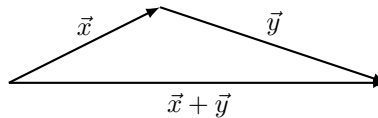
1.1. Was sind Vektoren?

Antwort Nr. 1: Vektoren sind Objekte \vec{x} , die eine Richtung im Raum und eine Länge besitzen: "Pfeile", "**Richtungsvektoren**", "freie Vektoren".

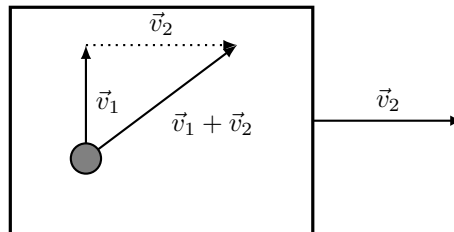
Beispiel. In der Physik: Kräften Geschwindigkeiten (Magnet)-felder, Wellenvektoren, Verschiebungen im Raum sind freie Vektoren.

Bemerkung. Richtungsvektoren kann man

- (i) mit einer Zahl (Skalar) **multiplizieren**, d.h., verlängern, verkürzen, umdrehen;
- (ii) addieren durch Zueinanderhängen:



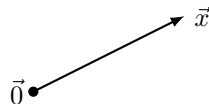
Beispiel. Bei Geschwindigkeiten sieht es folgenderweise aus:



Beispiel. Newton, Lex quartia: Das vierte Newtonsche Gesetz (**Superpositionsprinzip**), besagt, dass Kräfte wie Vektoren sich addieren.

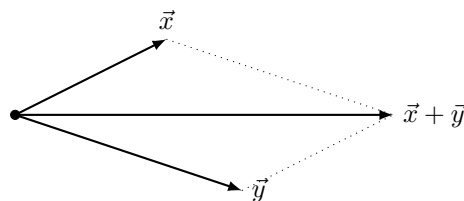
1.2. Was sind Vektoren?

Antwort Nr. 2: Vektoren sind Punkte im Raum relativ zu einem "Nullpunkt" $\vec{0}$. Diese Art von Vektoren nennt man "**Ortsvektoren**", sie werden folgenderweise symbolisiert:



Bemerkung. Richtungsvektoren kann man

- (i) mit einer Zahl (Skalar) **multiplizieren**, d.h., verlängern, verkürzen, umdrehen;
- (ii) mithilfe Parallelogrammen addieren:

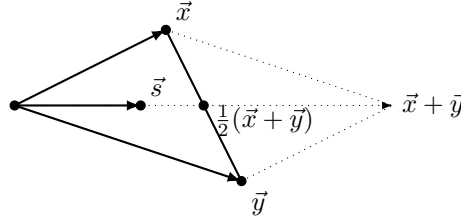


Beispiel. Wir platzieren in den Punkt \vec{x} eine Masse m_1 und in den Punkt \vec{y} eine Masse m_2 . Den Schwerpunkt \vec{s} dieser Konfiguration rechnet man so aus:

$$\vec{s} = \frac{m_1 \vec{x} + m_2 \vec{y}}{m_1 + m_2}.$$

Beispiel. Betrachte ein Dreieck mit Ecken $\vec{0}$, \vec{x} und \vec{y} . Auf jeden dieser Ecken stelle jeweils einen Masse von $1/3$. So ist der Schwerpunkt:

$$\vec{s} = \frac{1}{3}\vec{0} + \frac{1}{3}\vec{x} + \frac{1}{3}\vec{y} = \frac{1}{3}(\vec{x} + \vec{y}) = \frac{2}{3}\left(\frac{1}{2}\vec{x} + \frac{1}{2}\vec{y}\right).$$



Beispiel. Sei T eine Drehung des Raumes um $\vec{0}$. Dann gilt:

- (i) $T(\lambda\vec{x}) = \lambda T(\vec{x})$,
- (ii) $T(\vec{x} + \vec{y}) = T(\vec{x}) + T(\vec{y})$.

1.3. Was sind Vektoren?

Antwort Nr. 3: Vektoren sind Elemente in \mathbb{R}^n oder \mathbb{C}^n , also n -**Tupel** von Zahlen $\vec{x} = (x_1, x_2, \dots, x_n)$, $x_i \in \mathbb{R}$ (oder \mathbb{C}). Diese Art von Vektoren nennt man **Koordinatenvektoren**.

Bemerkung. Koordinatenvektoren kann man

- (i) mit einer Zahl λ (Skalar) **multiplizieren**:

$$\lambda(x_1, x_2, \dots, x_n) = (\lambda x_1, \lambda x_2, \dots, \lambda x_n);$$

- (ii) Addieren:

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

Beispiel. Koordinaten in der Ebene E_2 oder im Raum E_3 .

Beispiel. Zwei Substanzen haben spezifische Wärme c_1 und c_2 sowie spezifische Gewichte g_1 und g_2 . Wenn wir die Materialien im Verhältnis $\alpha_1, \alpha_2 \geq 0$, $\alpha_1 + \alpha_2 = 1$ mischen, so ergibt sich eine Substanz mit

$$\text{Spezifische Wärme:} \quad g = \alpha_1 g_1 + \alpha_2 g_2$$

$$\text{Spezifische Gewichte:} \quad c = \alpha_1 c_1 + \alpha_2 c_2.$$

Wie muss man für vorgegebene g und c mischen? Dies Frage führt zu linearen Gleichungssystemen. Bemerkte, dass die Abbildung

$$T: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \vec{\alpha} = (\alpha_1, \alpha_2) \mapsto T(\alpha_1, \alpha_2) = (g, c)$$

erfüllt

- (i) $T(\lambda\vec{\alpha}) = \lambda T(\alpha x)$,
- (ii) $T(\vec{\alpha} + \vec{\beta}) = T(\vec{\alpha}) + T(\vec{\beta})$.

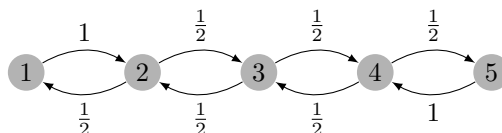
Unser Problem ist also ein $\vec{\alpha} \in \mathbb{R}^2$ mit $T(\vec{\alpha}) = (g, c)$ zu finden.

Beispiel. Stellen wir uns die folgende Situation vor. Ein Teilchen befindet sich in einem der fünf vorhandenen Kästchen, und in jeder Zeiteinheit springt es in ein benachbartes. Wir wissen aber nicht in was die genaue Regel für seine Bewegung sind, nur die Wahrscheinlichkeiten für die möglichen Sprünge sind bekannt. Die Kästchen nummerieren wir wie folgt durch:



Ist das Teilchen

- in 2, 3, 4, so springt es nach links bzw. nach rechts mit Wahrscheinlichkeit $\frac{1}{2}$;
- in 1, so springt es nach rechts (also nach 2) mit Wahrscheinlichkeit 1;
- in 5, so springt es nach links (also nach 4) mit Wahrscheinlichkeit 1;



Weiß man zu Beginn den Startpunkt, so weiß man nach n Zeiteinheiten nur Wahrscheinlichkeiten p_1, p_2, p_3, p_4, p_5 für den Aufenthalt in Kästchen 1, 2, 3, 4, 5. So ist $\vec{p} = (p_1, p_2, p_3, p_4, p_5)$ ein Vektor mit

$$p_i \geq 0, \quad \text{und} \quad \sum_{i=1}^5 p_i = 1,$$

ein so genannte **Wahrscheinlichkeitsvektor**. Dieser Vektor beschreibt die Wahrscheinlichkeitsverteilung des Teilchen. Ist $\vec{p} = (p_1, p_2, p_3, p_4, p_5)$ der Wahrscheinlichkeitsvektor zum Zeitpunkt n , so ist

$$T(\vec{p}) = \left(\frac{1}{2}p_2, p_1 + \frac{1}{2}p_3, \frac{1}{2}p_2 + \frac{1}{2}p_4, \frac{1}{2}p_3 + p_5, \frac{1}{2}p_4\right)$$

der Wahrscheinlichkeitsvektor zum Zeitpunkt $n+1$. Bemerke wieder, dass T die folgenden Eigenschaften hat:

- (i) $T(\lambda\vec{p}) = \lambda T(\vec{p})$,
- (ii) $T(\vec{p} + \vec{q}) = T(\vec{p}) + T(\vec{q})$.

Typische Fragestellungen sind:

1. Gegeben \vec{p} zum Zeitpunkt n was wahr der Wahrscheinlichkeitsvektor eine Sekunde früher? (Dies führt wieder zu einem Gleichungssystem.)
2. Gibt es einen invarianten Wahrscheinlichkeitsvektor $\vec{p} \in \mathbb{R}^5$, d.h. \vec{p} mit $T(\vec{p}) = \vec{p}$?

1.4. Was sind Vektoren?

Antwort Nr. 4: Vektoren sind Objekte, die man addieren und mit einem Skalar multiplizieren kann. Um dies präzise zu machen brauchen wir eine kleine Exkursion in Algebra. Sammeln wir zunächst welche gemeinsame Eigenschaften die Addition und Multiplikation von Zahlen haben. Addition $+$ und Multiplikation \cdot sind beide Operationen (Abbildungen) auf einer Menge, die zu zwei Elementen ein weiteres zuordnet. Sie erfüllen die Folgenden:

$$\left. \begin{aligned} (a+b)+c &= a+(b+c) = a+b+c \\ (a \cdot b) \cdot c &= a \cdot (b \cdot c) = a \cdot b \cdot c \end{aligned} \right\} \text{Assoziativität;}$$

d.h. egal wie wir Klammern einsetzen (die Elemente gruppieren) das Resultat bleibt dasselbe, deswegen ist die Schreibweise eigentlich ohne Klammern auch sinnvoll.

$$\left. \begin{aligned} 0+a &= a+0 = a \\ 1 \cdot a &= a \cdot 1 = a \end{aligned} \right\} \text{Einselement;}$$

d.h. es gibt ein Element, so dass die Operation mit diesem Element gar nichts tut: "seine Wirkung ist neutral"; Deswegen heißt so ein Element auch neutral, **neutrales Element**.

$$\left. \begin{aligned} a+(-a) &= (-a)+a = 0 \\ a \cdot a^{-1} &= a^{-1} \cdot a = 1 \end{aligned} \right\} \text{Inverselement;}$$

Statt $b+(-a)$ schreibt man auch $b-a$.

Definition. Sei G eine Menge und $m : G \times G \rightarrow G$ eine Abbildung (Operation) mit folgenden Eigenschaften:

$$(G1) \quad m(a, m(b, c)) = m(m(a, b), c) \text{ gilt für alle } a, b, c \in G.$$

(Assoziativität)

$$(G2) \quad \text{Es gibt ein Element } e \in G \text{ mit } m(a, e) = m(e, a) \text{ für alle } a \in G.$$

(Existenz des neutralen Elements)

(G3) Für alle $a \in G$ gibt es $b \in G$ mit $m(a, b) = m(b, a) = e$ (e ist das Element aus $(G2)$).
(Existenz des inverses Elements)

So heißt G (oder (G, m)) eine **Gruppe**. Gilt ferner $m(a, b) = m(b, a)$ für alle $a, b \in G$, so heißt die Gruppe **kommutativ** oder **Abelsch**.

Statt $m(a, b)$ ist manchmal bequemer $a \cdot b$, ab oder $a + b$ zu schreiben. Übrigens ist ein neutrales Element eindeutig, so redet man über *das* neutrale Element e der Gruppe G , und es wird manchmal auch als 1 oder als 0 bezeichnet.

Beispiel. Wie wir das oben besprochen haben, sind $(\mathbb{R}, +)$ und $(\mathbb{R} \setminus \{0\}, \cdot)$ beide kommutative Gruppen.

Auf \mathbb{R} sind die Addition und Multiplikation durch die Distributivgesetz verbindet:

Definition. Sei K eine Menge versehen mit zwei Operationen “+” und “·”, so dass $(K, +)$ eine kommutative Gruppe mit neutralem Element 0, und $(K \setminus \{0\}, \cdot)$ eine kommutative Gruppe (mit neutralem Element 1) ist, und so dass

$$(a + b) \cdot c = a \cdot c + b \cdot c \quad \text{für alle } a, b, c \in K \text{ gilt.}$$

(Distributivität)

So heißt $(K, +, \cdot)$ ein **Körper**.

Beispiel. Jede der folgenden Mengen ist ein Körper:

1. \mathbb{Q} mit + und ·, der Körper der **rationalen** Zahlen.
2. \mathbb{R} mit + und ·, der Körper der **reellen** Zahlen.
3. \mathbb{C} mit + und ·, der Körper der **komplexen** Zahlen.

Besonders wichtig sind (für uns) \mathbb{R} und \mathbb{C} .

A C H T U N G : ab jetzt bezeichnet \mathbb{K} immer einer der beiden Körper \mathbb{C} oder \mathbb{R} . Wir werden die Notation K für einen allgemeinen Körper behalten.

Definition. Ein **Vektorraum** V über einem Körper $(K, +, \cdot)$ ist eine kommutative Gruppe $(V, +)$ versehen mit einer weiteren (skalaren) Multiplikation $\cdot : K \times V \rightarrow V$, so dass die folgenden Bedingungen erfüllt sind:

(S1) Für alle $\lambda, \mu \in K$ und $\vec{x} \in V$ gilt $(\lambda + \mu) \cdot \vec{x} = \lambda \cdot \vec{x} + \mu \cdot \vec{x}$.

(S2) Für alle $\lambda \in K$ und $\vec{x}, \vec{y} \in V$ gilt $\lambda \cdot (\vec{x} + \vec{y}) = \lambda \cdot \vec{x} + \lambda \cdot \vec{y}$.

(S3) Für alle $\lambda, \mu \in K$ und $\vec{x} \in V$ gilt $(\lambda \cdot \mu) \cdot \vec{x} = \lambda \cdot (\mu \cdot \vec{x})$.

(S4) Für alle $\vec{x} \in V$ gilt $1 \cdot \vec{x} = \vec{x}$ (wobei 1 das neutrale Element in $(K \setminus \{0\}, \cdot)$ bezeichnet).

A C H T U N G : + und · haben zwei Bedeutungen, wie $\lambda + \mu$ und $\lambda \cdot \mu$ für Skalaren, und $\lambda \cdot \vec{x}$ und $\vec{x} + \vec{y}$ für Vektoren! Aus dem Kontext ist es aber immer eindeutig welche Bedeutung dann zu verstehen ist.

Die Elemente eines Vektorraums heißen **Vektoren**. Das neutrale Element der Gruppe $(V, +)$ heißt auch **0-Vektor**, notiert: $\vec{0}$.

Bemerkung. Es ist $0 \cdot \vec{v} = \vec{0}$, $\mu \cdot \vec{0} = \vec{0}$, $(-1)\vec{v} = -\vec{v}$ = das inverse Element zu \vec{v} . Die Gleichung

$$\vec{u} + \vec{x} = \vec{v}$$

ist eindeutig lösbar mit

$$\vec{x} = \vec{v} + (-\vec{u}) = \vec{v} - \vec{u}.$$

Definition. Seien V und W Vektorräume. Eine Abbildung

$$L : V \rightarrow W$$

heißt **lineare Abbildung**, oder **linear**, falls für $\vec{u}, \vec{v} \in V$, $\lambda \in K$ gilt

$$\text{(L1)} \quad L(\lambda\vec{v}) = \lambda L(\vec{v}),$$

$$\text{(L2)} \quad L(\vec{v} + \vec{u}) = L(\vec{v}) + L(\vec{u}).$$

Eine lineare Abbildung heißt auch **Vektorraum-Homomorphismus**. Ist L zusätzlich injektiv und surjektiv (also bijektiv), so heißt L ein **Vektorraum-Isomorphismus**. (Bijektiv bedeutet also: (1) **injektiv**: $L(\vec{x}) = L(\vec{y})$ genau dann, wenn $\vec{x} = \vec{y}$ und (2) **surjektiv**: für jedes $\vec{w} \in W$ existiert $\vec{v} \in V$ mit $\vec{w} = L(\vec{v})$.) Falls $W = \mathbb{K}$ so heißt L eine **Linearform**. Die Menge aller linearen Abbildungen von V nach W wird als $\mathcal{L}(V, W)$ bezeichnet.

Definition. Sei V ein Vektorraum über dem Körper K . Eine nicht leere Teilmenge $V_0 \subseteq V$ heißt **Untervektorraum** (oder einfach **Unterraum**, oder **linearer Teilraum**), falls gilt

$$\text{(TR1)} \quad \vec{u} \in V_0 \text{ und } \lambda \in K \text{ implizieren } \lambda\vec{u} \in V_0 \quad (\text{abgeschlossen bzgl. skalarer Multiplikation}),$$

$$\text{(TR2)} \quad \vec{u}, \vec{v} \in V_0 \text{ impliziert } \vec{u} + \vec{v} \in V_0 \quad (\text{abgeschlossen bzgl. Addition}).$$

Definition. Sei V ein Vektorraum über dem Körper K , und seien $\vec{v}_1, \dots, \vec{v}_n \in V$. Falls

$$\vec{v} = \lambda_1\vec{v}_1 + \dots + \lambda_n\vec{v}_n$$

für geeignete $\lambda_i \in K$, so heißt \vec{v} eine **Linearkombination** der Vektoren $\vec{v}_1, \dots, \vec{v}_n$. Die Zahlen $\lambda_1, \dots, \lambda_n$ heißen die **Koeffizienten** in der Linearkombination.

In dieser abstrakten Definition sind alle vorherigen Beispiele enthalten. Sie ermöglicht es aber auch, diejenige als Vektoren zu interpretieren, an die man vorher gar nicht gedacht hatte.

1.5. Was sind Vektoren?

Antwort Nr. 5: Auch Funktionen kann man addieren und mit Zahlen multiplizieren.

Beispiel. Betrachte die Mengen

$$\mathcal{P} := \{\text{Polynome } p : K \rightarrow K\}$$

$$\mathcal{P}_n := \{\text{Polynome } p : K \rightarrow K : \text{Grad}(p) \leq n\},$$

wobei \mathbb{K} den Körper \mathbb{R} (oder \mathbb{C}) bezeichnet. Mit $p_1 + p_2$ (Addition) und λp (skalarer Multiplikation) wie üblich sind \mathcal{P} , \mathcal{P}_n Vektorräume.

Ist p ein Polynom n -ten Grades, dann definiert

$$q(z) = p(z - z_0)$$

für $z \in \mathbb{K}$ fest auch ein Polynom n -ten Grades.

Beispiel. Die Translation S_{z_0} um z_0

$$S_{z_0} : \mathcal{P} \rightarrow \mathcal{P} \quad (\text{oder } \mathcal{P}_n \rightarrow \mathcal{P}_n)$$

definiert durch

$$S_{z_0}(p) = p(z - z_0)$$

ist also eine Abbildung.

Klar ist:

$$S_{z_0}(\lambda p) = \lambda S_{z_0}(p)$$

$$S_{z_0}(p_1 + p_2) = S_{z_0}(p_1) + S_{z_0}(p_2)$$

ist lineare Abbildung.

Aus der Schule wissen Sie:

Polynome kann man differenzieren:

$$p \mapsto p'$$

und es gilt

$$\begin{aligned} (\lambda p)' &= \lambda p' \\ (p_1 + p_2)' &= p_1' + p_2'. \end{aligned}$$

Ist $\text{Grad } p = k$, dann ist $\text{Grad } p' = k - 1$ (falls $p \neq \text{konstant}$). Also:

$$D : \mathcal{P} \rightarrow \mathcal{P} \quad (\text{oder } \mathcal{P}_n \rightarrow \mathcal{P}_n), \quad D : p \mapsto p'$$

ist eine lineare Abbildung. Genauso ist die k -te Ableitung

$$D^k : \mathcal{P}_n \rightarrow \mathcal{P}_n$$

linear und ebenso z.B. der **“Differentialoperator”**

$$T : D + 2D^2 + 4D^3 : p \mapsto p' + 2p'' + 4p'''.$$

Aufgabe. Sei p_0 ein Polynom. Finde Polynom p , so dass $p' + 2p'' + 4p''' = p_0$ (d.h. löse obige **Differentialgleichung**). Diese Aufgabe führt zu einem linearen Gleichungssystem $Tp = Tp_0$.

Die abstrakte Definition eines Vektorraumes erlaubt es, sehr verschiedene Objekte als Vektorraum zu erlauben und die Methoden und Anschauungen der linearen Algebra und Geometrie zu verwenden.

1.6. Kartesische Koordinaten (Vorläufiges). Zeichnen wir in der Ebene (Raum) einen Nullpunkt $\vec{0}$ aus, so wird jeder Punkt der Ebene ein Ortsvektor. Wir erhalten E_2, E_3 , wenn wir ein rechtwinkliges Koordinatensystem durch $\vec{0}$ legen. So lassen sich jedem Punkt \vec{x} zwei (drei) Koordinaten $(x_1, x_2) \in \mathbb{R}^2$ (oder $(x_1, x_2, x_3) \in \mathbb{R}^3$ zuordnen). Wir nennen x_1, x_2 (oder x_1, x_2, x_3) die kartesischen Koordinaten von x . Umgekehrt erhält $\mathbb{R}^2/\mathbb{R}^3$ die geometrische Interpretation einer Ebene/eines Raumes.

Analog kann es natürlich sein $(x_1, \dots, x_n) \in \mathbb{R}^n$ oder sogar aus \mathbb{C}^n als die Koordinaten eines Ortsvektors in einem abstrakten Raum aufzufassen und geometrisch zu denken. (Vgl. Raum-Zeit-Kontinuum in der Relativitätstheorie.)

Durch die Einführung von Koordinaten werden die geometrische Probleme zu analytischen/algebraischen Problemen.

Diese Idee stammt von Rene Descartes¹ (lat. Cartesius) (1596 La Haye/Touname - 1650 Stockholm).

1.7. Parameterdarstellung von Geraden. Sei V ein Vektorraum, stellen wir uns vor: $V = E_2$ oder E_3 . Für $\vec{x} \in V$ ist

$$\begin{aligned} G_0 &:= \{\lambda \vec{x} : \lambda \in K\} \\ &= \text{Menger der Punkte auf der Ursprungsgerade durch } \vec{x} \\ &= \text{Gerade durch } \vec{0} \text{ und } \vec{x} \end{aligned}$$

Wir behalten das Wort **“Gerade”** auch im allgemeinen bei.

Gerade durch zwei Punkte \vec{x} und \vec{y} : Offenbar gilt:

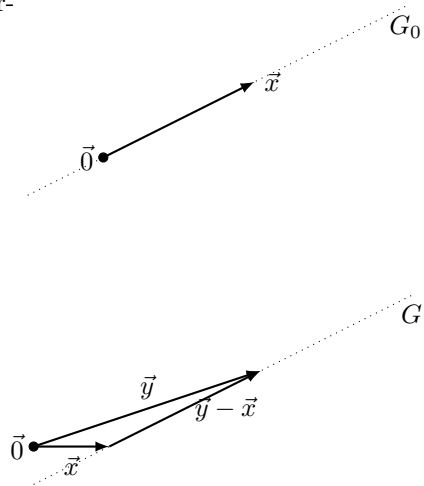
$$\vec{y} = \vec{x} + (\vec{y} - \vec{x}),$$

$$G := \{\vec{x} + \lambda(\vec{y} - \vec{x}) : \lambda \in K\} = \text{Gerade durch } \vec{0} \text{ und } \vec{y} - \vec{x} \text{ verschoben durch } \vec{x}.$$

Verbindungsstrecke zwischen \vec{x} und \vec{y} :

$$\{\vec{x} + \lambda(\vec{y} - \vec{x}) : 0 \leq \lambda \leq 1\} = \{(1 - \lambda)\vec{x} + \lambda\vec{y} : 0 \leq \lambda \leq 1\}.$$

Zum Beispiel Mittelpunkt: $\frac{1}{2}\vec{x} + \frac{1}{2}\vec{y}$



¹La Geometrie, 1637.

Bemerkung. In Koordinaten sieht das so aus:

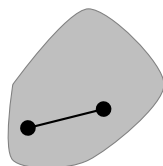
$$\vec{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad \vec{y} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} \quad G = \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \lambda \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} : \lambda \in \mathbb{R} \right\}.$$

Beispiel. Legen wir eine Masse m_1 in \vec{x} und eine Masse in \vec{y} . Dann ist der

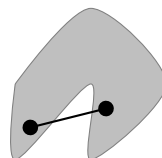
Schwerpunkt

$$\vec{s} = \frac{m_1 \vec{x} + m_2 \vec{y}}{m_1 + m_2} = \frac{m_1}{m_1 + m_2} \vec{x} + \frac{m_2}{m_1 + m_2} \vec{y}$$

Definition. Eine Menge $M \subseteq \mathbb{R}^2$ heißt **konvex**, falls für alle $x, y \in M$ die Verbindungsstrecke zwischen x und y in M liegt.



konvex



nicht konvex

1.8. Parameterdarstellung von Ebenen. Sei $V = E_3$ (oder "größer") und seien $\vec{x}_1, \vec{x}_2 \in V$. Dann ist

$$V_0 := \{ \lambda \vec{x}_1 + \mu \vec{x}_2 : \lambda, \mu \in K \}$$

= Menge der Punkte in der **Ebene** durch $\vec{0}$, \vec{x}_1 und \vec{x}_2

Ebene durch drei Punkte $\vec{x}_0, \vec{x}_1, \vec{x}_2$ kann man folgenderweise definieren:

$$V_1 := \{ \vec{x}_0 + \lambda(\vec{x}_1 - \vec{x}_0) + \mu(\vec{x}_2 - \vec{x}_0) : \lambda, \mu \in K \}$$

= Menge der Punkte in der **Ebene** durch $\vec{0}$, $\vec{x}_1 - \vec{x}_0$ und $\vec{x}_2 - \vec{x}_0$,
verschoben durch \vec{x}_0 .

(Auch für beliebiges V nennen wir E eine Ebene.)

1.9. Hyperebenen. Es geht gerade so weiter: Seien $\vec{x}_1, \dots, \vec{x}_n \in V$, dann heißt

$$H_0 = \{ \lambda_1 \vec{x}_1 + \lambda_2 \vec{x}_2 + \dots + \lambda_n \vec{x}_n : \lambda_i \in K, i = 1, 2, \dots, n \}$$

= **Hyperebene** durch $\vec{0}$ und $\vec{x}_1, \dots, \vec{x}_n$

$$H_1 = \{ \vec{x}_0 + \lambda_1(\vec{x}_1 - \vec{x}_0) + \dots + \lambda_n(\vec{x}_n - \vec{x}_0) : \lambda_i \in K, i = 1, 2, \dots, n \}$$

= **Hyperebene** durch die Punkte $\vec{x}_0, \vec{x}_1, \dots, \vec{x}_n$

Nebenbei ist $H_0 \subseteq V$ ein Untervektorraum, aber H_1 im Allgemeinen nicht. Man bezeichnet H_1 manchmal als **affiner Teilraum**.

1.10. Wie lang ist ein Vektor? (Norm).

Antwort: Das hängt ganz von seiner Bedeutung ab!

Beispiel. Betrachte ein Populationsmodell in n Städten.

Sei $x_i \geq 0, i = 1, 2, \dots, n$, die Größe der Population in Stadt "i" zu einem bestimmten Zeitpunkt. Ein sinnvolles Maß für die Länge/Größe des $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ könnte

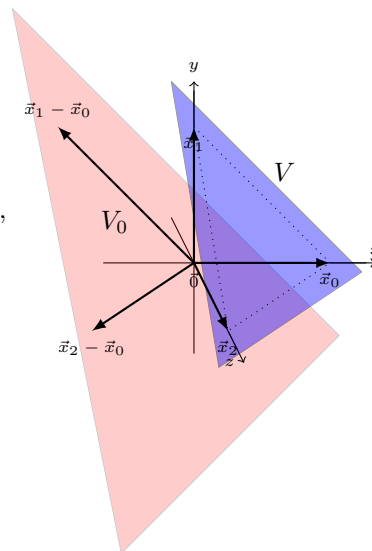
$$x_1 + x_2 + \dots + x_n$$

sein.

Beispiel. In einer Schraubenproduktion sollen Schrauben der Länge 5cm produziert werden.

Sei ℓ_i die wirkliche Länge der i -ten Schraube, $x_i = \ell_i - 5$ ihr Fehler. Ein sinnvolles Maß für die Länge/Größe des Fehlervektors $(x_1, x_2, \dots, x_m) \in \mathbb{R}^n$ könnte

$$\max\{|x_i| : 1 \leq i \leq n\}$$



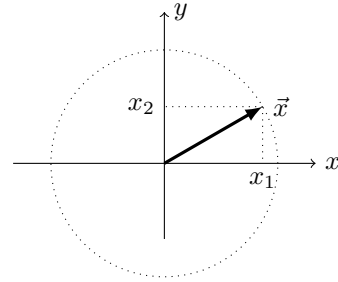
sein.

Beispiel. Sei $\vec{x} = (x_1, x_2) \in \mathbb{R}^2$ die Koordinaten eines Ortsvektors (in E_2).

Ein sinnvolles Maß für die Länge/Grösse von \vec{x} könnte

$$|(x_1, x_2)| = \sqrt{x_1^2 + x_2^2}$$

sein.



Definition. Sei V ein Vektorraum über \mathbb{K} ($= \mathbb{R}$ oder \mathbb{C}). Eine Abbildung

$$\|\cdot\| : V \rightarrow \mathbb{R}$$

heißt **Norm**, falls für jedes $\vec{x}, \vec{y} \in V$ und $\lambda \in \mathbb{K}$ gilt

$$\text{(N1)} \quad \|\vec{x}\| \geq 0, \quad \|\vec{x}\| = 0 \Leftrightarrow \vec{x} = \vec{0}$$

$$\text{(N2)} \quad \|\lambda \vec{x}\| = |\lambda| \cdot \|\vec{x}\|$$

$$\text{(N3)} \quad \|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$$

(Positivität)
(Homogenität)
(Dreiecksungleichung)

Bemerkung. Wir interpretieren $\|\vec{x}\|$ als die Länge des Vektors \vec{x} und $\|\vec{x} - \vec{y}\|$ als Abstand zwischen den Vektoren \vec{x} und \vec{y} .

Denn:

$$d : V \times V : (\vec{x}, \vec{y}) \mapsto d(\vec{x}, \vec{y}) := \|\vec{x} - \vec{y}\|$$

ist **Metrik** (siehe Analysis I) auf V .

Definition. Ist $\|\cdot\|$ Norm auf V , so heißt

$$\{\vec{x} \in V : \|\vec{x}\| \leq 1\}$$

Einheitskugel (zur Norm $\|\cdot\|$).

Bemerkung. 1. Einheitskugeln sind nicht immer rund (siehe unten). Ist $\|\vec{x}\| = 1$, so heißt \vec{x} ein **Einheitsvektor**. Ist $\vec{x} \neq \vec{0}$, so ist $\frac{\vec{x}}{\|\vec{x}\|}$ Einheitsvektor.

2. Die Einheitskugel ist konvex. (Dies folgt aus (N2) und (N3).)

Beispiel. Sei $V = \mathbb{R}^n$ oder \mathbb{C}^n .

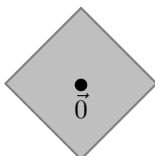
1. $\|(x_1, \dots, x_n)\|_1 := |x_1| + \dots + |x_n|$ ist Norm. (Übungsaufgabe)

2. $\|(x_1, \dots, x_n)\|_\infty := \max\{|x_i| : i = 1, 2, \dots, n\}$ ist Norm. (Übungsaufgabe)

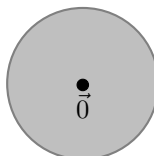
3. $\|(x_1, \dots, x_n)\|_2 := \sqrt{|x_1|^2 + \dots + |x_n|^2}$ ist Norm und heißt die **Euklidische Norm** auf \mathbb{R}^n bzw. auf \mathbb{C}^n ((N1),(N2) sind klar, (N3) nicht ganz).

Die Einheitskugel:

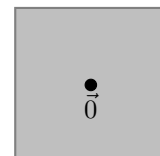
für die $\|\cdot\|_1$ -Norm



für die $\|\cdot\|_2$ -Norm



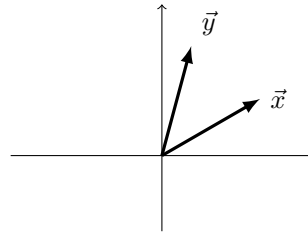
für die $\|\cdot\|_\infty$ -Norm



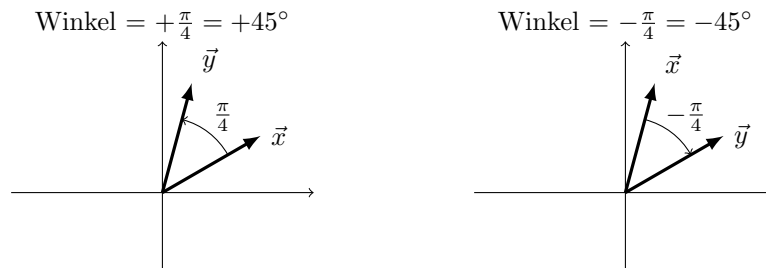
Definition. Sobald man eine Norm hat, kann man sagen was **Konvergenz** ist: Eine Folge $\vec{x}_n \in V$ heißt **konvergent** gegen \vec{x} , in Notation $\lim_n \vec{x}_n = \vec{x}$, falls für alle $\varepsilon > 0$ existiert ein Grenzindex $n_0 \in \mathbb{N}$, so dass für alle $n \geq n_0$ gilt $\|\vec{x}_n - \vec{x}\| \leq \varepsilon$ (siehe Analysis I).

1.11. Winkel. Sei V ein Vektorraum über \mathbb{R} und seien $\vec{x}, \vec{y} \in V$. Die Vektoren $\vec{0}, \vec{x}, \vec{y}$ spannen eine Ebene auf, die wir mit der Euklidischen Ebene E_2 identifizieren wollen. Also o.b.d.A können wir uns

das folgende Bild vorstellen:



Wir definieren den (orientierten) **Winkel** $\angle(\vec{x}, \vec{y})$ zwischen \vec{x} und \vec{y} als den kleineren Winkel von \vec{x} nach \vec{y} mit **positiver Orientierung**, also entgegen dem Uhrzeigersinn. Zum Beispiel



Bemerkung. Es gilt natürlich $|\angle(\vec{x}, \vec{y})| \leq \pi$.

Satz. Sei $\vec{0} \neq \vec{x} = (x_1, \dots, x_n)$ und $\vec{0} \neq \vec{y} = (y_1, \dots, y_n)$, dann gilt

$$\cos(\angle(\vec{x}, \vec{y})) = \frac{x_1 y_1 + x_2 y_2 + \dots + x_n y_n}{\|\vec{x}\|_2 \|\vec{y}\|_2}.$$

Beweis. Siehe Jänich Seite 39. ■

1.12. Das Standard-Skalarprodukt. Sei $V = \mathbb{R}^n$ oder \mathbb{C}^n . Für $\vec{x} = (x_1, \dots, x_n)$ und $\vec{y} = (y_1, \dots, y_n)$ setze

$$\langle \vec{x}, \vec{y} \rangle = x_1 \overline{y_1} + x_2 \overline{y_2} + \dots + x_n \overline{y_n}$$

das **Standard-Skalarprodukt** von \vec{x} und \vec{y} . Bemerke, dass die komplexe Konjugation nur im Falle $\mathbb{K} = \mathbb{C}$ nötig ist. Wir setzen weiterhin

$$\|\vec{x}\|_2^2 := \langle \vec{x}, \vec{x} \rangle.$$

Das Standard-Skalarprodukt erfüllt offenbar für $\vec{x}, \vec{y}, \vec{z} \in V$ und $\lambda, \mu \in \mathbb{R}$ (oder \mathbb{C}) die Folgenden:

(SP1) $\langle \lambda \vec{x} + \mu \vec{y}, \vec{z} \rangle = \lambda \langle \vec{x}, \vec{z} \rangle + \mu \langle \vec{y}, \vec{z} \rangle;$

(SP2) $\langle \vec{x}, \vec{y} \rangle = \overline{\langle \vec{y}, \vec{x} \rangle};$

(SP3) $\langle \vec{x}, \vec{x} \rangle \geq 0$, und $\langle \vec{x}, \vec{x} \rangle = 0$ genau dann, wenn $\vec{x} = \vec{0}$.

Das Standard-Skalarprodukt erfüllt also

$$\langle \vec{x}, \vec{y} \rangle = \|\vec{x}\|_2 \cdot \|\vec{y}\|_2 \cdot \cos \angle(\vec{x}, \vec{y})$$

und somit die **Cauchy-Schwarz-Ungleichung**:

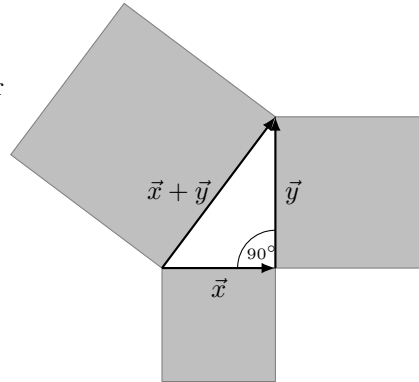
$$|\langle \vec{x}, \vec{y} \rangle| \leq \|\vec{x}\|_2 \cdot \|\vec{y}\|_2.$$

Definition. Die Vektoren \vec{x} und \vec{y} heißen **orthogonal** (oder **senkrecht**), falls $\langle \vec{x}, \vec{y} \rangle = 0$ gilt (wir schreiben $\vec{x} \perp \vec{y}$).

Sind \vec{x} und \vec{y} orthogonal, so gilt:

$$\begin{aligned} \|\vec{x} + \vec{y}\|_2^2 &= \langle \vec{x} + \vec{y}, \vec{x} + \vec{y} \rangle = \langle \vec{x}, \vec{x} \rangle + \langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{x} \rangle + \langle \vec{y}, \vec{y} \rangle \\ &= \langle \vec{x}, \vec{x} \rangle + \langle \vec{y}, \vec{y} \rangle + 0 + 0 = \|\vec{x}\|_2^2 + \|\vec{y}\|_2^2. \end{aligned}$$

Dies ist der **Satz von Pythagoras**.



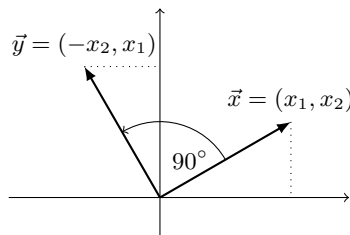
1.13. Zur geometrischen Interpretation des Skalarprodukts.

1. $\langle \vec{x}, \vec{y} \rangle = 0$ heißt also $\vec{x} \perp \vec{y}$. Seien $\vec{x} := (x_1, x_2) \in \mathbb{R}^2$ $\vec{y} := (-x_2, x_1) \in \mathbb{R}^2$. Dann gelten

$$\vec{x} \perp \vec{y}, \quad \langle \vec{x}, \vec{y} \rangle = x_1 x_2 - x_2 x_1 = 0, \quad \angle(\vec{x}, \vec{y}) = 90^\circ \quad \text{und} \quad \|\vec{y}\|_2 = \|\vec{x}\|_2.$$

Wir sehen also

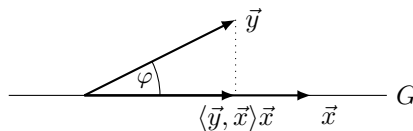
$\frac{\vec{y}}{\|\vec{y}\|_2}$ ist Einheitsvektor in dieselbe Richtung wie \vec{y} , **“Einheitsnormalenvektor”**.



2. Wir wissen bereits $\langle \vec{x}, \vec{y} \rangle = \|\vec{x}\|_2 \|\vec{y}\|_2 \cos \varphi$. Ist \vec{x} Einheitsvektor, d.h. $\|\vec{x}\|_2 = 1$, so ist

$$\langle \vec{y}, \vec{x} \rangle = \|\vec{y}\|_2 \cos \varphi$$

und somit $\langle \vec{y}, \vec{x} \rangle \vec{x}$ ist **orthogonale Projektion** von \vec{y} auf die Gerade $G = \{\lambda \vec{x} : \lambda \in \mathbb{K}\}$.



Sei nun \vec{x} beliebig aber nicht $\vec{0}$. Dann gilt:

$$\frac{\langle \vec{y}, \vec{x} \rangle}{\|\vec{x}\|_2} = \|\vec{y}\|_2 \cos \varphi.$$

Daraus folgt

$$\frac{\langle \vec{y}, \vec{x} \rangle}{\|\vec{x}\|_2} \frac{\vec{x}}{\|\vec{x}\|_2} = \frac{\langle \vec{y}, \vec{x} \rangle}{\|\vec{x}\|_2^2} \vec{x} =: \vec{y}_1.$$

D.h. \vec{y}_1 = **orthogonale Projektion** von \vec{y} auf die Gerade $G = \{\lambda \vec{x} : \lambda \in \mathbb{K}\}$.

Setze nun

$$\vec{y}_2 := \vec{y} - \vec{y}_1 = \vec{y} - \frac{\langle \vec{y}, \vec{x} \rangle}{\|\vec{x}\|_2^2} \vec{x}.$$

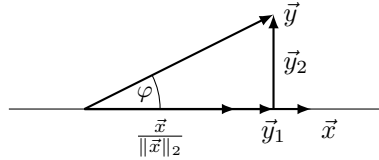
Dann gilt

$$\vec{y} = \vec{y}_1 + \vec{y}_2 \quad \text{mit} \quad \vec{y}_1, \vec{x} \perp \vec{y}_2$$

und \vec{y}_1 und \vec{x} **parallel** ($\vec{y}_1 \parallel \vec{x}$), d.h. \vec{y}_1 und \vec{x} liegen beide auf einer Ursprungsgeraden (sind also skalarer Vielfacher von einander). Denn

$$\langle \vec{y}_2, \vec{x} \rangle = \left\langle \vec{y} - \frac{\langle \vec{y}, \vec{x} \rangle}{\|\vec{x}\|_2^2} \vec{x}, \vec{x} \right\rangle = \langle \vec{y}, \vec{x} \rangle - \frac{\langle \vec{y}, \vec{x} \rangle}{\|\vec{x}\|_2^2} \langle \vec{x}, \vec{x} \rangle = 0.$$

Dies ergibt die **orthogonale Zerlegung** von \vec{y} längst \vec{x} .



1.14. Nochmal Geraden und Ebenen. Sei $\vec{0} \neq \vec{a} = (a_1, a_2) \in \mathbb{R}^2$ gegeben. Setze

$$G_0 := \{\lambda \vec{a} : \lambda \in \mathbb{R}\} = \text{Gerade durch Ursprung und } \vec{a}.$$

Behauptung. Setze $\vec{b} = (-a_2, a_1)$. Die Gerade G ist alternativ gegeben durch:

$$G_0 = \{\vec{x} : \langle \vec{x}, \vec{b} \rangle = 0\}.$$

Beweis. "⊆": Ist $\vec{x} = \lambda \vec{a}$, so gilt

$$\langle \vec{x}, \vec{b} \rangle = \langle \lambda \vec{a}, \vec{b} \rangle = \lambda \langle \vec{a}, \vec{b} \rangle = \lambda \cdot 0 = 0,$$

also G ist in der Menge auf der rechten Seite enthalten.

"⊇": Sei $\vec{x} \perp \vec{b}$ und betrachte die orthogonale Zerlegung von \vec{x} längst \vec{a} :

$$\vec{x} = \vec{x}_1 + \vec{x}_2 = \lambda_1 \vec{a} + \lambda_2 \vec{b}.$$

Es gilt nun:

$$0 = \langle \vec{x}, \vec{b} \rangle = \lambda_2 \|\vec{b}\|_2^2, \quad \text{und somit} \quad \lambda_2 = 0, \quad \text{d.h.} \quad \vec{x} = \lambda_1 \vec{a}.$$

Also $\vec{x} \in G$. ■

Sei jetzt $\alpha \in \mathbb{R}$ und setze

$$\vec{x}_0 := \alpha \frac{\vec{b}}{\|\vec{b}\|_2^2}.$$

Dann gelten die Äquivalenzen:

$$\langle \vec{x}, \vec{b} \rangle = \alpha \iff \langle \vec{x}, \vec{b} \rangle = \langle \vec{x}_0, \vec{b} \rangle \iff \langle \vec{x} - \vec{x}_0, \vec{b} \rangle = 0.$$

Das bedeutet:

$$\begin{aligned} G &:= \{\vec{x} \in \mathbb{R}^2 : \langle \vec{x} - \vec{x}_0, \vec{b} \rangle = 0\} = \{\vec{x} : \langle \vec{x}, \vec{b} \rangle = \alpha\} \\ &= \text{Gerade mit Ursprungsvektor } \vec{a} \text{ verschoben durch } \vec{x}_0. \end{aligned}$$

1.15. Hesse'sche Normalform (einer Geraden). Sei $\vec{b} \in \mathbb{R}^2$ mit $\vec{b} \neq \vec{0}$. Setze

$$\vec{n} := \frac{\vec{b}}{\|\vec{b}\|_2} = \text{normaler Einheitsvektor in Richtung } \vec{b}.$$

Setze ferner $\alpha := \langle \vec{x}, \vec{b} \rangle$. So kann man die obige Gerade auch folgenderweise darstellen:

$$G = \left\{ \vec{x} : \langle \vec{x}, \vec{n} \rangle = \frac{\alpha}{\|\vec{b}\|_2} \right\}.$$

Diese heißt die **Hesse-Normalform** der Geraden G . Dabei heißt \vec{n} ein **Normalvektor** der Geraden.

Analog in \mathbb{R}^3 (und \mathbb{R}^n). Sei $\vec{n} \in \mathbb{R}^n$ mit $\|\vec{n}\|_2 = 1$ und sei $\alpha \in \mathbb{R}$. Wir haben:

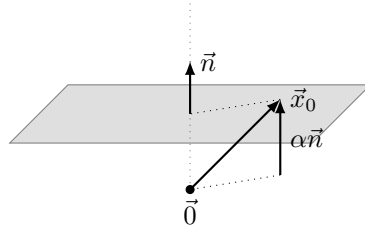
$$E_0 := \{x \in \mathbb{R}^n : \langle \vec{x}, \vec{n} \rangle = 0\} = \text{(Hyper)ebene "durch Ursprung" senkrecht zu } \vec{n}$$

$$\begin{aligned} E &:= \{\vec{x} \in \mathbb{R}^n : \langle \vec{x}, \vec{n} \rangle = \alpha\} = \{\text{Vektoren } \vec{x} \in \mathbb{R}^n \text{ mit orthogonaler Projektion auf } \vec{n} \text{ gleich } \langle \vec{x}, \vec{n} \rangle\} \\ &= \text{(Hyper)ebene senkrecht zu } \vec{n} \text{ durch } \alpha \vec{n}. \end{aligned}$$

Dies bedeutet: eine Hyperebene in \mathbb{R}^n können wir mit der Gleichung

$$\langle \vec{x}, \vec{n} \rangle = \alpha,$$

beschreiben (die **Hesse-Normalform**), wobei die Hyperebene senkrecht zu \vec{n} steht (\vec{n} ist ein **Normalvektor** der Hyperebene), und der Abstand zwischen der Hyperebene und dem Ursprung ist α .



1.16. Kanonische Basis. In $V = \mathbb{R}^n$ (oder \mathbb{C}^n) seien

$$\vec{e}_1 := (1, 0, 0, \dots, 0), \quad \vec{e}_2 := (0, 1, 0, \dots, 0), \quad \dots, \quad \vec{e}_n := (0, 0, 0, \dots, 1),$$

d.h. $\vec{e}_i := (0, 0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{K}^n$.
 \uparrow
i-te Koord.

Noch anders gesagt:

$$e_{ij} := \begin{cases} 1 & i = j \\ 0 & \text{sonst} \end{cases} = \mathbf{Kronecker-Delta} = \text{die Koordinaten von } \vec{e}_j.$$

Behauptung. Sei $V = \mathbb{R}^n$ oder \mathbb{C}^n . Ist $\vec{x} \in V$, $\vec{x} = (x_1, \dots, x_n)$, dann ist

$$\langle \vec{x}, \vec{e}_i \rangle = x_1 \overline{e_{i1}} + x_2 \overline{e_{i2}} + \dots + x_n \overline{e_{in}} = x_i.$$

Das bedeutet:

$$\vec{x} = \sum_{i=1}^n \overline{x_i} \vec{e}_i = \sum_{i=1}^n \langle \vec{x}, \vec{e}_i \rangle \vec{e}_i$$

Diese Darstellung ist sogar wegen koordinatenweiser Addition eindeutig: Falls

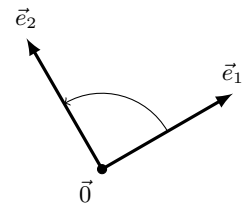
$$\vec{x} = \sum_{i=1}^n \alpha_i \vec{e}_i$$

Dann gilt $\alpha_i = x_i$.

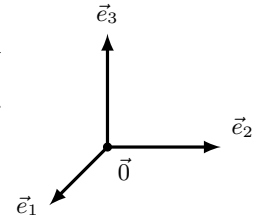
Zusammenfassend: \vec{x} ist also kanonischerweise Linearkombination der Vektoren $\vec{e}_i : i = 1, \dots, n$. Die Vektoren heißen \vec{e}_i **kanonische Basisvektoren** und $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ die **kanonische Basis** in \mathbb{K}^n . Die Zahlen x_1, \dots, x_n heißen die Koordinaten bzgl. der kanonischen Basis.

1.17. Orientierung.

1. Sei $\{\vec{e}_1, \vec{e}_2\}$ kanonische Basis von \mathbb{R}^2 . Bzgl. eines Nullpunktes $\vec{0}$ sei \vec{e}_1 ein Ortsvektor in E_2 . Schaut man von $\vec{0}$ nach \vec{e}_1 , dann sei \vec{e}_2 links von uns. Wir sagen das Koordinatensystem sei **positiv orientiert**. Die Drehung von \vec{e}_1 nach \vec{e}_2 nach links definiert einen Drehsinn/Orientierung zur Winkelmessung (dies haben wir schon festgelegt).



1. Sei $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ kanonische Basis von \mathbb{R}^3 . Bzgl. $\vec{0}$ sei \vec{e}_1 ein Ortsvektor in E_3 und \vec{e}_2 ein zweiter Ortsvektor senkrecht zu \vec{e}_1 : In der Ebene aufgespannt von \vec{e}_1 und \vec{e}_2 seien \vec{e}_1, \vec{e}_2 positiv Orientiert (siehe 1.). Nun sei \vec{e}_3 senkrecht zu dieser Ebene und es zeige nach oben. (“ \vec{e}_2 nach links” “ \vec{e}_3 nach oben”). Dies definiert die **positive Orientierung des Raums** (“**Rechte-Hand-Regel**”).



1.18. Vektoren und Linearformen. Sei $V = \mathbb{K}^n$ und sei $\vec{v} \in V$ beliebig (\mathbb{K} bezeichnet wie immer \mathbb{R} oder \mathbb{C}). Setze

$$L_v : V \rightarrow \mathbb{K} : \vec{x} \mapsto \langle \vec{x}, \vec{v} \rangle,$$

so ist L eine **Linearform**. Wir sagen, dass v die Linearform L_v **induziert**. Umgekehrt sei

$$L : V \rightarrow \mathbb{K}$$

eine Linearform, und sei $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ die kanonische Basis in \mathbb{K}^n . Setze

$$v_i := \overline{L(\vec{e}_i)} \in \mathbb{K} \quad \text{und} \quad \vec{v} := (v_1, v_2, \dots, v_n) \in \mathbb{K}^n.$$

Es gilt

$$L(\vec{x}) = L\left(\sum_{i=1}^n x_i \vec{e}_i\right) = \sum_{i=1}^n x_i L(\vec{e}_i) = \sum_{i=1}^n x_i \overline{v_i} = \langle \vec{x}, \vec{v} \rangle.$$

Wir haben also bewiesen:

Behauptung. Sei $V = \mathbb{K}^n$. Dann gibt es eine 1 zu 1 Beziehung zwischen den Vektoren in V und der Linearformen auf V .

(Dies ist ein Spezialfall von Satz von Riesz-Fréchet.)

Bemerkung. Der Raum

$$V^* := \mathcal{L}(V, \mathbb{K}) := \text{Linearformen auf } V$$

ist auch selber ein Vektorraum, und heißt der **Dualraum** von V . Die Vektorraum-Operationen sind folgenderweise definiert: Für $L_1, L_2 \in V^*$ und $\lambda \in \mathbb{K}$ sei

$$(L_1 + \lambda L_2)(\vec{x}) = L_1(\vec{x}) + \lambda L_2(\vec{x}).$$

Nach obiger Behauptung identifiziert man manchmal V und V^* .

Bemerkung. Zur Unterscheidung nennt man oft

- Vektoren in V “**Kontravariante Vektoren**” oder “**Kontravariante Tensoren 1-Stufe**”.
- Vektoren in V^* “**Kovariante Vektoren**” oder “**Kovariante Tensoren 1-Stufe**”.

Ab jetzt schreiben wir Vektoren in V als Spaltenvektoren, und Vektoren in V^* als Zeilenvektoren:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in V, \quad (v_1, v_2, \dots, v_n) \in V^*.$$

1.19. Multilineare Abbildungen. Seien V_1, \dots, V_k, W Vektorräume über dem Körper \mathbb{K} .

Definition. Eine Abbildung

$$M : V_1 \times \dots \times V_k \rightarrow W$$

heißt **multilinear** (genauer: **k-linear**), falls

$$\begin{aligned} M(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_{i-1}, \alpha \vec{x}_i + \vec{y}_i, \vec{x}_{i+1}, \dots, \vec{x}_n) = \\ \alpha M(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_{i-1}, \vec{x}_i, \vec{x}_{i+1}, \dots, \vec{x}_n) + M(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_{i-1}, \vec{y}_i, \vec{x}_{i+1}, \dots, \vec{x}_n) \end{aligned}$$

gilt für alle $\vec{x}_j, \vec{y}_j \in V, j = 1, \dots, n$ und für alle $\alpha \in \mathbb{K}$. Für $k = 2$ heißt M **bilinear**. M heißt dabei **symmetrisch**, falls

$$M(\vec{x}_1, \dots, \vec{x}_i, \dots, \vec{x}_j, \dots, \vec{x}_n) = M(\vec{x}_1, \dots, \vec{x}_j, \dots, \vec{x}_i, \dots, \vec{x}_n)$$

M heißt **antisymmetrisch** (oder: **alternierend**), falls

$$M(\vec{x}_1, \dots, \vec{x}_i, \dots, \vec{x}_j, \dots, \vec{x}_n) = (-1)M(\vec{x}_1, \dots, \vec{x}_j, \dots, \vec{x}_i, \dots, \vec{x}_n).$$

Für $V_1 = V_2 = \dots = V_k$ heißt M ein **Produkt**. Ist $W = \mathbb{K}$, so heißt M eine **k -Form**, **Tensor k -ter Stufe**.

Beispiel. $M : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} \quad (x, y) \mapsto \langle x, y \rangle$ ist eine Bilinearform, also eine 2-Form.

1.20. Parallelogramme- und Kreuzprodukt in der Ebene.

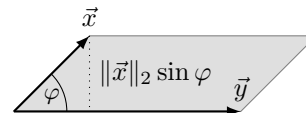
Vorbemerkung: es gibt auch negative Flächeninhalte/Längen wie z.B. beim Integral

$$\int_a^b f(x)dx = - \int_b^a f(x)dx.$$

Um negative Flächeninhalte interpretieren zu können, braucht man die Orientierung des Raums.

Sei also E_2 die euklidische Ebene mit orientiertem Koordinatensystem. $\vec{x}, \vec{y} \in E_2$ haben Koordinaten $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ bzw. $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$. Sei $\varphi = \angle(\vec{x}, \vec{y}) =$ Orientierter Winkel zwischen \vec{x} und \vec{y} . Wir erinnern uns an die Definition: $\vec{y}^\perp := \begin{pmatrix} -y_2 \\ y_1 \end{pmatrix}$. Dann gilt:

$$\begin{aligned} x_1 y_2 - x_2 y_1 &= \langle \vec{x}, -\vec{y}^\perp \rangle = \|\vec{x}\|_2 \cdot \|\vec{y}^\perp\|_2 \cos \angle(\vec{x}, \vec{y}^\perp) \\ &= \|\vec{x}\|_2 \cdot \|\vec{y}\|_2 \cos(90^\circ - \varphi) \\ &= \|\vec{x}\|_2 \cdot \|\vec{y}\|_2 \sin \varphi \\ &= \text{Oberfläche des durch } \vec{x} \text{ und } \vec{y} \text{ aufgespannten Parallelogramms.} \end{aligned}$$



Setze

$$R : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad R(\vec{y}) := -\vec{y}^\perp.$$

So ist R eine lineare Abbildung (eigentlich eine Drehung). Die Abbildung

$$b : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}, \quad b(\vec{x}, \vec{y}) := \langle \vec{x}, R(\vec{y}) \rangle$$

ist bilinear, und heißt das **Kreuzprodukt** von \vec{x} und \vec{y} . Mit der obigen Notation gilt

$$b(\vec{x}, \vec{y}) = x_1 y_2 - x_2 y_1,$$

und daraus folgt auch, dass b antisymmetrisch ist. Eine bequeme Merkregel (später bekommt diese auch ihre tatsächliche Bedeutung)

$$b(\vec{x}, \vec{y}) = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = x_1 y_2 - x_2 y_1.$$

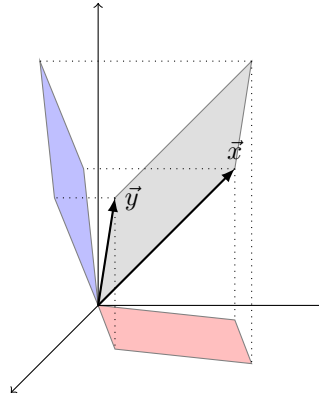
1.21. Das Vektorprodukt/Kreuzprodukt im \mathbb{R}^3 . Für

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3, \quad \text{und} \quad \vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \in \mathbb{R}^3$$

definieren wir das **Vektorprodukt**

$$\vec{x} \times \vec{y} := \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix} \in \mathbb{R}^3.$$

Die Koordinaten sind also orientierte Oberflächeninhalte der an die Koordinatenebenen projizierten, von \vec{x} und \vec{y} aufgespannten Parallelograms.



Eine formale Merkregel:

$$\vec{x} \times \vec{y} = \begin{vmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = \vec{e}_1 \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} - \vec{e}_2 \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} + \vec{e}_3 \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}.$$

Rechenregeln. Für $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^3$ und $\lambda \in \mathbb{R}$ gelten:

1. $(\vec{x} + \vec{y}) \times \vec{z} = \vec{x} \times \vec{z} + \vec{y} \times \vec{z}.$
 $(\lambda \vec{x}) \times \vec{y} = \lambda(\vec{x} \times \vec{y}) = \vec{x} \times (\lambda \vec{y}).$

Die Abbildung

$$B : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad B(\vec{x}, \vec{y}) := \vec{x} \times \vec{y}$$

ist also bilinear.

2. Die Abbildung B ist alternierend

$$\vec{x} \times \vec{y} = -\vec{y} \times \vec{x}.$$

3. (Zyklische Vertauschung)

$$\langle \vec{x} \times \vec{y}, \vec{z} \rangle = \langle \vec{y} \times \vec{z}, \vec{x} \rangle = \langle \vec{z} \times \vec{x}, \vec{y} \rangle.$$

Insbesondere

$$\langle \vec{x} \times \vec{y}, \vec{y} \rangle = \langle \vec{y} \times \vec{y}, \vec{x} \rangle = 0,$$

d.h. $\vec{x} \times \vec{y} \perp \vec{x}, \vec{y}.$

4. $\langle \vec{x} \times \vec{y}, \vec{u} \times \vec{v} \rangle = \langle \vec{x}, \vec{u} \rangle \langle \vec{y}, \vec{v} \rangle - \langle \vec{x}, \vec{v} \rangle \langle \vec{y}, \vec{u} \rangle.$
5. $(\vec{x} \times \vec{y}) \times \vec{z} = \langle \vec{x}, \vec{z} \rangle \vec{y} - \langle \vec{y}, \vec{z} \rangle \vec{x}.$

Beweis. 1) Folgt aus der Bilinearität des Kreuzprodukts b auf \mathbb{R}^2 . Betrachte nämlich die lineare Abbildungen (die Projektionen auf die Koordinatenebenen)

$$P_1 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} := \begin{pmatrix} x_2 \\ x_3 \end{pmatrix}, \quad P_2 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} := \begin{pmatrix} x_1 \\ x_3 \end{pmatrix}, \quad P_3 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} := \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Somit gilt

$$B(\vec{x}, \vec{y}) = \vec{x} \times \vec{y} = \vec{e}_1 b(P_1(\vec{x}), P_1(\vec{y})) - \vec{e}_2 b(P_2(\vec{x}), P_2(\vec{y})) + \vec{e}_3 b(P_3(\vec{x}), P_3(\vec{y})).$$

2)–5) Man zeigt diese Eigenschaften zunächst für $\vec{x}, \vec{y}, \vec{z}$ jeweils einer der Vektoren $\vec{e}_1, \vec{e}_2, \vec{e}_3$. Dann verwendet man 1), also die Bilinearität von B , um diese Eigenschaften für beliebigen $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^3$ zu zeigen. ■

Beispiel. Es gilt

$$\vec{e}_1 \times \vec{e}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \vec{e}_3$$

Also

$$\langle \vec{e}_1 \times \vec{e}_2, \vec{e}_2 \rangle = 0, \quad \text{und} \quad \langle \vec{e}_1 \times \vec{e}_2, \vec{e}_3 \rangle = \langle \vec{e}_3, \vec{e}_3 \rangle = 1.$$

Geometrische Interpretation.

1. $\vec{x} \times \vec{y} = \vec{0}$ gilt genau dann, wenn \vec{x} und \vec{y} parallel sind (also $\vec{0}$, \vec{x} , \vec{y} liegen auf einer Geraden). Denn: Ist z.B. $\vec{y} = \lambda \vec{x}$, so gilt $\vec{x} \times \vec{y} = \lambda(\vec{x} \times \vec{x}) = 0$. Umgekehrt: ist $\vec{x} \times \vec{y} = \vec{0}$, dann gilt

$$\vec{0} = (\vec{x} \times \vec{y}) \times \vec{x} = \langle \vec{x}, \vec{x} \rangle \vec{y} - \langle \vec{y}, \vec{x} \rangle \vec{x} = \langle \vec{y}, \vec{x} \rangle \vec{x}, \quad \text{d.h.} \quad \vec{y} = \frac{\langle \vec{y}, \vec{x} \rangle \vec{x}}{\langle \vec{x}, \vec{x} \rangle},$$

also $\vec{0}$, \vec{x} und \vec{y} liegen auf einer Geraden, sind also parallel.

2. Es gilt $\vec{x} \times \vec{y} \perp \vec{x}$ und $\vec{x} \times \vec{y} \perp \vec{y}$.
3. Es gilt

$$\begin{aligned} \|\vec{x} \times \vec{y}\|_2^2 &= \langle \vec{x} \times \vec{y}, \vec{x} \times \vec{y} \rangle = \langle \vec{x}, \vec{x} \rangle \langle \vec{y}, \vec{y} \rangle - \langle \vec{x}, \vec{y} \rangle^2 = \\ &= \|\vec{x}\|_2^2 \|\vec{y}\|_2^2 - \cos^2 \angle(\vec{x}, \vec{y}) \|\vec{x}\|_2^2 \|\vec{y}\|_2^2 = \|\vec{x}\|_2^2 \|\vec{y}\|_2^2 \sin^2 \angle(\vec{x}, \vec{y}). \end{aligned}$$

4. Wie wir es oben teilweise gesehen haben, es gelten:

$$\vec{e}_1 \times \vec{e}_2 = \vec{e}_3, \quad \vec{e}_2 \times \vec{e}_3 = \vec{e}_1, \quad \vec{e}_3 \times \vec{e}_1 = \vec{e}_2,$$

“Rechte-Hand-Regel”.

Ergänzung. Ist

$$s : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3 \quad \text{bilinear und antisymmetrisch}$$

mit

$$s(\vec{e}_1, \vec{e}_2) = \vec{e}_3, \quad s(\vec{e}_2, \vec{e}_3) = \vec{e}_1, \quad s(\vec{e}_3, \vec{e}_1) = \vec{e}_2,$$

so ist $s(\vec{x}, \vec{y}) = \vec{x} \times \vec{y}$.

- 1.22. Das Spatprodukt im \mathbb{R}^3 .** Für $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^3$ setzen wir

$$[\vec{x}, \vec{y}, \vec{z}] := \langle \vec{x} \times \vec{y}, \vec{z} \rangle \in \mathbb{R},$$

das **Spatprodukt** von $\vec{x}, \vec{y}, \vec{z}$. Die Rechenregeln sind aus 1.21 zu entnehmen. Insbesondere ist

$$[\cdot, \cdot, \cdot] : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R} \quad \text{eine alternierende Trilinearform,}$$

und es gilt

$$[\vec{x}, \vec{y}, \vec{z}] = [\vec{y}, \vec{z}, \vec{x}] = [\vec{z}, \vec{x}, \vec{y}],$$

also man darf zyklisch die Argumente vertauschen.

Interpretation 1. Für $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^3$ kann man $[\vec{x}, \vec{y}, \vec{z}] \in \mathbb{R}$ als orientiertes Volumen, des von $\vec{x}, \vec{y}, \vec{z}$ aufgespannten Parallelepipeds (Spats) auffassen.

$$P := \{ \lambda \vec{x} + \mu \vec{y} + \nu \vec{z} : 0 \leq \lambda, \mu, \nu \leq 1 \}$$

auffassen.

Das Volumen des Parallelepipeds ist das Produkt der Grundfläche und der Höhe des Parallelepipeds. Die Grundfläche ist

$$\|\vec{x}\|_2 \cdot \|\vec{y}\|_2 \cdot |\sin \alpha| = \|\vec{x} \times \vec{y}\|_2,$$

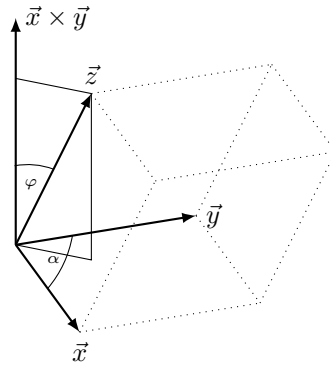
wobei $\alpha = \angle(\vec{x}, \vec{y})$. Die Höhe des Parallelepipeds ist

$$|\cos \varphi| \cdot \|\vec{z}\|_2,$$

wobei $\varphi = \angle(\vec{x} \times \vec{y}, \vec{z})$. Daher gilt

$$|[\vec{x}, \vec{y}, \vec{z}]| = |\langle \vec{x} \times \vec{y}, \vec{z} \rangle| = \|\vec{x} \times \vec{y}\|_2 \cdot \|\vec{z}\|_2 \cos \varphi = \text{Volumen von } P.$$

Die Orientierung des Volumen legen wir so fest, dass wir sagen, dass das Volumen von P positiv sei, wenn $[\vec{x}, \vec{y}, \vec{z}] \geq 0$.



Bemerkung. Mit dieser geometrischen Interpretation ist es leicht einzusehen, dass $[\vec{x}, \vec{y}, \vec{z}] = 0$ genau dann gilt, wenn $\vec{0}$, und $\vec{x}, \vec{y}, \vec{z}$ auf einer Ebene liegen.

Interpretation 2. Sei \vec{z} der Geschwindigkeitsvektor eines konstanten Flusses. Sei P das Parallelogramm aufgespannt von \vec{x} und \vec{y} . Dann fließt durch P pro Zeiteinheit das Volumen $[\vec{x}, \vec{y}, \vec{z}]$.

Ergänzungen.

1. Eine Merkregel:

$$[\vec{x}, \vec{y}, \vec{z}] = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix} = x_1 \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} - x_2 \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} + x_3 \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}$$

$$= x_1(x_2y_3 - x_3y_2) - x_2(x_1y_3 - x_3y_1) + x_3(x_1y_2 - x_2y_1).$$

2. Ist

$$s : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R} \quad \text{eine alternierende Trilinearform}$$

mit $s(\vec{e}_1, \vec{e}_2, \vec{e}_3) = 1$, so gilt $s(\vec{x}, \vec{y}, \vec{z}) = [\vec{x}, \vec{y}, \vec{z}]$. Wir haben gesehen, dass auch hier spielt die Orientierung des Raums eine wichtige Rolle. Bemerke aber an dieser Stelle, dass die Orientierung des Raums \mathbb{R}^3 bedeutet, dass wir eine alternierende Trilinearform auf \mathbb{R}^3 und die kanonische Basis $\vec{e}_1, \vec{e}_2, \vec{e}_3$ mit den obigen Eigenschaften festlegen.

3. Sei $\vec{e}_1, \dots, \vec{e}_n \in \mathbb{R}^n$ die kanonische Basis. Es gibt genau eine alternierende n -Linearform auf \mathbb{R}^n mit der Eigenschaft $s(\vec{e}_1, \dots, \vec{e}_n) = 1$. s heißt **Volumenform**. (Siehe Determinante 4.7).

1.23. Lineare Gleichungssysteme – Vorläufiges.

Definition. Ein **lineares Gleichungssystem** (LGS) ist ein System von k Gleichungen für n Unbekannten x_1, \dots, x_n der Form
indexlineares Gleichungssystem!homogen

$$(*) \quad \begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{k1}x_1 + a_{k2}x_2 + \dots + a_{kn}x_n &= b_k \end{aligned}$$

Gegeben sind also die Koeffizienten $a_{ij} \in K$, $b_i \in K$ in dem Körper K , und gesucht sind x_1, \dots, x_n aus dem Körper K . Dabei heißt $(*)$ **homogenes** LGS, falls

$$b_1 = b_2 = \dots = b_k = 0,$$

sonst **inhomogenes**. Ist $(*)$ inhomogen, so heißt das LGS mit denselben a_{ij} aber mit $b_1 = b_2 = \dots = b_k := 0$ das **zugehörige homogene LGS**.

Beispiele und Interpretation.

1. Seien

$$\vec{a}_1 := \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{k1} \end{pmatrix} \in K^k, \quad \vec{a}_2 := \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{k2} \end{pmatrix} \in K^k, \dots, \vec{a}_n := \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{kn} \end{pmatrix} \in K^k, \quad \vec{b} := \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{pmatrix} \in K^k$$

Finde $x_1, \dots, x_n \in K$, so dass

$$x_1\vec{a}_1 + x_2\vec{a}_2 + \dots + x_n\vec{a}_n = \vec{b}.$$

Das Lösen des LGSs ist also gleichwertig wie bestimmen Koeffizienten einer Linearkombination.

2. Wir erinnern uns an die Definition von Hyperebenen:

$$H_i := \left\{ \vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n : x_1a_{i1} + \dots + x_na_{in} = b_i \right\} = \left\{ \vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n : \langle \vec{x}, \vec{a}_i \rangle = b_i \right\}$$

= eine Hyperebene.

Wir sehen also: $\vec{x} \in K^n$ löst das LGS $(*)$, falls

$$\vec{x} \in H_1 \cap H_2 \cap \dots \cap H_k.$$

Bestimmen von Schnittmengen von Hyperebenen ist gleichwertig wie finden aller Lösungen eines LGSs.

3. Die Abbildung

$$A : K^n \rightarrow K^k, \quad \vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} \sum_{i=1}^n a_{1i}x_i \\ \vdots \\ \sum_{i=1}^n a_{ki}x_i \end{pmatrix} = A(\vec{x}) \in K^k$$

ist linear. Für gegebenes $\vec{b} \in K^k$ eine Lösung \vec{x} von (*) ist also ein Urbild von \vec{b} unter der Abbildung A , d.h.

$$A(\vec{x}) = \vec{b}.$$

1.24. Lösung von LGS – das Gauß'sche Eliminationsverfahren.

Beobachtung. Falls (*) von der Form

$$\begin{aligned} a_{12}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\ a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2 \\ a_{33}x_3 + \cdots + a_{3n}x_n &= b_3 \\ &\vdots \end{aligned}$$

ist, so ist die Lösung leicht zu finden: Beginne von unten und setze ein.

Definition. Das LGS (*) hat **Zeilenstufenform**, falls für jedes i, j mit $1 \leq i < j \leq k$ die j -te Zeile links mehr Nullen hat als die i -te Zeile, oder lauter Nullen besteht. Also:

$$\begin{array}{cccccccc} 0 & \boxed{\alpha_{1j_1}x_{j_1}} & \dots & & & & & = b_1 \\ \dots & \dots & 0 & \boxed{\alpha_{2j_2}x_{j_2}} & \dots & & & = b_2 \\ & \dots & \dots & 0 & & & & \\ & & \dots & \dots & 0 & \boxed{\alpha_{ij_i}x_{j_i}} & \dots & = b_i \\ & & & \dots & \dots & 0 & & \\ & & & & \dots & \dots & 0 & \boxed{\alpha_{lj_\ell}x_{j_\ell}} & \dots & = b_\ell \\ 0 & \dots & & & & & & 0 & = b_{l+1} \\ \vdots & & & & & & & & \vdots \\ 0 & \dots & & & & & & 0 & = b_k \end{array}$$

Klar ist also:

1. Falls $b_{l+1} \neq 0$, oder $b_{l+2} \neq 0$, ..., oder $b_k \neq 0$, dann gibt es *keine* Lösung.
2. Andernfalls kann man das LGS auflösen (beginne von unten und setze ein).

Gauß'sches Eliminationsverfahren. Bringe (*) in ein äquivalentes LGS, das Zeilenstufenform hat:

1. Durch Vertauschen von Zeilen erreicht man: *keine* Zeile geht *weiter* nach links als die erste.
2. Durch Addition geeigneter Vielfacher der ersten Zeile zu den anderen erreicht man: *Jede* Zeile ab der Zweiten geht *weniger* weit nach links als die erste.
3. "Vergiss" die erste Zeile, betrachte die zweite als erste und fahre mit 1. und 2. fort usw...

Beispiel. Löse das LGS

$$\begin{aligned} x_1 + 3x_2 - 5x_3 &= 1 \\ x_1 + 2x_2 - x_3 &= 2 \\ x_1 + x_2 + 2x_3 &= 3 \end{aligned}$$

Schritt 1. und 2. auf die erste Zeile:

$$\begin{aligned} x_1 + 3x_2 - 5x_3 &= 1 \\ -x_2 + 4x_3 &= 1 \\ -2x_2 + 7x_3 &= 2 \end{aligned}$$

Schritt 1. und 2. auf die zweite Zeile:

$$\begin{aligned}x_1 + 3x_2 - 5x_3 &= 1 \\- x_2 + 4x_3 &= 1 \\-x_3 &= 0\end{aligned}$$

Dies hat Zeilenstufenform! Durch einsetzen bekommt man die Lösungen:

$$\begin{aligned}x_3 &= 0 \\x_2 &= -1 \\x_1 &= 4.\end{aligned}$$

2. LINEARE UNABHÄNGIGKEIT, BASEN

2.1. Problematik.

- Sei $E = \{\vec{x} = \vec{a} + \lambda\vec{b} + \mu\vec{c} : \lambda, \mu \in \mathbb{R}\}$ eine Ebene. Wirklich? Falls nicht, wieso?
 $\vec{y} \in E \rightarrow \lambda = ?, \mu = ? \rightarrow \lambda, \mu$ eindeutig bestimmt \rightsquigarrow **Unabhängigkeit**
- Auch Funktionen sind Vektoren.
 Also eine typische Aufgabe könnte sein: gegeben Funktionen $g_1, \dots, g_n, f : [0, 1] \rightarrow \mathbb{R}$, finde $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ mit $\lambda_1 g_1 + \lambda_2 g_2 + \dots + \lambda_n g_n = f$. \rightsquigarrow Wir brauchen **Koordinaten**.
- Es gibt so viele Vektorräume. Wie verschieden sind sie und wie groß? \rightsquigarrow **Dimension**

A C H T U N G : ab jetzt verzichten wir auf der kleinen Pfeile wenn wir Vektoren notieren möchten.
Wir schreiben also x statt \vec{x} .

2.2. Lineare Hülle. Sei V ein Vektorraum über dem Körper K und $W \subseteq V$. Setze

$\text{Lin}(W) :=$ Menge aller Linearkombinationen aus W

$$= \left\{ v \in V : \exists n \in \mathbb{N} \exists \lambda_1, \dots, \lambda_n \in K \exists w_1, \dots, w_n \text{ s.d. } v = \sum_{i=1}^n \lambda_i w_i \right\}.$$

Eine Konvention: $\text{Lin}(\emptyset) = \{0\}$. Die Menge $\text{Lin}(W) \subseteq V$ ist Untervektorraum von V und heißt **lineare Hülle** oder **Aufspann** von W . Dabei heißt W **erzeugendes System** für $\text{Lin}(W)$.

Bemerkung. Häufig besteht W aus endlich vielen Elementen $\{w_1, \dots, w_m\}$. Dann ist

$$\text{Lin}(W) = \left\{ \sum_{i=1}^m \lambda_i w_i : \lambda_i \in K \right\}.$$

Beispiel. Zur Illustration der obigen Begriffe betrachte die folgenden Beispiele:

1. Sei $V = \mathbb{R}^n$, $K = \mathbb{R}$, und sei $W = \{e_1, \dots, e_n\}$ die kanonische Basis. Dann gilt

$$\text{Lin}(W) = V.$$

2. Sei $V = E_3$, $K = \mathbb{R}$, und sei

$$W = \left\{ \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

Dann gilt

$$\text{Lin}(W) = x\text{-}y \text{ Ebene.}$$

3. Sei $V = \{f : \mathbb{R} \rightarrow \mathbb{R} \text{ Funktionen}\}$, $K = \mathbb{R}$, und sei $W = \{x \mapsto x^n : n \in \mathbb{N} \cup \{0\}\}$. Dann gilt

$$\text{Lin}(W) = \text{Vektorraum der Polynome.}$$

4. Offensichtlich gilt $\text{Lin}(V) = V$.

Satz 2.2.1. Seien $v_1, \dots, v_n \in V$ ($n \geq 1$) n -Vektoren. Dann sind äquivalent:

(i) Ist

$$v = \sum_{i=1}^n \lambda_i v_i,$$

so sind die **Koeffizienten** $\lambda_i \in K$ eindeutig bestimmt. D.h., falls

$$\sum_{i=1}^n \lambda_i v_i = \sum_{i=1}^n \mu_i v_i$$

dann gilt für alle $i = 1, \dots, n$

$$\lambda_i = \mu_i.$$

(ii) Falls

$$\sum_{i=1}^n \lambda_i v_i = 0$$

gilt für $\lambda_i \in K$, $i = 1, \dots, n$, so ist

$$\lambda_i = 0 \quad \text{für alle } i = 1, \dots, n.$$

(iii) Keiner der Vektoren v_1, \dots, v_n ist Linearkombination der übrigen.

(iv) Die lineare Hülle jeder echten Teilsystem von v_1, \dots, v_n ist echt kleiner als $\text{Lin}(\{v_1, \dots, v_n\})$.

Beweis. (i) \Rightarrow (ii): $\sum_{i=1}^n 0 \cdot v_i = 0 = \sum_{i=1}^n \lambda_i v_i \Rightarrow \lambda_i = 0$ nach (i).

(ii) \Rightarrow (i): Ist $v = \sum_{i=1}^n \lambda_i v_i$ und $v = \sum_{i=1}^n \mu_i v_i$, dann $0 = v - v = \sum_{i=1}^n (\lambda_i - \mu_i) v_i$. Somit gilt $(\lambda_i - \mu_i) = 0$ für alle $i = 1, \dots, n$ nach (ii).

(ii) \Rightarrow (iii) (also $\neg(\text{iii}) \Rightarrow \neg(\text{ii})$): O.B.d.A. sei $v_n = \sum_{i=1}^{n-1} \lambda_i v_i$, dann

$$0 = \sum_{i=1}^{n-1} \lambda_i v_i - v_n,$$

also gilt (ii) auch nicht.

(iii) \Rightarrow (ii) (also $\neg(\text{ii}) \Rightarrow \neg(\text{iii})$) Ist $0 = \sum_{i=1}^n \lambda_i v_i$ mit $\lambda_n \neq 0$, dann

$$v_n = - \sum_{i=1}^{n-1} \frac{\lambda_i}{\lambda_n} v_i,$$

also gilt (iii) auch nicht.

(iii) \Leftrightarrow (iv): $v_n = \sum_{i=1}^{n-1} \lambda_i v_i$ gilt genau dann, wenn $\text{Lin}(\{v_1, \dots, v_n\}) = \text{Lin}(\{v_1, \dots, v_{n-1}\})$. ■

Definition. a) Das System der Vektoren v_1, \dots, v_n heißt **linear unabhängig**, falls sie die äquivalenten Bedingungen des Satzes erfüllen, sonst **linear abhängig**.

b) Allgemeiner heißt eine Familie $(v_i)_{i \in I}$ linear unabhängig, falls jedes endliche Teilsystem linear unabhängig ist. (Konvention: \emptyset sei linear unabhängig.)

c) Ist darüber hinaus

$$V = \text{Lin}(\{v_1, \dots, v_n\})$$

(oder $V = \text{Lin}(\{v_i : i \in I\})$), d.h. jeder Vektor in V ist eine endliche Linearkombination der v_i , dann heißt $\{v_i : i \in I\}$ eine **Basis** von V .

Bemerkung. Ist v_i , $i \in I$ eine Basis in V , so läßt sich jeder Vektor in V auf *genau eine Weise* als endliche Linearkombination von Vektoren in $\{v_i : i \in I\}$ schreiben.

Bemerkung. Ist $\{v_1, \dots, v_n\}$ eine Basis von V und

$$w = \sum_{i=1}^n \lambda_i v_i,$$

so heißen die Koeffizienten λ_i auch die **Koordinaten** von w bezüglich dieser Basis. Der Vektor

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

heißt der **Koordinatenvektor** von v bzgl. der gegebenen Basis.

Bemerkung. Ist $\{v_1, \dots, v_n\}$ eine Basis von V , so ist die **Dimension** von V über K eben n . Schreibweise: $\text{Dim}(V) = n$. (Später: Dimension ist unabhängig von der Basis.) z.B. $\{e_1, \dots, e_n\}$ als kanonische Basis in \mathbb{R}^n , daraus folgt: \mathbb{R}^n ist n -dimensional über \mathbb{R} .

2.3. Diskussion.

1. $\{v\}$ ist linear abhängig $\Leftrightarrow v = 0$.

2. $\{v_1, v_2\}$ linear abhängig $\Leftrightarrow v_1 = 0$ oder $v_2 = 0$ oder $\lambda v_1 = v_2$.
(**parallele** Vektoren), die zwei Vektoren liegen auf einer Ursprungsgeraden.
3. $\{x_1, x_2, x_3\} \subseteq \mathbb{R}^3$
 - linear unabhängig \Leftrightarrow Linearkombinationen erzeugen \mathbb{R}^3 .
 - linear abhängig \Leftrightarrow Liegen in einer Ebene durch 0 (oder sogar auf einer Geraden).
4. Falls $0 \in \{v_1, \dots, v_n\}$, dann sind $\{v_1, \dots, v_n\}$ linear abhängig.
5. $\{v_i : i \in I\}$ linear unabhängig \Rightarrow jede Teilmenge ist linear unabhängig.
6. $\{v_i : i \in I\}$ linear unabhängig $\Rightarrow \{v_i : i \in I\}$ ist Basis von

$$W = \text{Lin}(\{v_i : i \in I\}).$$
7. $\{v_i : i \in I\}$ Basis von V , $w \in V \Rightarrow$ Das System $v_i, i \in I$ zusammen mit w ist linear abhängig.

Beispiel. 1. Sei K ein Körper und $e_i \in K^n, i = 1, \dots, n$ die **kanonische Basisvektoren**, d.h.

$$e_{ij} = j\text{-te Koordinate von } e_i = \delta_{ij} = \mathbf{Kronecker-Delta}.$$

Dann ist $\{e_i : i = 1, \dots, n\} \subseteq K^n$ eine Basis von K^n .

2. Sei $p_n : \mathbb{R} \ni x \mapsto x^n \in \mathbb{R}$ (bzw. $p_n : \mathbb{C} \ni z \mapsto z^n \in \mathbb{C}$), dann ist $\{p_n : n \in \mathbb{N} \cup \{0\}\}$ im Vektorraum der Funktionen linear unabhängig.

Beweis: $\sum_{i=0}^n \lambda_i p_i = 0 \Rightarrow \sum_{i=0}^n \lambda_i p_i(x) = \sum_{i=0}^n \lambda_i x^i = 0$.

Das Polynom hat mehr als n Nullstellen \Rightarrow ist das Null-Polynom, also $\lambda_i = 0$.

3. Die Funktionen $f = \text{const} = 1, \{\cos nx : n \in \mathbb{N}\}$ und $\{\sin nx : n \in \mathbb{N}\}$ sind auf $[0, 2\pi]$ linear unabhängig. Der Beweis kommt später.
4. Sind $a_1, \dots, a_k \in K^n$, so dass a_j links mehr Nullen hat als a_i für $j > i$, dann sind a_i, \dots, a_k linear unabhängig.

2.4. Orthonormalsysteme.

Definition. Ein System $\{e_1, \dots, e_k\} \in \mathbb{C}^n$ (oder \mathbb{R}^n) heißt **Orthonormalsystem** (ONS), falls

$$\langle e_i, e_j \rangle = \delta_{ij} := \begin{cases} 1 & \text{für } i = j, \\ 0 & \text{für } i \neq j, \end{cases}$$

(δ_{ij} wird **Kronecker-Delta** genannt).

Also: $e_i \perp e_j$ (orthogonal) und $\|e_i\|_2 = \sqrt{\langle e_i, e_i \rangle} = 1$ (normiert)

Beispiel. 1. Die kanonische Basisvektoren in \mathbb{C}^n bilden ein Orthonormalsystem.

2. $\begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}, \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix}$ bilden ein Orthonormalsystem.

Satz 2.4.1. Sei $\{e_1, \dots, e_k\}$ ein Orthonormalsystem $\Rightarrow \{e_1, \dots, e_k\}$ linear unabhängig.

Beweis. Sei

$$0 = \sum_{i=1}^k \lambda_i e_i.$$

Wir müssen zeigen, dass die Koeffizienten λ_i alle 0 sind. Multipliziere mit e_j für ein $j = 1, \dots, n$:

$$0 = \left\langle \sum_{i=1}^k \lambda_i e_i, e_j \right\rangle = \sum_{i=1}^k \lambda_i \langle e_i, e_j \rangle = \lambda_j. \quad \blacksquare$$

2.5. Warum sind linear unabhängige Vektoren wichtig? Wir wollen zeigen: Mit linear unabhängigen Vektoren kann man Koordinaten einführen und wie im \mathbb{R}^n (oder im \mathbb{C}^n) rechnen. Sei V Vektorraum über dem Körper K , und seien $v_1, \dots, v_n \in V$. Definiere:

$$T : K^n \rightarrow V, \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \mapsto \sum_{i=1}^n \lambda_i v_i \in V.$$

Dann ist T offensichtlich linear.

Satz 2.5.1. a) $\text{Lin}(\{v_1, \dots, v_n\}) = V \Leftrightarrow T$ ist surjektiv.

- b) $\{v_1, \dots, v_n\}$ linear unabhängig $\Leftrightarrow T$ ist injektiv.
 c) $\{v_1, \dots, v_n\}$ Basis von $V \Leftrightarrow T$ ist bijektiv.

Beweis. a) “ \Rightarrow ”: Ist T surjektiv so existieren für $w \in V$ Elemente $\lambda_1, \dots, \lambda_n \in K$ mit

$$w = \sum_{i=1}^n \lambda_i v_i = T \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

“ \Leftarrow ”: Für $w \in V$ existieren $\lambda_1, \dots, \lambda_n \in K$ mit

$$w = T \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \sum_{i=1}^n \lambda_i v_i,$$

d.h. T ist surjektiv.

b) Es gilt

$$T \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = T \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$$

genau dann wenn

$$\sum_{i=1}^n \lambda_i v_i = \sum_{i=1}^n \mu_i v_i \Leftrightarrow \sum_{i=1}^n (\lambda_i - \mu_i) v_i = 0.$$

Also: T nicht injektiv $\Leftrightarrow \{v_1, \dots, v_n\}$ nicht linear unabhängig.

c) folgt aus a) und b). ■

Definition. Seien V, W Vektorräume über K . Ist $T : V \rightarrow W$ linear und bijektiv, so heißt T (**Vektorraum-**) **Isomorphismus**. Die lineare Abbildung T^{-1} ist dann ebenfalls ein Vektorraum-Isomorphismus. V und W heißen **isomorph**, falls es einen Vektorraum-Isomorphismus $T : V \rightarrow W$ gibt. Wir schreiben $V \simeq W$. Isomorphe Vektorräume sind als Vektorräume gleich. Also: Ist $\{v_1, \dots, v_n\}$ eine Basis von V (wobei V ein Vektorraum über K ist), so ist V isomorph zu K^n .

Beispiel. Sei $V = \{\text{Polynome vom Grad } \leq n\}$ über $K = \mathbb{R}$. Betrachte die Polynome

$$p_0 = 1, p_1(x) = x, \dots, p_n(x) = x^n.$$

Dann ist p_0, \dots, p_n eine Basis in V . Die Abbildung

$$T : \mathbb{R}^{n+1} \rightarrow V, \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} \mapsto a_0 + a_1 x + \dots + a_n x^n$$

ist ein Vektorraum-Isomorphismus.

2.6. Frage: Besitzt jeder Vektorraum eine Basis?

Definition. Ein Vektorraum V heißt **endlich erzeugt**, falls es endlich viele Vektoren $v_1, \dots, v_k \in V$ gibt mit $V = \text{Lin}(\{v_1, \dots, v_k\})$.

Hauptsatz. Sei V ein Vektorraum.

- Ist V endlich erzeugt, so besitzt V eine Basis.
- Besitzt V eine Basis aus n Elementen, so sind je $n + 1$ Vektoren in V linear abhängig.
- Besitzt V eine Basis aus n Elementen, so hat jede Basis von V genau n Elemente.

Beweis. a) Durch Ausdünnen: Sei $V = \text{Lin}(\{v_1, \dots, v_k\})$. Streiche solange wie möglich Vektoren aus $\{v_1, \dots, v_n\}$ heraus, so dass die lineare Hülle des Restes immer noch V ergibt. Wenn nichts mehr geht, hat man eine Basis (vgl. 2.2.1 (iii)). Dies geschieht nach endlich vielen Schritten.

b) Wir wissen $V \simeq K^n$. Es reicht also, die Behauptung in K^n zu beweisen. Seien

$$a_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}, \dots, a_{n+1} = \begin{pmatrix} a_{1(n+1)} \\ \vdots \\ a_{n(n+1)} \end{pmatrix}$$

$n + 1$ Vektoren in K^n . a_i sind linear abhängig $\Leftrightarrow \sum_{i=1}^{n+1} \lambda_i a_i = 0$ hat nicht-triviale Lösungen $\lambda_1, \dots, \lambda_n \in K$. Dies ist genau dann, wenn das LGS

$$\begin{array}{rcccc} \lambda_1 a_{11} & + \cdots + & \lambda_{n+1} a_{1(n+1)} & = & 0 \\ \vdots & & \vdots & & \vdots \\ \lambda_1 a_{n1} & + \cdots + & \lambda_{n+1} a_{n(n+1)} & = & 0 \end{array}$$

mit n Zeilen und $n + 1$ Spalten nicht triviale Lösungen hat. Umformen auf Zeilenstufenform \Rightarrow Behauptung (vgl. 1.24).

Zur Illustration $n = 2$:

$$\begin{array}{l} \lambda_1 a_{11} + \lambda_2 a_{12} + \lambda_3 a_{13} = 0 \\ \lambda_1 a_{21} + \lambda_2 a_{22} + \lambda_3 a_{23} = 0 \end{array} \rightsquigarrow \begin{array}{l} \lambda_1 a_{11} + \lambda_2 a_{12} + \lambda_3 a_{13} = 0 \\ \lambda_2 b_{22} + \lambda_3 b_{23} = 0 \end{array} \rightsquigarrow \text{nicht-triviale Lösung}$$

c) Folgt aus b): Seien $\{v_1, \dots, v_n\} \subseteq V$ und $\{w_1, \dots, w_k\} \subseteq V$ Basen. Dann:

$$k > n \Rightarrow \{w_1, \dots, w_k\} \text{ linear abhängig} \Rightarrow \text{Widerspruch;}$$

$$n > k \Rightarrow \{v_1, \dots, v_n\} \text{ linear abhängig} \Rightarrow \text{Widerspruch.} \quad \blacksquare$$

F O L G E R U N G : Ein endlich erzeugter Vektorraum V über K ist isomorph zu einem Vektorraum der Form K^n für genau ein n .

Definition. Ist $\{v_1, \dots, v_n\} \subseteq V$ Basis von V , dann heißt n die **Dimension von V** über K . Schreibweise:

$$\text{Dim}(V) = n.$$

Hier ist n eindeutig bestimmt. Besitzt V keine endliche Basis, dann heißt V **unendlich-dimensional**, $\text{Dim}(V) = \infty$.

2.7. Ergänzung.

1. Hauptsatz 2.6.a) gilt allgemein: Jeder Vektorraum hat eine Basis! Beweis mit Zornschem Lemma.
2. Meist leitet man Hauptsatz 2.6 aus dem **Basisergänzungssatz** ab, der folgenderweise lautet: Seien $v_1, \dots, v_k \in V$ linear unabhängig und $w_1, \dots, w_\ell \in V$ so, dass

$$\text{Lin}(\{v_1, \dots, v_k, w_1, \dots, w_\ell\}) = V.$$

Dann kann man durch Hinzunahme geeigneter Vektoren aus $\{w_1, \dots, w_\ell\}$ die Menge $\{v_1, \dots, v_k\}$ in eine Basis ergänzen.

Mithilfe dieses Satzes ist klar:

Korollar. Sei V ein n -dimensionaler Vektorraum, $U \subseteq V$ ein Unterraum, dann ist

$$\text{Dim } U \leq \text{Dim}(V); \quad \text{und} \quad \text{Dim } U = \text{Dim}(V) \quad \text{genau dann, wenn} \quad U = V.$$

2.8. Skalarprodukte. Berechnung der Koeffizienten ist am Besten bzgl. Orthonormalbasen!

Definition. Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ (oder $= \mathbb{C}$). Eine Abbildung

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{K} : (u, v) \mapsto \langle u, v \rangle$$

heißt (abstraktes) **Skalarprodukt** auf V , falls für alle $u, v, w \in V$ und $\lambda, \mu \in \mathbb{K}$ gilt:

$$\text{(SP1)} \quad \langle \lambda u + \mu v, w \rangle = \lambda \langle u, w \rangle + \mu \langle v, w \rangle,$$

$$\text{(SP2)} \quad \langle u, v \rangle = \overline{\langle v, u \rangle},$$

$$\text{(SP3)} \quad \langle u, u \rangle \geq 0 \text{ und } \langle u, u \rangle = 0 \Leftrightarrow u = 0,$$

Ein Vektorraum mit Skalarprodukt $(V, \langle \cdot, \cdot \rangle)$ heißt auch **prä-Hilbertraum**, **euklidischer Raum** ($\mathbb{K} = \mathbb{R}$) oder **unitärer Raum** ($\mathbb{K} = \mathbb{C}$).

Diskussion.

1. (SP1) bedeutet: für alle $w \in V$ fest ist

$$V \rightarrow \mathbb{K} : u \mapsto \langle u, w \rangle \quad \text{linear.}$$

2. Im Falle $\mathbb{K} = \mathbb{C}$, (SP1) + (SP2) implizieren: für alle $w \in V$ ist

$$V \rightarrow \mathbb{K} : u \mapsto \langle w, u \rangle \quad \text{antilinear (oder konjugiert-linear),}$$

d.h.

$$\langle w, \lambda u + \mu v \rangle = \bar{\lambda} \langle w, u \rangle + \bar{\mu} \langle w, v \rangle.$$

3. Es gilt genau dann $\langle u, v \rangle = 0$ für alle $v \in V$, wenn $u = 0$ (setze $v = u$).

4. In der Physik sind Skalarprodukte oft antilinear in der ersten Komponente (Dirac-Schreibweise in der Quantenmechanik) \rightsquigarrow Geht aber genauso.

Beispiel. 1. Das Standard-Skalarprodukt $\langle \cdot, \cdot \rangle_{\mathbb{R}^n}$ auf \mathbb{R}^n ($\langle \cdot, \cdot \rangle_{\mathbb{C}^n}$ auf \mathbb{C}^n) ist tatsächlich ein Skalarprodukt (siehe 1.12).

2. Sei $[a, b] \subseteq \mathbb{R}$ ein Intervall. Für $f, g : [a, b] \rightarrow \mathbb{C}$ stetig setze

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

Allgemeiner: Sei $w : [a, b] \rightarrow \mathbb{R}_+$ strikt positiv "Gewichtsfunktion". Setze:

$$\langle f, g \rangle_w := \int_a^b f(x) \overline{g(x)} w(x) dx.$$

So sind beide Skalar-Produkte auf dem Vektorraum

$$V = \left\{ \text{Stetige Funktionen auf } [a, b] \right\}.$$

Skalarprodukte sind also auch wichtig in Analysis.

2.9. Cauchy-Schwartz-Ungleichung (vgl. 1.12).

Satz. Für $x, y \in V$ gilt

$$|\langle x, y \rangle| \leq \langle x, x \rangle^{1/2} \cdot \langle y, y \rangle^{1/2}$$

und die Gleichheit

$$|\langle x, y \rangle| = \langle x, x \rangle^{1/2} \cdot \langle y, y \rangle^{1/2}$$

gilt genau dann, wenn x, y linear abhängig sind.

Beweis. 1. Ist $\langle x, y \rangle = 0$, so ist die Ungleichung trivial, und auch die Aussage über die Gleichheit folgt leicht.

2. Nehme an $\langle x, y \rangle \neq 0$: Wir setzen

$$\alpha = \frac{\overline{\langle x, y \rangle}}{|\langle x, y \rangle|^2},$$

dann gilt $\alpha \in \mathbb{K}$ mit $|\alpha| = 1$. Betrachte:

$$p(\lambda) := \langle \alpha x + \lambda y, \alpha x + \lambda y \rangle = \langle x, x \rangle + \alpha \bar{\lambda} \langle x, y \rangle + \bar{\alpha} \lambda \langle y, x \rangle + |\lambda|^2 \langle y, y \rangle.$$

Für $\lambda \in \mathbb{R}$ gilt:

$$p(\lambda) = \langle x, x \rangle + \lambda \cdot 2 \operatorname{Re} \alpha \langle x, y \rangle + \lambda^2 \cdot \langle y, y \rangle = \langle x, x \rangle + 2|\langle x, y \rangle| \cdot \lambda + \langle y, y \rangle \cdot \lambda^2.$$

Also ist $p(\lambda)$ ein Polynom mit Grad $p = 2$ und $p(\lambda) \geq 0$ gilt für alle $\lambda \in \mathbb{R}$. D.h. $p(\lambda) = 0$ hat maximal eine \mathbb{R} -Lösung und die Diskriminante " $b^2 - 4ac$ " ist nicht-positiv:

$$4|\langle x, y \rangle|^2 - 4\langle x, x \rangle \langle y, y \rangle \leq 0.$$

Daraus folgt die behauptete Ungleichung. Hat man hier "=", so heißt das, dass $p(\lambda) = 0$ genau eine Lösung $\lambda_0 \in \mathbb{R}$ hat, also

$$p(\lambda_0) = 0 \stackrel{\text{(SP3)}}{\implies} \alpha x + \lambda_0 y = 0 \implies x = -\frac{\lambda_0}{\alpha} y \implies \text{linear abhängig} \quad \blacksquare$$

Bemerkung. Sei V ein Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$, dann ist

$$\|x\| := \langle x, x \rangle^{1/2}$$

eine Norm auf V , denn:

1. $\|x\| = \langle x, x \rangle^{1/2} \geq 0$.

2. $\|\lambda x\| = \langle \lambda x, \lambda x \rangle^{1/2} = |\lambda| \langle x, x \rangle^{1/2} = |\lambda| \cdot \|x\|.$

3. Zum Beweis der Dreiecksungleichung verwendet man die Cauchy-Schwartz-Ungleichung:

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

2.10. Quadratische Form.

Definition. Die Abbildung

$$Q : V \rightarrow R_+, \quad v \mapsto Q(v) := \langle v, v \rangle = \|v\|^2$$

heißt auch die zugehörige **quadratische Form** (denn $Q(\lambda x) = |\lambda|^2 Q(x)$).

Ist $\mathbb{K} = \mathbb{R}$ so ist:

$$\langle x, y \rangle = \frac{1}{4}(Q(x+y) - Q(x-y))$$

Ist $\mathbb{K} = \mathbb{C}$ so ist:

$$\langle x, y \rangle = \frac{1}{4}(Q(x+y) - Q(x-y) - iQ(x-iy) + iQ(x+iy))$$

Beweis. durch Nachrechnen ■

WICHTIG: Polarisierung: Ein Skalarprodukt ist also durch die zugehörige quadratische Form schon vollständig festgelegt.

2.11. Orthogonalität (vgl. 1.12). Sei V ein Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$.

Definition. a) $x, y \in V$ heißen **orthogonal** (in Notation: $x \perp y$), falls $\langle x, y \rangle = 0$.

b) $M, N \subseteq V$ heißen **orthogonal** ($M \perp N$), falls

$$x \perp y \quad \text{für alle } x \in M, y \in N.$$

c) Für $M \subseteq V$ heißt

$$M^\perp := \{y \in V : y \perp x \quad \forall x \in M\}$$

das **orthogonale Komplement** von M . M^\perp ist ein linearer Unterraum in V (auch wenn M nur eine Teilmenge ist).

d) Eine Familie $(v_i)_{i \in I}$ heißt **Orthonormalsystem** (ONS), falls

$$v_i \perp v_j \quad i \neq j \quad \text{(Orthogonal)}$$

$$\|v_i\| = 1. \quad \text{(Normalisiert)}$$

Ist $\{e_1 \dots e_n\}$ Orthonormalsystem und Basis, so heißt es auch **Orthonormalbasis** (ONB).

Bemerkung. 1. Satz von Pythagoras: $x \perp y \Rightarrow \|x + y\|^2 = \|x\|^2 + \|y\|^2$ (vgl. 1.12)

2. $(v_i)_{i \in I}$ ist Orthonormalsystem $\Rightarrow \{v_i \mid i \in I\}$ ist linear unabhängig (vgl. 1.12)

2.12. Entwicklung nach Orthonormalsysteme.

Satz. Sei $\{e_1, \dots, e_n\}$ Orthonormalbasis für V .

a) Ist

$$v = \sum_{i=1}^n \lambda_i e_i,$$

so gilt $\lambda_i = \langle v, e_i \rangle$, d.h. $v = \sum_{i=1}^n \langle v, e_i \rangle e_i$.

b) Ist

$$u = \sum_{i=1}^n \lambda_i e_i, \quad v = \sum_{i=1}^n \mu_i e_i$$

so gilt

$$\langle u, v \rangle = \sum_{i=1}^n \lambda_i \cdot \bar{\mu}_i = \left\langle \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}, \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} \right\rangle_{\mathbb{C}^n}$$

Insbesondere gilt also

$$\|u\|^2 = \langle u, u \rangle = \sum_{i=1}^n |\lambda_i|^2 = \sum_{i=1}^n |\langle u, e_i \rangle|^2. \quad (\text{Parsevalsche Gleichung})$$

Beweis. a) Sei $v = \sum_{i=1}^n \lambda_i e_i$, und (skalar-)multipliziere mit e_j . So erhält man

$$\langle v, e_j \rangle = \left\langle \sum_{i=1}^n \lambda_i e_i, e_j \right\rangle = \sum_{i=1}^n \lambda_i \langle e_i, e_j \rangle = \lambda_j.$$

b) Seien u, v wie in der Behauptung, und betrachte ihr Skalarprodukt:

$$\langle u, v \rangle = \left\langle \sum_{i=1}^n \lambda_i e_i, \sum_{j=1}^n \mu_j e_j \right\rangle = \sum_{i,j=1}^n \lambda_i \bar{\mu}_j \langle e_i, e_j \rangle = \sum_{i=1}^n \lambda_i \bar{\mu}_i. \quad \blacksquare$$

Bemerkung. Die Koordinaten-Abbildung

$$T : \mathbb{K}^n \rightarrow V : \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \mapsto \sum_{i=1}^n \lambda_i e_i$$

führt das Standard-Skalarprodukt auf \mathbb{K}^n über in $\langle \cdot, \cdot \rangle_V$, d.h., T "erhält das Skalarprodukt und die Norm".

2.13. Beispiele.

1. Betrachte $V = C([0, 2\pi]) = \{f : [0, 2\pi] \rightarrow \mathbb{R} \text{ stetig}\}$, und

$$\langle f, g \rangle = \int_0^{2\pi} f(x)g(x)dx, \quad \|f\| = \langle f, f \rangle^{1/2} = \left(\int_0^{2\pi} f(x)^2 dx \right)^{1/2}.$$

Betrachte ferner die Funktionen:

$$e_0(x) = \frac{1}{\sqrt{2\pi}} \text{ (konstante Funktion), } \quad e_{2n}(x) = \frac{1}{\sqrt{\pi}} \cos(nx), \quad e_{2n-1}(x) = \frac{1}{\sqrt{\pi}} \sin(nx) \quad \text{für } n \geq 1.$$

Dann ist $\{e_n : n \geq 0\}$ ein Orthonormalsystem (insbesondere linear unabhängig). Also gilt: Ist $f \in \text{Lin}\{e_n : n \geq 0\}$, dann ist

$$f = \sum \lambda_n e_n = c \cdot \mathbb{1} + \sum_n a_n \sin nx + b_n \cos nx$$

eine endliche Summe mit

$$\begin{aligned} c &= \langle f, e_0 \rangle = \int_0^{2\pi} f(x) \cdot \frac{1}{\sqrt{2\pi}} dx, \\ a_n &= \langle f, e_{2n-1} \rangle = \frac{1}{\sqrt{\pi}} \int_0^{2\pi} f(x) \sin(nx) dx, \\ b_n &= \langle f, e_{2n} \rangle = \frac{1}{\sqrt{\pi}} \int_0^{2\pi} f(x) \cos(nx) dx. \end{aligned}$$

Theorie der Fourier-Reihen: Läßt man auch unendliche Summen zu, so kann man praktisch "jede" Funktion so entwickeln. (Fourier-Reihen sind also LA).

2. In \mathbb{R}^2 ist

$$e_1 := \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}, \quad e_2 := \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix}$$

eine Orthonormalbasis (vgl. 1.12). Also ist z.B.

$$\begin{pmatrix} 2 \\ 0 \end{pmatrix} = \lambda_1 e_1 + \lambda_2 e_2 \quad \text{mit} \quad \lambda_1 = 2 \cos \varphi, \quad \lambda_2 = -2 \sin \varphi.$$

Allgemein:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \lambda_1 e_1 + \lambda_2 e_2 \quad \text{mit} \quad \begin{aligned} \lambda_1 &= x \cos \varphi + y \sin \varphi, \\ \lambda_2 &= -x \sin \varphi + y \cos \varphi. \end{aligned}$$

2.14. Orthogonale Projektion. Sei V ein Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$, $\{e_1, \dots, e_n\}$ ein Orthonormalsystem und setze

$$V_0 := \text{Lin}\{e_1, \dots, e_m\}, \quad x \in V, \quad \lambda_i := \langle x, e_i \rangle.$$

Frage: Falls $V_0 \neq V$, was ist $\sum_{i=1}^m \lambda_i e_i$? Wir erinnern uns: $\lambda_i e_i$ ist die orthogonale Projektion auf die von e_i aufgespannte Gerade, siehe 1.15.

Satz (Wichtig!). Sei $x \in V$ mit

$$x = \sum_{i=1}^n \lambda_i e_i, \quad \text{d.h.} \quad \lambda_i = \langle x, e_i \rangle \quad \text{für } i = 1, \dots, n.$$

Für $y \in V_0$ sind äquivalent:

- a) $y = \sum_{i=1}^m \lambda_i e_i$.
- b) $x = y + z$ und $x - y = z \perp V_0$.
- c) $\|x - y\| \leq \|x - y'\|$ gilt für alle $y' \in V_0$.

Beweis. (**Nicht so wichtig!**) a) \Rightarrow b): Setze $z = x - y$ und sei $y' \in V_0$, d.h., y' kann so geschrieben werden: $y' = \sum_{i=1}^m \mu_i e_i$. Dann gilt

$$\langle z, y' \rangle = \langle x - y, y' \rangle = \langle x, y' \rangle - \langle y, y' \rangle = \sum_{i=1}^m \bar{\mu}_i \overbrace{\langle x, e_i \rangle}^{\lambda_i \text{ nach Def.}} - \sum_{i=1}^m \lambda_i \bar{\mu}_i = 0.$$

Also $z \perp V_0$.

b) \Rightarrow a): Nehme an $y \in V_0$ mit $z := x - y \perp V_0$. Dann

$$\langle y, e_i \rangle = \langle x - z, e_i \rangle = \langle x, e_i \rangle - \underbrace{\langle z, e_i \rangle}_{=0} = \langle x, e_i \rangle = \lambda_i \quad \text{für } i = 1, \dots, m.$$

b) \Rightarrow c): Betrachte die Zerlegung $x = y + z$ mit $z \perp V_0$ wie in b). Sei nun $y' \in V_0$ beliebig, dann $y' = y + \tilde{y}$ mit $\tilde{y} \in V_0$, also

$$\begin{aligned} \|x - y'\|^2 &= \|y + z - y - \tilde{y}\|^2 = \|z - \tilde{y}\|^2 = \|z\|^2 + \|\tilde{y}\|^2 \quad \text{da } z \perp \tilde{y} \\ &\geq \|z\|^2 = \|x - y\|^2. \end{aligned}$$

c) \Rightarrow a): Sei $\tilde{y} = \sum_{i=1}^m \lambda_i e_i$. Wegen der Äquivalenz "a) \Leftrightarrow b)" gilt $z := x - \tilde{y} \perp V_0$. Es gilt $y' := y - \tilde{y} \in V_0$, also wegen der Annahme c) gilt:

$$\begin{aligned} \|x - y\|^2 &\geq \|x - \tilde{y} - y'\|^2 = \|x - \tilde{y}\|^2 + \|y'\|^2 \quad \text{da } x - \tilde{y} \perp y' \\ &\geq \|x - y\|^2, \end{aligned}$$

und daher steht hier überall "=", d.h. $\|y'\| = 0$, also $y = \tilde{y}$. ■

Bemerkung. Also für jedes $x \in V$ existiert ein eindeutiges $y \in V_0$ mit $\|x - y\|$ minimal. Es ist

$$y = \sum_{i=1}^m \langle x, e_i \rangle e_i,$$

und die Abbildung

$$P : V \rightarrow V_0 : x \mapsto \sum_{i=1}^m \langle x, e_i \rangle e_i$$

ist linear und heißt die **Orthogonale Projektion** von V auf V_0 .

2.15. Das Gram-Schmidtsche Orthonormalisierungsverfahren.

Satz. Sei V ein Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und $\{v_1, v_2, \dots\}$ eine endliche oder abzählbare Familie linear unabhängiger Vektoren. Dann gibt es ein Orthonormalsystem $\{e_1, e_2, \dots\}$ in V so dass

$$\text{Lin}\{v_1, v_2 \dots v_k\} = \text{Lin}\{e_1, e_2 \dots e_k\} \quad \text{gilt für alle } k \in \mathbb{N}, k \leq V.$$

Beweis. (**konstruktiv, wichtig!**)

1. Sei $e_1 := \frac{v_1}{\|v_1\|}$.
2. Sei $f_2 := v_2 - \langle v_2, e_1 \rangle e_1$ und $e_2 = \frac{f_2}{\|f_2\|}$. Dann $e_2 \perp e_1$, $e_2 \perp v_1$ (vgl. 2.14).
3. Seien e_1, \dots, e_{k-1} schon konstruiert mit

$$\text{Lin}\{v_1 \dots v_{k-1}\} = \text{Lin}\{e_1 \dots e_{k-1}\}.$$

Dann $v_k \notin \text{Lin}\{e_1, \dots, e_{k-1}\}$. Setze

$$f_k := v_k - \sum_{i=1}^{k-1} \langle v_k, e_i \rangle e_i. \quad (\text{vgl. 2.24!})$$

Dann gilt $f_k \neq 0$ und $f_k \perp \text{Lin}\{e_1, \dots, e_{k-1}\}$. Wir können also setzen

$$e_k := \frac{f_k}{\|f_k\|}. \quad \blacksquare$$

Folgerung: Ist $\text{Dim}(V) < \infty$, so besitzt V eine Orthonormalbasis und jedes Orthonormalsystem läßt sich zu Orthonormalbasis ergänzen.

3. LINEARE ABBILDUNGEN UND MATRIZEN

3.1. Erinnerung: Lineare Abbildungen und lineare Teilräume. $T : V \rightarrow W$ heißt linear, falls

- $T(u + v) = T(u) + T(v)$ gilt für alle $u, v \in V$,
- $T(\lambda u) = \lambda T(u)$ gilt für alle $\lambda \in K, u \in V$

Die Menge

$$\mathcal{L}(V, W) := \{T : V \rightarrow W \text{ linear}\}$$

ist selber ein Vektorraum, mit den üblichen Operationen (siehe unten). $V_0 \subseteq V$ heißt linearer Teilraum, falls $u + v \in V_0$ für $u, v \in V_0$ und $\lambda u \in V_0$ für $u \in V_0, \lambda \in K$

Definition. Sei $T : V \rightarrow W$ linear. Wir definieren

$\text{Kern}(T) := \{u \in V : T(u) = 0\}$ ist linearer Teilraum von V und heißt **Kern von T** .

$\text{Bild}(T) := \{T(u) : u \in V\}$ ist linearer Teilraum von W und heißt **Bild von T** .

Allgemein: Ist $V_0 \subseteq V$ linearer Teilraum, dann

$$T(V_0) := \{T(u) : u \in V_0\} \text{ linearer Teilraum.}$$

Ist $W_0 \subseteq W$ linearer Teilraum, dann

$$T^{-1}(W_0) := \{u \in V : T(u) \in W_0\} \text{ linearer Teilraum.}$$

Es gilt also

$$\text{Kern}(T) = T^{-1}(\{0\}).$$

3.2. Operationen auf linearen Abbildungen. Seien U, V, W Vektorräume.

- $I : V \rightarrow V : x \mapsto x$ heißt **identische Abbildung/Identität**.
- $O : V \rightarrow W : x \mapsto 0 \in W$ (0 ist das Nullelement in W) heißt **Nullabbildung**.
- Sind $T, S : V \rightarrow W$ lineare Abbildungen, so sind

$$\lambda T : V \rightarrow W, x \mapsto \lambda Tx \quad \text{und} \quad T + S : V \rightarrow W : x \mapsto Tx + Sx$$

auch lineare Abbildungen. $\mathcal{L}(V, W)$ ist selbst wieder ein Vektorraum.

- Sind $T : V \rightarrow W, S : U \rightarrow V$ lineare Abbildungen, so ist

$$T \circ S : U \rightarrow W, \quad TS : x \mapsto T(S(x))$$

eine lineare Abbildung, das **Produkt** (oder die **Verkettung**) von T und S . Man schreibt auch $T \circ S, T \cdot S$ oder TS . Das Produkt ist **assoziativ**, d.h. $T(SR) = (TS)R$.

3.3. Injektive lineare Abbildungen.

Satz. Für $T : V \rightarrow W$ linear und äquivalent sind:

- (i) $\text{Kern}(T) = \{0\}$.
- (ii) T ist injektiv.

Beweis. (i) \Rightarrow (ii): Seien $u, v \in V$ mit $Tu = Tv$. Dann gilt $T(u - v) = 0$, d.h. $u - v \in \text{Kern}(T)$. Wegen (i) ist also $u = v$, und T ist injektiv.

(ii) \Rightarrow (i): Sei $v \in \text{Kern}(T)$, d.h. $Tv = 0 = T0$. Also wegen Injektivität, muß $v = 0$ gelten. Dies bedeutet $\text{Kern}(T) = \{0\}$. ■

3.4. Die Dimensionsformel.

Satz. Sei $T : V \rightarrow W$ linear und $\text{Dim}(V)$ endlich. Dann gilt

$$\text{Dim}(V) = \text{Dim}(\text{Bild}(T)) + \text{Dim}(\text{Kern}(T)).$$

Beweis. Ist T injektiv, so ist $T : V \rightarrow \text{Bild}(T)$ ein Vektorraum-Isomorphismus, also $\text{Dim}(V) = \text{Dim}(\text{Bild}(T))$. Somit ist in diesem Fall die Behauptung klar. Sei also T nicht injektiv, und sei $\{b_1, \dots, b_k\}$ eine Basis für $\text{Kern}(T) \neq \{0\}$. Ergänze zur Basis

$$\{b_1, \dots, b_k, b_{k+1}, \dots, b_n\}$$

für V (siehe 2.6). Dann gilt

$$\text{Bild}(T) = \text{Lin}\{T(b_1), T(b_2), \dots, T(b_n)\}.$$

Wir zeigen, dass $T(b_{k+1}), \dots, T(b_n)$ linear unabhängig sind, und so sind wir mit dem Beweis fertig. Sei also

$$\lambda_{k+1}T(b_{k+1}) + \dots + \lambda_n T(b_n) = 0,$$

und

$$v = \lambda_{k+1}b_{k+1} + \dots + \lambda_nb_n.$$

Dann $v \in \text{Kern}(T)$. Jedes $v \in \text{Kern}(T)$ ist aber auf eindeutiger Weise eine Linearkombination von $b_1, \dots, b_k, b_{k+1}, \dots, b_n$. Daher bekommen wir:

$$\lambda_{k+1} = \dots = \lambda_n = 0,$$

also lineare Unabhängigkeit folgt. ■

Folgerung: Im Beweis haben wir gesehen: T injektiv $\Rightarrow \text{Dim}(V) = \text{Dim}(\text{Bild}(T))$. Ist also $\{b_1, \dots, b_n\}$ Basis von V so ist $T(b_1), \dots, T(b_n)$ eine Basis von $\text{Bild}(T)$.

Bemerkung. $\text{Rang}(T) := \text{Dim}(\text{Bild}(T))$ heißt der **Rang** von T .

3.5. Lineare Abbildungen auf Basen.

Der nächste Satz besagt, dass wir auf linearen unabhängigen Vektoren die Werte einer linearen Abbildung beliebig vorschreiben können.

Satz. Sei $\{b_1 \dots b_n\}$ Basis von V und $\{w_1 \dots w_n\}$ Vektoren in W . Es existiert genau eine lineare Abbildung

$$T : V \rightarrow W \quad \text{mit} \quad T(b_1) = w_1, \dots, T(b_n) = w_n.$$

Beweis. $T(\sum_{i=1}^n \lambda_i b_i) := \sum_{i=1}^n \lambda_i w_i$ ist wohldefiniert, da Koeffizienten λ_i eindeutig bestimmt sind. Linearität und Eindeutigkeit sind klar. ■

Bemerkung. Spezialfall $V = K^n, b_i = e_i$ (vgl. 2.5).

Satz. Wie in Satz 2.5.1 gilt:

- a) $\text{Lin}\{w_1, \dots, w_n\} = W \Leftrightarrow T$ ist surjektiv
- b) $\{w_1, \dots, w_n\}$ linear unabhängig $\Leftrightarrow T$ ist injektiv
- c) $\{w_1, \dots, w_n\}$ Basis in $W \Leftrightarrow T$ ist bijektiv (Isomorphismus)

Beweis. Wie in 2.5: ersetze aber

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \quad \text{durch} \quad \sum_{i=1}^n \lambda_i b_i. \quad \text{■}$$

3.6. Die Matrix einer linearen Abbildung.

Sei $T : V \rightarrow W$ eine lineare Abbildung

$$\mathcal{L} := \{b_1, \dots, b_n\} \quad \text{Basis von } V$$

und $w_i = T(b_i) \in W$. Nach 3.5 ist T die einzige lineare Abbildung von V nach W mit $b_i \mapsto w_i$.

Sei

$$\mathcal{L}' := \{c_1, \dots, c_m\} \quad \text{Basis von } W,$$

dann ist

$$T(b_j) = w_j = \alpha_{1j}c_1 + \alpha_{2j}c_2 + \dots + \alpha_{mj}c_m$$

und w_j ist durch die Koordinaten $(\alpha_1, \dots, \alpha_m)$ eindeutig bestimmt.

Sind die Basen \mathcal{L} von V und \mathcal{L}' von W vorgegeben, so ist $T : V \rightarrow W$ durch die $n \times m$ Zahlen $\{\alpha_{ij} : 1 \leq i \leq m, 1 \leq j \leq n\}$ vollständig bestimmt

Konvention. Schreibe die α_{ij} in die Form:

$$M := M_T^{\mathcal{L}, \mathcal{L}'} := \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix} = (\alpha_{ij})_{\substack{i=1, \dots, m \\ j=1, \dots, n}}$$

In den Spalten stehen also die Koordinaten der Bilder der Basisvektoren. So eine quadratische Anordnung von $m \cdot n$ Zahlen in m -Zeilen und n -Spalten, heißt eine $m \times n$ -**Matrix**.

Definition. $M = M_T^{\alpha, \alpha'}$ heißt die $m \times n$ **Matrix** der linearen Abbildung T bzgl. der Basen \mathcal{L} von V und \mathcal{L}' von W . α_{ij} heißen die **Einträge** oder die Koeffizienten der Matrix M .

Bemerkung. 1. $n = \text{Dim}(V) = \text{Anzahl der Spalten}$, $m = \text{Dim}(W) = \text{Anzahl der Zeilen}$.

2. Bei festen Basen \mathcal{L} von V und \mathcal{L}' von W hat man also eine 1-1 Beziehung zwischen linearen Abbildungen von V nach W und den $m \times n$ Matrizen

$$T \leftrightarrow M_T^{\mathcal{L}, \mathcal{L}'}$$

Setze

$$M_{m,n} = M_{m,n}(K) := \{m \times n \text{ Matrizen mit Einträge in } K\}.$$

3. Seien $\mathcal{E} := \{e_1, \dots, e_n\}$ und $\mathcal{E}' = \{f_1, \dots, f_m\}$ die kanonischen Basen von K^n und K^m . Dann gehört insbesondere zu jeder $m \times n$ -Matrix M genau eine lineare Abbildung

$$T_M : K^n \rightarrow K^m \quad \text{mit} \quad M = M_{T_M}^{\mathcal{E}, \mathcal{E}'}$$

In den Spalten von $M_{T_M}^{\mathcal{E}, \mathcal{E}'}$ stehen die Vektoren $T_M(e_j)$, $j = 1, \dots, n$

4. Sei $M = M_T^{\mathcal{E}, \mathcal{E}'}$ mit \mathcal{L} und \mathcal{L}' Basen von V bzw. W ; $\text{Dim}(V) = n$ und $\text{Dim}(W) = m$. So ist dieselbe Matrix M auch die Matrix (bzgl. \mathcal{L} und \mathcal{L}') genau einer linearen Abbildung

$$R : V \rightarrow W, \quad M = M_R^{\mathcal{L}, \mathcal{L}'}$$

Seien

$$K_n : K^n \rightarrow V, \quad e_i \mapsto b_i, \quad K_m : K^m \rightarrow W \quad f_j \mapsto c_j$$

die Koordinatenabbildungen (vgl. 2.5). Dann sind K_n , und K_m Vektorraum-Isomorphismen, und das folgende Diagramm ist kommutativ:

$$\begin{array}{ccc} K^n & \xrightarrow{T} & K^m \\ K_n \downarrow & & \downarrow K_m \\ V & \xrightarrow{R} & W \end{array}$$

Das heißt

$$R \circ K_n = K_m \circ T \quad (\text{oder } R = K_m \circ T \circ K_n^{-1}).$$

Das werden wir bald besser verstehen.

Wir merken uns schon jetzt: R und T unterscheiden sich nur durch Vektorraum-Isomorphismen. Daraus folgt: $\text{Dim}(\text{Kern})$ und $\text{Rang} = \text{Dim}(\text{Bild})$ sind für R und T gleich. Daher genügt es oft, eine Matrix als die Matrix einer linearen Abbildung $T : K^n \rightarrow K^m$ zu interpretieren. Wer will, kann sich das im folgenden so klar machen, auch wenn es allgemeiner formuliert ist.

3.7. Beispiel für das Aufstellen von Matrizen.

1. Die Nullabbildung $O : V \rightarrow W$ hat in jeder Basis die Matrix

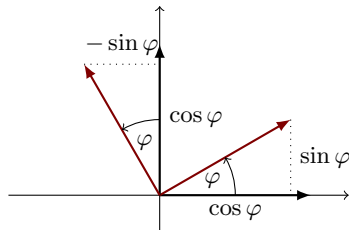
$$O_{m \times n} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}_{m \times n}$$

2. Die identische Abbildung $I : V \rightarrow V$ bzgl. einer Basis \mathcal{L} hat die Matrix

$$I_{n \times n} = M_I^{\mathcal{L}, \mathcal{L}} = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}_{n \times n}.$$

A C H T U N G: $M = M_I^{\mathcal{L}, \mathcal{L}'}$ sieht anders aus, falls $\mathcal{L} \neq \mathcal{L}'$, und diese Tatsache wird später ganz wichtig sein. (Siehe Basiswechsel 3.16).

3. Sei $D : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ Drehung um Winkel φ , und sei $\mathcal{E} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ die kanonische Basis.



Die Matrix von D bestimmt man so:

$$D \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}, \quad D \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix}.$$

Also:

$$M_D^{\mathcal{E}, \mathcal{E}} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}.$$

4. Orthogonale Projektion P von \mathbb{R}^3 auf x - y -Ebene:

$$P \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}, \quad \text{also} \quad P \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad P \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad P \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Daher ist

$$M_P^{\mathcal{E}, \mathcal{E}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Frage:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \text{oder} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{gehört zu??}$$

5. Die $m \times 1$ Matrix

$$M = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}$$

ist die Matrix einer linearen Abbildung

$$T : K \rightarrow K^m \quad \text{bezüglich der kanonischen Basen, nämlich:} \quad T(\lambda) = \lambda \begin{pmatrix} \lambda \alpha_1 \\ \vdots \\ \lambda \alpha_m \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}.$$

T ist injektiv und somit ein Isomorphismus auf $\text{Bild}(T)$. Vektoren in K^n schreiben wir also als Spaltenvektoren.

6. Die $1 \times n$ Matrix

$$M := (\alpha_1, \dots, \alpha_n)$$

ist die Matrix einer linearen Abbildung

$$S : K^n \rightarrow K \quad \text{bezüglich der kanonischen Basen, nämlich:} \quad S(e_i) = \alpha_i,$$

also

$$S \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n.$$

Bemerke, dass S eine Linearform ist, d.h. $S \in (K^n)^*$. Vektoren in $(K^n)^*$ schreiben wir also als Zeilenvektoren. So ist der Wert einer Linearform auf einem Vektor sehr bequem auszurechnen

$$(\alpha_1 \dots \alpha_m) \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} = \alpha_1 \beta_1 + \alpha_2 \beta_2 + \dots + \alpha_m \beta_m \quad \text{“Zeile} \times \text{Spalte”}.$$

3.8. Berechnung des Bildes eines allgemeinen Vektors. Sei $T : V \rightarrow W$ linear, seien $\mathcal{L} = \{b_1, \dots, b_n\}$ und $\mathcal{L}' = \{c_1, \dots, c_m\}$ Basen von V bzw. W . Setze

$$M := M_T^{\mathcal{L}, \mathcal{L}'} = (\alpha_{ij})_{i,j}.$$

Für $v \in V$ schreibe

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

Wir benutzen die Notationen

$$v_{\mathcal{L}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \quad \text{und} \quad v = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}_{\mathcal{L}},$$

und analog für Vektoren $w \in W$ bzgl. der Basis \mathcal{L}' . Dann ist

$$T(v) = T\left(\sum_{j=1}^n \lambda_j b_j\right) = \sum_{j=1}^n \lambda_j T(b_j) = \sum_{j=1}^n \lambda_j \sum_{i=1}^m \alpha_{ij} c_i = \sum_{i=1}^m \left(\sum_{j=1}^n \alpha_{ij} \lambda_j\right) c_i$$

Wir sehen also:

$$T : V \rightarrow W,$$

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}_{\mathcal{L}} \mapsto \begin{pmatrix} \sum_{j=1}^n \alpha_{1j} \lambda_j \\ \vdots \\ \sum_{j=1}^n \alpha_{mj} \lambda_j \end{pmatrix}_{\mathcal{L}'} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}_{\mathcal{L}}$$

d.h. wir rechnen:

1. Koordinate des Bildvektors $T(v)$ ist gleich 1. Zeile mal Koordinatenvektor $v_{\mathcal{L}}$
2. Koordinate des Bildvektors $T(v)$ ist gleich 2. Zeile mal Koordinatenvektor $v_{\mathcal{L}}$
- \vdots
- m . Koordinate des Bildvektors $T(v)$ ist gleich m . Zeile mal Koordinatenvektor $v_{\mathcal{L}}$.

3.9. Beispiele.

1. Sei

$$M_T^{\mathcal{L}, \mathcal{L}'} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad \text{und} \quad v = \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix}_{\mathcal{L}}.$$

Dann hat $T(v)$ in Basis \mathcal{L}' den Koordinatenvektor

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix} = \begin{pmatrix} 1 \cdot 7 + 2 \cdot 8 + 3 \cdot 9 \\ 4 \cdot 7 + 5 \cdot 8 + 6 \cdot 9 \end{pmatrix} = \begin{pmatrix} 50 \\ 122 \end{pmatrix}.$$

2. Betrachte die Drehung D auf \mathbb{R}^2 . Bzgl. kanonischer Basis ist:

$$D \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} \xi \cos \varphi - \eta \sin \varphi \\ \xi \sin \varphi + \eta \cos \varphi \end{pmatrix}.$$

3. Das lineare Gleichungssystem

$$\begin{aligned} \alpha_{11}x_1 + \alpha_{12}x_2 + \dots + \alpha_{1n}x_n &= y_1 \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \dots + \alpha_{2n}x_n &= y_2 \\ &\vdots \\ \alpha_{m1}x_1 + \alpha_{m2}x_2 + \dots + \alpha_{mn}x_n &= y_m \end{aligned}$$

kann auch geschrieben werden als:

$$(*) \quad \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

Also ist $A = (\alpha_{ij})_{i,j}$ die Matrix von $T : K^n \rightarrow K^m$ bzgl. kanonischer Basen, $x \in K^n, y \in K^m$, dann heißt $(*)$: Für die lineare Abbildung T und für $y \in K^m$ suche $x \in K^n$ mit

$$Tx = y,$$

also ein Urbild von y unter T . Dadurch wird die Theorie von linearen Gleichungssystemen sehr einfach.

3.10. Vielfaches und Summen von Matrizen. Gegeben seien V mit Basis \mathcal{L}, W mit Basis \mathcal{L}' und $S, T : V \rightarrow W$ lineare Abbildungen mit Matrizen

$$M_S^{\mathcal{L},\mathcal{L}'} = (\alpha_{ij})_{i,j}, \quad M_T^{\mathcal{L},\mathcal{L}'} = (\beta_{ij})_{i,j}.$$

Satz. a) Für $\lambda \in K$ ist die Matrix von λS gegeben durch:

$$M_{\lambda S}^{\mathcal{L},\mathcal{L}'} = (\lambda\alpha_{ij})_{i,j} = \begin{pmatrix} \lambda\alpha_{11} & \lambda\alpha_{12} & \dots & \lambda\alpha_{1n} \\ \lambda\alpha_{21} & \lambda\alpha_{22} & \dots & \lambda\alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda\alpha_{m1} & \lambda\alpha_{m2} & \dots & \lambda\alpha_{mn} \end{pmatrix}$$

b) Die Matrix von $S + T$ ist gegeben durch:

$$\begin{aligned} M_{S+T}^{\mathcal{L},\mathcal{L}'} = (\alpha_{ij} + \beta_{ij})_{i,j} &= \begin{pmatrix} \alpha_{11} + \beta_{11} & \alpha_{12} + \beta_{12} & \dots & \alpha_{1n} + \beta_{1n} \\ \alpha_{21} + \beta_{21} & \alpha_{22} + \beta_{22} & \dots & \alpha_{2n} + \beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} + \beta_{m1} & \alpha_{m2} + \beta_{m2} & \dots & \alpha_{mn} + \beta_{mn} \end{pmatrix} \\ &= \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{m1} & \beta_{m2} & \dots & \beta_{mn} \end{pmatrix} = M_S^{\mathcal{L},\mathcal{L}'} + M_T^{\mathcal{L},\mathcal{L}'}. \end{aligned}$$

Beweis. Durch Hinschauen: An der Stelle (i, j) steht die i -te Koordinate des Bildes des j -ten Basisvektors. ■

3.11. $M_{m,n}$ als Vektorraum. Der Satz im 3.10 liefert, dass $M_{m,n}(K)$ ein Vektorraum über dem Körper K ist. Wir untersuchen kurz diesen Vektorraum weiter.

Satz. a) Die $m \times n$ -Elementarmatrizen

$$E_{ij} := \begin{pmatrix} \vdots & & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots \\ 0 & \dots & 0 & 1 & 0 & \dots \\ 0 & \dots & 0 & 0 & 0 & \dots \\ \vdots & & \vdots & \vdots & \vdots \end{pmatrix} \leftarrow i\text{-te Zeile}$$

\uparrow
 j -te Spalte

bilden eine Basis in $M_{m,n}(K)$. Daher ist die Dimension

$$\dim(M_{m,n}(K)) = mn.$$

b) Seien V und W Vektorräume über K mit $\dim(V) = n$ und $\dim(W) = m$. Seien $\mathcal{L} \subseteq V$ und $\mathcal{L}' \subseteq W$ Basen. Die Abbildung

$$J: \mathcal{L}(V, W) \rightarrow M_{m,n}(K), \quad J(T) := M_T^{\mathcal{L}, \mathcal{L}'}$$

ist ein Vektorraum-Isomorphismus. Also $\dim(\mathcal{L}(V, W)) = mn$.

Beweis. a) ist klar. b) folgt aus 3.10 und 3.6, Bemerkung 2. ■

Bemerkung. Spezialfall: Sei $W = K$. Dann $\mathcal{L}(V, W) = V^*$ ist ein n -dimensionaler Vektorraum, der **Dualraum** der **Linearformen** auf V (vgl. 1.18).

Ist $\mathcal{L} = \{b_1, \dots, b_n\}$ Basis von V , so ist die **duale Basis**

$$\mathcal{L}^* := \{b_1^*, \dots, b_n^*\}$$

von V^* durch die Matrizen

$$(1, 0, \dots, 0), (0, 1, \dots, 0) \quad \dots \quad (0, 0, \dots, 1) \quad \text{gegeben,}$$

also $b_i^*(b_j) = \delta_{ij}$ (siehe auch 4.1).

3.12. Matrixmultiplikation. Seien U, V, W Vektorräume über K mit Basen $\mathcal{L} = \{a_1, \dots, a_n\} \subseteq U$, $\mathcal{L}' = \{b_1, \dots, b_m\} \subseteq V$ und $\mathcal{L}'' = \{c_1, \dots, c_l\} \subseteq W$. Seien ferner $T: U \rightarrow V$ und $S: V \rightarrow W$ lineare Abbildungen mit Matrizen

$$B = M_T^{\mathcal{L}, \mathcal{L}'} = (\beta_{ij}) \quad \text{und} \quad A = M_S^{\mathcal{L}', \mathcal{L}''} = (\alpha_{ij}).$$

Dann ist $ST: U \rightarrow W$ auch linear. Wir bestimmen nun die Matrix $M_{ST}^{\mathcal{L}, \mathcal{L}''}$. Nach Definition gilt

$$(ST)(a_k) = S\left(\sum_{i=1}^m \beta_{ik} b_i\right) = \sum_{i=1}^m \beta_{ik} S(b_i) = \sum_{i=1}^m \beta_{ik} \sum_{j=1}^l \alpha_{ji} c_j = \sum_{j=1}^l \left(\sum_{i=1}^m \alpha_{ji} \beta_{ik}\right) c_j.$$

Das heißt, für die Matrix $C = M_{ST}^{\mathcal{L}, \mathcal{L}''}$ gilt $C = (\gamma_{jk}) \in M_{l,n}(K)$ mit

$$\gamma_{jk} = \sum_{i=1}^m \alpha_{ji} \beta_{ik}.$$

Definition. Gegeben seien Matrizen $A \in M_{l,m}(K)$, $A = (\alpha_{ij})$, $B \in M_{m,n}(K)$, $B = (\beta_{ij})$. Dann ist das **Produkt** AB von A und B die $l \times n$ -Matrix $C = (\gamma_{jk})$ mit

$$\gamma_{jk} = \sum_{i=1}^m \alpha_{ji} \beta_{ik}.$$

Wir haben also die Definition so gestaltet, dass der folgende Satz gültig wird.

Satz. Es gilt

$$M_{ST}^{\mathcal{L}, \mathcal{L}''} = M_S^{\mathcal{L}', \mathcal{L}''} M_T^{\mathcal{L}, \mathcal{L}'}$$

A C H T U N G: AB kann nur gebildet werden, falls Spaltenanzahl von A =Zeilenanzahl von B gilt. Das Element in der j -ten Zeile und der k -ten Spalte der Produktmatrix AB ist das Produkt der j -ten Zeile von A und der k -ten Spalte von B .

Beispiel. Hier sind Produkte einiger Matrizen:

$$1. \quad \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

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

$$3. \quad \begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & 1 & 2 & -2 \\ 2 & -1 & 0 & -3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

$$4. \quad (1 \ 2 \ -1 \ 1) \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} = (1) = 1$$

$$5. \quad \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} (1 \ 2 \ -1 \ 1) = \begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ -1 & -2 & 1 & -1 \end{pmatrix}$$

3.13. Rechenregeln Falls die Produkte und Summen gebildet werden können gilt:

1. Für $A \in M_{m,n}$ gilt

$$I_{m \times m} A = A = A I_{n \times n}, \quad O_{k \times m} A = O_{k \times n}, \quad A O_{n \times k} = O_{m \times k}.$$

2. Für $A \in M_{m,n}$, $B \in M_{n,k}$ und $C \in M_{k,l}$ gilt

$$A(BC) = (AB)C.$$

3. Für $A, B \in M_{m,n}$, $C \in M_{n,k}$ gilt

$$(A + B)C = AC + BC.$$

Für $C \in M_{k,m}$ gilt

$$C(A + B) = CA + CB.$$

Beweis. Dies analoge Aussagen gelten für lineare Abbildungen. Mithilfe vom Satz 3.11.b) kann man diese Eigenschaften auf Matrizen übertragen. ■

A C H T U N G I: Es gilt in Allgemeinem nicht $AB = BA$ (selbst für $n \times n$ -Matrizen).

Beispiel. Betrachte

$$S_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \text{und} \quad S_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Dann gilt

$$S_3 S_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad S_1 S_3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

A C H T U N G II: Aus $AB = 0$ folgt nicht $A = 0$ oder $B = 0$.

Beispiel. Betrachte

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

Dann gilt

$$A^2 = AA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0.$$

Da eine Potenz von A gleich die Nullmatrix 0) ist, heißt A **nilpotent**.

3.14. Zusammenfassung. Ist U ein Vektorraum mit der Basis $\mathcal{L}_1 = \{b_1, b_2, \dots, b_n\}$ und $u \in U$ mit $u = \sum_{i=1}^n \lambda_i b_i$, dann sei

$$u_{\mathcal{L}_1} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in K^n$$

der **Koordinatenvektor** von u bzgl. der Basis \mathcal{L}_1 . Ebenso schreiben wir

$$u = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}_{\mathcal{L}_1} \quad \text{für den Vektor } u \in U \text{ mit Koordinatenvektor } \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \text{ bzgl. } \mathcal{L}_1.$$

Seien V und W Vektorräume mit Basen $\mathcal{L}_2 = \{c_1, \dots, c_m\}$ bzw. $\mathcal{L}_3 = \{d_1, \dots, d_k\}$. Seien ferner

$$T : U \rightarrow V \quad \text{und} \quad S : V \rightarrow W \quad \text{lineare Abbildungen.}$$

Dann

$$M_T^{\mathcal{L}_1, \mathcal{L}_2} = (T(b_1)_{\mathcal{L}_2}, \dots, T(b_n)_{\mathcal{L}_2})$$

↑
Spaltenvektor

die Matrix der Abbildung T bzgl. der Basen $\mathcal{L}_1, \mathcal{L}_2$ ($T(b_j)$ als Spaltenvektoren). Für den Bild von u gilt:

$$(Tu)_{\mathcal{L}_2} = M_T^{\mathcal{L}_1, \mathcal{L}_2} u_{\mathcal{L}_1}$$

und für die Verkettung ST

$$M_{ST}^{\mathcal{L}_1, \mathcal{L}_3} = M_S^{\mathcal{L}_2, \mathcal{L}_3} M_T^{\mathcal{L}_1, \mathcal{L}_2}.$$

3.15. Inverse von Matrizen. Sei $T : U \rightarrow V$ eine lineare Abbildung mit der Matrix $M := M_T^{\mathcal{L}_1, \mathcal{L}_2}$.

Definition. Die Abbildung T heißt **invertierbar** (oder **bijektiv**), falls eine lineare Abbildung $T^{-1} : V \rightarrow U$ mit $TT^{-1} = I_V$ und $T^{-1}T = I_U$ existiert. (Hier bezeichnet I_U bzw. I_V die identische Abbildung auf U und V .) (Übrigens ist "invertierbar" eine alternative Terminologie für "Isomorphismus".)

Satz. T ist genau dann invertierbar, wenn M invertierbar ist, d.h. wenn eine Matrix M^{-1} mit $M^{-1}M = I_{m \times m}$ und $MM^{-1} = I_{n \times n}$ existiert. Darüber hinaus gilt:

$$M^{-1} = M_{T^{-1}}^{\mathcal{L}_2, \mathcal{L}_1}.$$

Beweis. Ist T invertierbar, dann setze $M^{-1} := M_{T^{-1}}^{\mathcal{L}_2, \mathcal{L}_1}$. Nach dem Obigen gilt:

$$I_{m \times m} = M_I^{\mathcal{L}_1, \mathcal{L}_2} = M_{TT^{-1}}^{\mathcal{L}_1, \mathcal{L}_2} = M_T^{\mathcal{L}_1, \mathcal{L}_2} M_{T^{-1}}^{\mathcal{L}_2, \mathcal{L}_1} = MM^{-1}.$$

Analog zeigt man $M^{-1}M = I_{n \times n}$.

Umgekehrt nehme an, das M^{-1} mit $M^{-1}M = I_{m \times m}$ und $MM^{-1} = I_{n \times n}$ existiert. Sei T^{-1} die lineare Abbildung mit Matrix M^{-1} . Das obige Rechnen ergibt dann auch $TT^{-1} = I_V$ und $T^{-1}T = I_U$. ■

Bemerkung. Ist $T : U \rightarrow V$ bijektiv, dann gilt $\dim U = \dim(V)$, also M ist quadratisch.

3.16. Basiswechsel.

Definition. Sei V Vektorraum mit den Basen \mathcal{L} und $\mathcal{L}' = \{b_1, \dots, b_n\}$ und $\{b'_1, \dots, b'_n\}$, dann heißt $S = M_I^{\mathcal{L}, \mathcal{L}'}$ die **Transformationsmatrix** für den Basiswechsel von \mathcal{L} auf \mathcal{L}' .

Eigenschaften.

1. S ist invertierbar, und $S^{-1} = M_I^{\mathcal{L}', \mathcal{L}}$ (siehe 3.15).
2. Sei $v \in V$ mit

$$v = \sum_{i=1}^n \lambda_i b_i \quad \text{also} \quad v_{\mathcal{L}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

und

$$v = \sum_{i=1}^n \lambda'_i b'_i \quad \text{also} \quad v_{\mathcal{L}'} = \begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_n \end{pmatrix}.$$

Dann gilt

$$v_{\mathcal{L}'} = S v_{\mathcal{L}} = M_I^{\mathcal{L}, \mathcal{L}'} v_{\mathcal{L}} \quad \text{und} \quad v_{\mathcal{L}} = S^{-1} v_{\mathcal{L}'}.$$

3. Ist $T : V \rightarrow V$ linear, dann gilt

$$M_T^{\mathcal{L}', \mathcal{L}'} = M_{IT'I}^{\mathcal{L}', \mathcal{L}'} = M_I^{\mathcal{L}, \mathcal{L}'} M_T^{\mathcal{L}, \mathcal{L}} M_I^{\mathcal{L}', \mathcal{L}} = S M_T^{\mathcal{L}, \mathcal{L}} S^{-1}$$

4. **Allgemeiner:** Sei $T : U \rightarrow V$, \mathcal{M} und \mathcal{M}' Basen von U . Setze

$$M = M_T^{\mathcal{M}, \mathcal{L}}, \quad S = M_{IV}^{\mathcal{L}, \mathcal{L}}, \quad R = M_{IU}^{\mathcal{M}, \mathcal{M}'}$$

Dann gelten die folgenden Aussagen:

$$\begin{aligned} M_T^{\mathcal{M}, \mathcal{L}'} &= M_{IV}^{\mathcal{L}, \mathcal{L}'} M_T^{\mathcal{M}, \mathcal{L}} = SM \\ M_T^{\mathcal{M}', \mathcal{L}} &= M_T^{\mathcal{L}, \mathcal{L}} M_{IU}^{\mathcal{M}', \mathcal{M}} = MR^{-1} \\ M_T^{\mathcal{M}', \mathcal{L}'} &= M_{IV}^{\mathcal{L}, \mathcal{L}'} M_T^{\mathcal{M}, \mathcal{L}} M_{IU}^{\mathcal{M}', \mathcal{M}} = SMR^{-1}. \end{aligned}$$

3.17. Äquivalenz und Ähnlichkeit von Matrizen.

Behauptung. Sei $S \in M_{n,n}$ eine invertierbare Matrix und $\mathcal{L} = \{b_1, \dots, b_n\}$ Basis von V . Dann existiert eine Basis \mathcal{L}' von V , so dass S die Transformationsmatrix $S = M_{IV}^{\mathcal{L}, \mathcal{L}'}$ ist.

Beweis. In den Spalten der invertierbaren Matrix S^{-1} stehen n linear unabhängige Koordinatenvektoren, die wir bzgl. \mathcal{L} als die Koordinaten von Vektoren $\{b'_1, \dots, b'_n\} = \mathcal{L}'$ auffassen. So ist \mathcal{L}' eine Basis (linear unabhängig und $\text{Dim}(V) = n$) und es gilt

$$S^{-1} = M_{IV}^{\mathcal{L}', \mathcal{L}}, \quad S = M_{IV}^{\mathcal{L}, \mathcal{L}'}. \quad \blacksquare$$

Jede invertierbare Matrix lässt sich als Transformationsmatrix interpretieren. Multiplikation von rechts bzw. links ergibt eine Koordinaten-Transformation im Urbildraum bzw. Bildraum.

Definition. a) Zwei Matrizen $A, B \in M_{m,n}$ heißen **äquivalent**, wenn es eine invertierbare $m \times m$ -Matrix R und eine invertierbare $n \times n$ -Matrix S gibt mit

$$B = RAS^{-1}.$$

A und B sind also genau dann äquivalent, wenn sie bzgl. geeigneter Basen dieselbe lineare Abbildung beschreiben (siehe 3.16.4).

b) Zwei quadratische Matrizen $A, B \in M_{n,n}$ heißen **ähnlich**, wenn es eine invertierbare Matrix $S \in M_{n,n}$ gibt mit

$$B = SAS^{-1},$$

d.h. $A = M_T^{\mathcal{L}, \mathcal{L}}$ und $B = M_T^{\mathcal{L}', \mathcal{L}'}$ für geeignete $\mathcal{L}, \mathcal{L}'$ für eine lineare Abbildung T .

Beispiel. Sei $V = \mathbb{R}^2$ und $\mathcal{L} = \{b_1, b_2\}$ die Basis gegeben durch

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

und $\mathcal{L}' = \{b'_1, b'_2\}$ die Basis gegeben durch

$$b'_1 := \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad b'_2 := \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Es gelten:

$$b_1 = \frac{1}{2}b'_1 + \frac{1}{2}b'_2, \quad b_2 = \frac{1}{2}b'_1 - \frac{1}{2}b'_2,$$

also

$$S = M_I^{\mathcal{L}, \mathcal{L}'} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

Weiterhin kann man S^{-1} auch bestimmen:

$$b'_1 = 1 \cdot b_1 + 1 \cdot b_2, \quad b'_2 = 1 \cdot b_1 - 1 \cdot b_2,$$

und somit

$$S^{-1} = M_I^{\mathcal{L}', \mathcal{L}} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Probe:

$$S^{-1}S = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Sei nun T die lineare Abbildung mit Matrix

$$M_T^{\mathcal{L}, \mathcal{L}} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Dann ist

$$M_T^{\mathcal{L}', \mathcal{L}'} = SM_T^{\mathcal{L}, \mathcal{L}}S^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

eine **Diagonalmatrix**! Bezüglich \mathcal{L}' ist T besonders einfach: Streckung der Basisvektoren.

Eines der Hauptprobleme der Linearen Algebra ist für ein gegebenes T eine Basis zu finden, so dass $M_T^{\mathcal{L}, \mathcal{L}}$ besonders einfach wird, wenn möglich diagonal.

3.18. Transformation von Orthonormalbasen.

Definition. a) Sei

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix} \in M_{m,n}$$

eine $m \times n$ -Matrix. Dann heißt die $n \times m$ -Matrix definiert durch

$$A^\top := \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix} \in M_{n,m}$$

die **transponierte** Matrix von A .

b) Sei $A \in M_{m,n}(\mathbb{K})$ (also die Einträge der Matrix sind reell oder komplex). Setze

$$A^* := \begin{pmatrix} \overline{\alpha_{11}} & \overline{\alpha_{12}} & \dots & \overline{\alpha_{1n}} \\ \overline{\alpha_{21}} & \overline{\alpha_{22}} & \dots & \overline{\alpha_{2n}} \\ \vdots & & & \vdots \\ \overline{\alpha_{m1}} & \overline{\alpha_{m2}} & \dots & \overline{\alpha_{mn}} \end{pmatrix} = \overline{A}^\top.$$

So heißt A^* die **adjungierte** der Matrix A .

Bemerkung: Ist $\mathbb{K} = \mathbb{R}$ so gilt $A^\top = A^*$.

c) A heißt **symmetrisch**, falls gilt

$$A = A^\top.$$

A heißt **selbstadjungiert**, falls gilt

$$A = A^*.$$

A heißt **orthogonal** ($\mathbb{K} = \mathbb{R}$), falls gilt

$$A^{-1} = A^\top.$$

A heißt **unitär** ($\mathbb{K} = \mathbb{C}$), falls gilt

$$A^{-1} = A^*.$$

Sei $\mathcal{L} = \{b_1, \dots, b_n\}$ Orthonormalbasis in V und sei $T: V \rightarrow V$, so dass $b'_i := T(b_i)$, $\mathcal{L}' = \{b'_1, \dots, b'_n\}$ wieder eine Orthonormalbasis in V ist. Dann ist $M = M_T^{\mathcal{L}, \mathcal{L}} = M_T^{\mathcal{L}', \mathcal{L}'}$ quadratisch. In den Spalten von

M steht die Orthonormalbasis $(b'_1)_{\mathcal{L}}, \dots, (b'_n)_{\mathcal{L}}$ und in den Zeilen von M^* stehen die (Zeilen)-Vektoren $(\overline{b'_1})_{\mathcal{L}}^\top, \dots, (\overline{b'_n})_{\mathcal{L}}^\top$. Daraus folgt

$$M^*M = \begin{pmatrix} 1 & 0 & \dots & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix} = I_{n \times n},$$

d.h. M ist unitär.

Ist also S die Transformationsmatrix zwischen zwei Orthonormalbasen, dann ist S unitär (oder, im Falle $K = \mathbb{R}$, orthogonal). Damit ist S^{-1} sehr leicht auszurechnen: $S^{-1} = S^*$.

4. RECHNEN MIT MATRIZEN

In diesem Kapitel werden wir zunächst die so genannten **elementaren Umformungen** studieren, die es ermöglichen eine Matrix auf besonders einfache Gestalt zu bringen: auf **Zeilenstufenform** (Hausaufgabe: das Gaußsche Eliminationsverfahren wiederholen). Die elementaren Umformungen dienen auch dazu, inhomogene und homogene lineare Gleichungssysteme einfach behandeln zu können. Am Ende des Kapitels führen wir die **Determinante** einer Matrix ein.

Als Aufwärmung betrachten wir eine weitere Operation die man mit Matrizen bzw. mit linearen Abbildungen ausführen kann.

4.1. Das Dual einer linearen Abbildung. Gegeben seien die endlichdimensionalen Vektorräume V, W über dem Körper K . Wir betrachten die Dualräume von V und W , d.h., die Räume der Linearformen auf V und W . Die üblichen Notationen dafür sind V^* und W^* . Ist $T : V \rightarrow W$ eine lineare Abbildung, so können wir jede Linearform $\ell \in W^*$ mit T verketteten, so erhalten wir eine neue Linearform jetzt aber auf V :

$$V \xrightarrow{T} W \xrightarrow{\ell} K, \quad \ell \circ T(v) = \ell(Tv).$$

Da T und ℓ beide linear sind, so ist tatsächlich auch ihre Verkettung, d.h. $\ell \circ T \in V^*$, und somit können wir die folgende Definition betrachten.

Definition. Sei $T \in \mathcal{L}(V, W)$. Die Abbildung, die zu jeder Linearform $\ell \in W^*$ die Linearform $\ell \circ T \in V^*$ zuordnet, heißt die zu T **duale Abbildung**. Also $T'\ell = \ell \circ T$.

Man rechnet nach, dass T' wieder eine lineare Abbildung ist und W^* nach V^* abbildet.

Nun wollen wir es untersuchen, wie die Matrizen von T und T' zusammenhängen. Dazu brauchen wir Basen in V^* und W^* . Die interessante Wahl ist jene von dualen Basen. Seien

$$\mathcal{L} = \{b_1, b_2, \dots, b_n\} \subseteq V \quad \text{und} \quad \mathcal{M} = \{c_1, c_2, \dots, c_m\} \subseteq W$$

Basen in V bzw. in W . Betrachte die zugehörigen **dualen Basen**

$$\mathcal{L}^* = \{b_1^*, b_2^*, \dots, b_n^*\} \subseteq V^* \quad \text{und} \quad \mathcal{M}^* = \{c_1^*, c_2^*, \dots, c_m^*\} \subseteq W^*$$

in V^* und W^* . D.h. es gelten $b_j^*(b_k) = \delta_{jk}$ und $c_i^*(c_p) = \delta_{ip}$ für alle $j, k = 1, \dots, n$ und $i, p = 1, \dots, m$ (δ bezeichnet das Kronecker-Delta). Ist $A \in M_{m,n}(K)$, $A = (\alpha_{ij})$ die Matrix der Abbildung T bzgl. der Basen \mathcal{L} und \mathcal{M} , so können wir die Matrix $B = (\beta_{pk}) \in M_{m,n}(K)$ von T' bzgl. \mathcal{L}^* und \mathcal{M}^* folgenderweise bestimmen: Die k -te Koordinate von $v' \in V^*$ bzgl. der Basis \mathcal{L}^* ist $v'(b_k)$. Daraus bekommen wir

$$\beta_{kp} = (T'c_p^*)(b_k) = c_p^*(Tb_k) = c_p^*\left(\sum_{i=1}^m \alpha_{ik}c_i\right) = \sum_{i=1}^m \alpha_{ik}c_p^*(c_i) = \alpha_{pk}.$$

Dies bedeutet aber, dass B die Transponierte von A ist, also

$$B = A^\top = \begin{pmatrix} \alpha_{11} & \alpha_{21} & \dots & \alpha_{m1} \\ \alpha_{12} & \alpha_{22} & \dots & \alpha_{m2} \\ \vdots & \vdots & & \\ \alpha_{1m} & \alpha_{2m} & \dots & \alpha_{mn} \end{pmatrix}$$

Das heißt, die Zeilenvektoren von B sind genau die Spaltenvektoren von A . Diese einfache Tatsache werden wir später oft ausnutzen.

4.2. Elementare Umformungen. Gegeben sei eine Matrix

$$A = (\alpha_{ij})_{\substack{i=1, \dots, m \\ j=1, \dots, n}} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & & \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix} \in M_{m,n}(K).$$

Die Zeile bzw. die Spalten möchten wir als Vektoren in K^n bzw. in K^m auffassen (Zeilen- bzw. Spaltenvektoren).

Definition. Der **Zeilenrang** von A ist die Dimension des Vektorraums aufgespannt durch die Zeilenvektoren:

$$\text{Z-Rang}(A) := \text{Dim Lin}\{\text{Zeilenvektoren von } A\} = \text{Dim Bild}(A).$$

Analog definiert man den **Spaltenrang**:

$$\text{S-Rang} := \text{Dim Lin}\{\text{Spaltenvektoren von } A\}.$$

Beispiel.

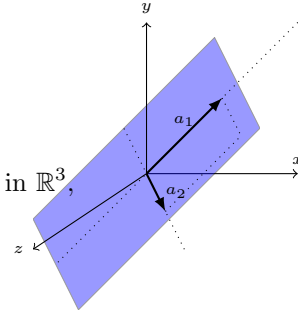
Für die Matrix

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

gilt:

$\text{Lin}\{\text{Zeilenvektoren von } A\} = \{x, y, z\} \in \mathbb{R}^3 : y + z = x\} = \text{Ebene in } \mathbb{R}^3,$
und $\text{Lin}\{\text{Spalten von } A\} = \mathbb{R}^2.$

Daher ist $\text{Z-Rang}(A) = 2$ und $\text{S-Rang}(A) = 2$.



Wir werden zeigen, dass für jede Matrix die Gleichheit gilt. Dazu brauchen wir aber einige Vorbereitungen.

Definition. Eine **elementare Zeilenumformung** (abgekürzt EZU) bzw. eine **elementare Spaltenumformung** (ESU) vom

Typ I (Typ I') ist Multiplikation einer Zeile (Spalte) mit einem Skalar $0 \neq \lambda \in K$;

Typ II (Typ II') ist die Addition des λ -fachen ($\lambda \in K$) einer Zeile (Spalte) zu einer anderen Zeile (Spalte).

Typ III (Typ III') ist das Vertauschen zweier Zeilen (Spalten).

Entsteht eine Matrix B aus A durch endliche Anzahl von elementaren Zeilen- (Spalten-, oder beiden) Umformungen, so schreiben wir

$$A \underset{\text{EZU}}{\sim} B \quad (A \underset{\text{ESU}}{\sim} B, A \underset{\text{EU}}{\sim} B).$$

Man vergleiche diese Operationen mit dem Gauß'schen Eliminationsverfahren. Die folgenden Aussagen lassen sich sehr leicht zu beweisen und werden dem Leser als Übungsaufgaben überlassen.

Behauptung. a) Es gilt $A \underset{\text{EZU}}{\sim} A$.

b) Ist $A \underset{\text{EZU}}{\sim} B$, so gilt auch $B \underset{\text{EZU}}{\sim} A$.

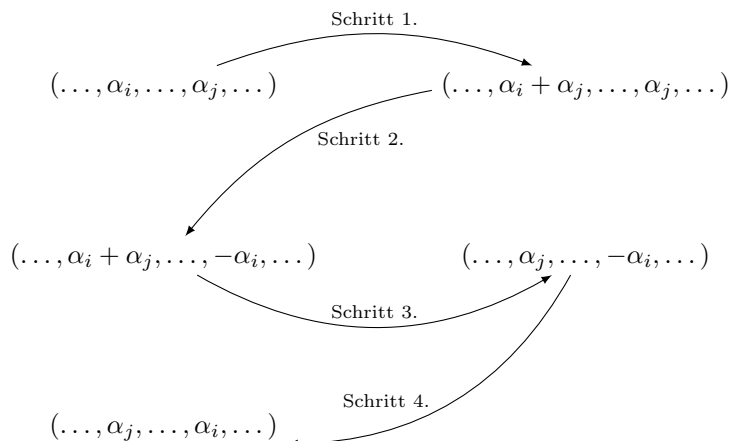
c) Sind $A \underset{\text{EZU}}{\sim} B$ und $B \underset{\text{EZU}}{\sim} C$, so auch $A \underset{\text{EZU}}{\sim} C$.

Man sagt, dass $\underset{\text{EZU}}{\sim}$ eine Äquivalenzrelation auf der Menge $M_{m,n}(K)$ ist. Diese Eigenschaften stimmen natürlich auch für $\underset{\text{ESU}}{\sim}$ und $\underset{\text{EU}}{\sim}$.

Die folgenden Beobachtungen werden wir oft benutzen.

Beobachtung. a) Eine elementare Umformung vom Typ III (III') läßt sich aus elementaren Umformungen von Typ I und Typ II (I' und I'') zusammenstellen. Wenn wir zum Beispiel die j -te und die i -te Spalten miteinander vertauschen möchten können wir folgenderweise vorgehen.

1. Wir addieren die j -te Spalte zu der i -ten (Typ II').
2. Wir subtrahieren die i -te Spalte aus der j -ten (Typ II').
3. Wir addieren die j -te Spalte zu der i -ten (Typ II').
4. Wir multiplizieren die j -te Spalte mit -1 (Typ I').



- b) Durch Transposition gehen elementare Zeilenumformungen in elementare Spaltenumformungen über und umgekehrt.
 c) Seien a_1, \dots, a_m die Zeilenvektoren von A . Bezeichne durch A_I bzw. durch A_{II} die Matrix nach einer (beliebigen) elementaren Zeilenumformung (vom Typ I bzw. II). D.h.

$$A = \begin{pmatrix} \vdots \\ a_i \\ \vdots \end{pmatrix}, \quad A_I = \begin{pmatrix} \vdots \\ \lambda a_i \\ \vdots \end{pmatrix}, \quad A_{II} = \begin{pmatrix} \vdots \\ a_i + \lambda a_j \\ \vdots \end{pmatrix}.$$

Wir setzen

$$S_I := \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \lambda & \ddots & \\ & & & \ddots & 0 \\ 0 & \dots & & 0 & 1 \end{pmatrix} \leftarrow i\text{-te Zeile} \quad S_{II} := \begin{pmatrix} \ddots & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & & 0 \\ 0 & \ddots & \ddots & \ddots & \\ & \lambda & & 1 & 0 \\ 0 & \dots & & 0 & \ddots \end{pmatrix} \leftarrow \begin{array}{l} j\text{-te Zeile} \\ i\text{-te Zeile} \end{array}$$

Mit dieser Notationen gilt $A_I = S_I A$ und $A_{II} = S_{II} A$. [Dies ist leicht einsehbar, wenn man die rechten Seiten ausmultipliziert.] Analog entsteht jede elementare Spaltenumformung durch Multiplikation von rechts mit Matrizen S_I und S_{II} der obigen Form. Bemerke auch dass die Matrizen S_I und S_{II} invertierbar sind. Die Inversen sind leicht anzugeben:

$$S_I^{-1} := \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & \ddots & 0 & & 0 \\ 0 & \ddots & \frac{1}{\lambda} & \ddots & \\ & & 0 & \ddots & 0 \\ 0 & \dots & & 0 & 1 \end{pmatrix} \leftarrow i\text{-te Zeile} \quad S_{II}^{-1} := \begin{pmatrix} \ddots & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & & 0 \\ 0 & \ddots & \ddots & \ddots & \\ & -\lambda & & 1 & 0 \\ 0 & \dots & & 0 & \ddots \end{pmatrix} \leftarrow \begin{array}{l} j\text{-te Zeile} \\ i\text{-te Zeile} \end{array}$$

Da für eine invertierbare Matrix S gilt $\text{Bild}(A) = \text{Bild}(A \cdot S)$ und $\text{Dim Bild}(A) = \text{Dim Bild}(S \cdot A)$, haben wir also den folgenden Satz bewiesen:

- Satz 4.2.1.** a) Elementare Zeilenumformungen entsprechen der Multiplikation von links mit invertierbaren Matrizen, also einem Basiswechsel in Bildraum;
 a') und elementare Spaltenumformungen entsprechen der Multiplikation von rechts mit invertierbaren Matrizen, also einem Basiswechsel in Urbildraum.
 b') Ist A eine Matrix und B eine andere die durch elementare Spaltenumformungen aus A entsteht, so gilt

$$\text{Lin}\{\text{Spaltenvektoren von } A\} = \text{Lin}\{\text{Spaltenvektoren von } B\} \quad \text{und} \quad \text{Dim Lin}\{\text{Zeilenvektoren von } A\} = \text{Dim Lin}\{\text{Zeilenvektoren von } B\}$$

- (iv) Durch elementare Zeilenumformungen vom Typ I und II lässt sich A auf die Form der Einheitsmatrix bringen, d.h.,

$$A \underset{\text{EZU}}{\sim} \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

- c) Durch elementare Umformungen lässt sich A auf dem folgenden Gestalt bringen:

$$A \underset{\text{EU}}{\sim} \left(\begin{array}{ccc|cc} \alpha_{11} & & 0 & \dots & 0 \\ \vdots & \ddots & & & \vdots \\ 0 & \dots & \alpha_{\ell\ell} & \dots & 0 \\ \hline 0 & \dots & 0 & 0 & \dots \\ \vdots & & \vdots & \vdots & 0 \end{array} \right)$$

Beweis. a) Wie haben diese Aussage beim Gauß'schen Verfahren gesehen.

b) "(i) \Rightarrow (ii)": Ist A invertierbar, so auch quadratisch (also $n \times n$). Bringe A auf Zeilenstufenform. Die neue Matrix \tilde{A} hat den gleichen Spaltenrang wie A . Aber \tilde{A} ist invertierbar, und daher sind die Spalten von A linear unabhängig, also $\text{S-Rang}(A) = n$. Hätte \tilde{A} nicht die gewünschte Form, so wäre $\alpha_{nn} = 0$. Das würde auch $\text{S-Rang}(\tilde{A}) < n$ bedeuten, was aber einen Widerspruch liefert.

"(ii) \Rightarrow (iii)": Durch Zeilenumformungen vom Typ II können wir alle nicht Null Elemente über der Diagonale eliminieren. (Beginne von unten mit dem n -ten Zeile).

"(iii) \Rightarrow (iv)": Multipliziere jede Spalte (ESU Typ I) mit dem jeweiligen Skalar $\frac{1}{\lambda_i}$.

"(iv) \Rightarrow (i)": Der Spaltenrang von A ist gleich wie der Spaltenrang der identischen $n \times n$ -Matrix, also $\text{S-Rang}(A) = n$. Dies bedeutet aber, dass die Spaltenvektoren von A eine Basis in ihrem linearen Aufspan bilden, also A ist invertierbar.

- c) Durch elementare Zeilenumformungen können wir A auf Zeilenstufenform bringen und dann durch Vertauschen der Spalten erreichen, dass die Matrix die folgende Form hat:

$$A \underset{\text{EU}}{\sim} \left(\begin{array}{ccc|cc} \alpha_{11} & & * & * & * \\ \vdots & \ddots & & * & * \\ 0 & \dots & \alpha_{\ell\ell} & * & * \\ \hline 0 & \dots & 0 & 0 & \dots \\ \vdots & & \vdots & \vdots & 0 \end{array} \right)$$

mit $\alpha_{ii} \neq 0$ für alle $i = 1, \dots, l$. Durch Zeilenumformungen (oder Spaltenumformungen) vom Typ II können wir es erreichen, dass im ersten Block (oben, links) nur die Diagonale nicht Null wird:

$$\underset{\text{EZU}_{\text{II}}}{\sim} \left(\begin{array}{ccc|cc} \alpha_{11} & & 0 & * & * \\ \vdots & \ddots & & * & * \\ 0 & \dots & \alpha_{\ell\ell} & * & * \\ \hline 0 & \dots & 0 & 0 & \dots \\ \vdots & & \vdots & \vdots & 0 \end{array} \right),$$

dann durch Spaltenumformungen vom Typ II bekommen wir die gewünschte Form. ■

Anwendungen.

Satz 4.3.2. Für jede Matrix A sind der Zeilenrang und der Spaltenrang gleich. Also können wir vom Rang reden:

$$\text{Rang}(A) = \text{Z-Rang}(A) = \text{S-Rang}(A)$$

Beweis. Die elementare Umformungen verändern weder den Zeilenrang noch den Spaltenrang. Durch solche Umformungen können wir die Matrix A auf die Form

$$\left(\begin{array}{ccc|cc} \alpha_{11} & & 0 & \dots & 0 \\ \vdots & \ddots & & & \vdots \\ 0 & \dots & \alpha_{\ell\ell} & \dots & 0 \\ \hline 0 & \dots & 0 & 0 & \dots \\ \vdots & & \vdots & \vdots & 0 \end{array} \right)$$

bringen. Für solche Matrizen gilt aber $\text{Z-Rang} = \ell = \text{S-Rang}$, und somit ist der Beweis fertig. ■

Aufgabe. Gegeben Vektoren $a_1, \dots, a_m \in K^n$ bestimme eine Basis für den linearen Aufspann

$$\text{Lin}\{a_1, \dots, a_m\},$$

(insbesondere: teste auf lineare Unabhängigkeit, bestimme Dimension der linearen Hülle).

Lösung. Wir fassen die Vektoren als die Zeilenvektoren der Matrix A auf, und bringen A auf Zeilenstufenform. Die nicht Null Zeilen bilden eine gesuchte Basis. ■

Aufgabe. Sei A eine quadratische ($n \times n$) Matrix. Stelle fest, ob A invertierbar ist, und bestimme gegebenenfalls die Inverse.

Lösung. Bringe A auf Zeilenstufenform, so kann man die Invertierbarkeit ablesen: A ist genau dann invertierbar, wenn $\ell = n$.

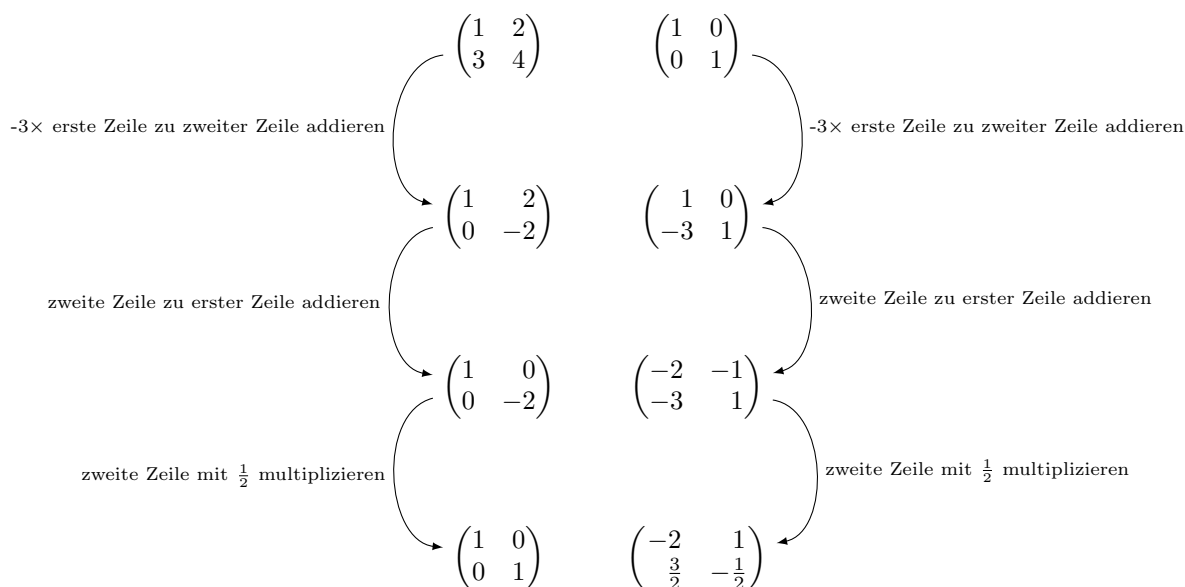
Um die Inverse auszurechnen bringe A durch elementaren Zeilenumformungen auf die Einheitsmatrix. Führe parallel dazu dieselbe elementare Zeilenumformungen an der Einheitsmatrix I durch, dann ergibt sich daraus die Inverse A^{-1} von A . Diese Method funktioniert wegen der folgenden Tatsache. Elementare Zeilenumformungen entsprechen Multiplikation mit invertierbaren Matrizen der Form S_I und S_{II} aus 4.2. Ist A invertierbar, so gibt es S_1, \dots, S_N alle von diesen Gestalt, so dass

$$I = S_1 \cdot S_2 \cdots S_N \cdot A$$

Daher

$$A^{-1} = S_1 \cdot S_2 \cdots S_N \cdot I.$$

Zur Illustration betrachte folgenden Beispiel. Für $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, die vorgeschlagene Methode sieht folgenderweise aus:



■

4.4. Lineare Gleichungssysteme, Bezeichnungen.

Betrachte das lineare Gleichungssystem

$$\begin{aligned}
 & \alpha_{11}x_1 + \alpha_{12}x_2 + \cdots + \alpha_{1n}x_n = \beta_1 \\
 (*) \quad & \alpha_{21}x_1 + \alpha_{22}x_2 + \cdots + \alpha_{2n}x_n = \beta_2 \\
 & \vdots \\
 & \alpha_{m1}x_1 + \alpha_{m2}x_2 + \cdots + \alpha_{mn}x_n = \beta_m
 \end{aligned}$$

Sei $A = (\alpha_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$ die **Koeffizientenmatrix** (also $A \in M_{m,n}$), ferner seien

$$x := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n, \quad \text{und} \quad b := \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}.$$

Dadurch wird (*) zu $Ax = b$ übersetzt. Das lineare Gleichungssystem $Ax = 0$ heißt das zu (*) gehörige homogene lineare Gleichungssystem. Sei

$$T = T_A : \mathbb{K}^n \rightarrow \mathbb{K}^m \quad \text{die lineare Abbildung,}$$

so dass bzgl. der kanonischen Basen $\mathcal{E}_n, \mathcal{E}_m$ A die Matrix von T ist, d.h. $A = M_T^{\mathcal{E}_n, \mathcal{E}_m}$. Nochmal: dies bedeutet

$$T_A(x) = Ax \quad \text{und die Spaltenvektoren von } A \text{ sind die Bilder der Basisvektoren in } \mathcal{E}_n.$$

4.5. Homogenes lineares Gleichungssystem, abstrakt. Betrachte die Gleichung $Ax = 0$, welche $x = 0$ =Nullvektor immer als Lösung hat. Ob der Nullvektor $0 \in \mathbb{K}^n$ die einzige Lösung ist, läßt sich folgenderweise charakterisieren.

Satz 4.5.1. Für $A \in M_{m,n}$ sind die folgenden Aussagen äquivalent.

- Der Vektor $x = 0$ ist einzige Lösung der Gleichung $Ax = 0$.
- Kern $T_A = \{0\}$
- T_A ist injektiv
- $\text{Rang}(A) = n$

Beweis. Die Äquivalenz zwischen a) und b) folgt aus

$$\text{Kern } T_A = \{x : T_A(x) = 0\} = \{x : Ax = 0\}.$$

Wir haben schon gesehen (Abschnitt 3.3), dass die Injektivität von T_A zu $\text{Kern } T_A = \{0\}$ äquivalent ist (wie ging der Beweis?).

Wir haben auch gesehen (Abschnitt 3.4), dass T_A genau dann injektiv ist, wenn T_A linear unabhängige Vektoren auf linear unabhängige Vektoren abbildet. Dies bedeutet aber, dass die Injektivität von T_A dazu äquivalent ist, dass die Spaltenvektoren von A linear unabhängig sind. Dies ist genau dann, wenn $\text{Rang}(A) = n$. ■

Aus dem Beweis bekommen wir den folgenden Korollar.

Satz 4.5.2. Es gilt $\text{Dim}(\text{Bild } A) = \text{Rang}(A)$ und $\text{Dim}(\text{Kern } T_A) = n - \text{Rang}(A)$.

Beweis. Wie oben zeigt man dass $\text{Rang}(A) = \text{Dim}(\text{Bild } A)$. Die andere Gleichheit ist nur die Dimensionsformel. ■

Aufgabe. Löse das homogene lineare Gleichungssystem $Ax = 0$. (Finde *alle* Lösungen. N.B. $x = 0$ ist immer Lösung.)

Lösung. 1) Bringe A durch elementare Zeilenumformungen auf Zeilenstufenform.

2) Bringe die Matrix durch Vertauschen von Spalten auf die Form

$$\tilde{A} = \left(\begin{array}{ccc|cc} \alpha_{11} & & 0 & * & * \\ \vdots & \ddots & & * & * \\ 0 & \dots & \alpha_{\ell\ell} & * & * \\ \hline 0 & \dots & 0 & 0 & \dots \end{array} \right) \quad \text{mit } \alpha_{ii} \neq 0 \text{ für } 1 \leq i \leq \ell$$

(für die Variablen bedeutet dies nur eine Ummummerierung). Dieser Schritt ist natürlich nicht wirklich notwendig (nur die Schreibweise wird somit wesentlich einfacher).

3) Löse die Gleichung $\tilde{A}x = 0$: Da über die letzten $m - \ell$ Zeilen keine Bedingung gestellt ist, wähle $x_{\ell+1} = \mu_1, x_{\ell+2} = \mu_2, \dots, x_n = \mu_{n-\ell}$ beliebig. Halte diese Wahl fest, so kann man eindeutig die restlichen Variablen durch Rekursion bestimmen:

$$\begin{array}{ll} x_\ell & \text{aus } \ell\text{-ter Zeile} \qquad \qquad \qquad \alpha_{\ell\ell}x_\ell + \alpha_{\ell,(\ell+1)}\mu_1 + \dots + \alpha_{\ell n}\mu_{n-\ell} = 0 \\ x_{\ell-1} & \text{aus } (\ell-1)\text{-ter Zeile} \qquad \alpha_{(\ell-1),(\ell-1)}x_{\ell-1} + \alpha_{(\ell-1,\ell)}x_\ell + \alpha_{\ell-1,\ell+1}\mu_1 + \dots + \alpha_{\ell-1,n}\mu_{n-\ell} = 0 \\ \vdots & \\ x_1 & \text{aus } 1\text{-ter Zeile} \qquad \qquad \qquad a_{11}x_1 + \dots + \alpha_{1\ell}x_\ell + \alpha_{1,(\ell+1)}\mu_1 + \dots + \alpha_{1n}\mu_{n-\ell} = 0 \end{array}$$

So bekommt man

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_\ell \\ \mu_1 \\ \vdots \\ \mu_{n-\ell} \end{pmatrix},$$

eine Lösung der homogenen Gleichung $Ax = 0$. ■

Mithilfe der obigen Methode ist eine Basis in dem Lösungsraum sehr leicht zu finden. Wir haben also die freie Wahl für $\mu_1, \dots, \mu_{n-\ell}$. Für

$$\begin{array}{ll} \text{(B)} & (\mu_1, \dots, \mu_{n-\ell}) = (1, 0, 0, \dots, 0) \quad \text{bestimme Lösung } y_1 \\ & (\mu_1, \dots, \mu_{n-\ell}) = (0, 1, 0, \dots, 0) \quad \text{bestimme Lösung } y_2 \\ & \vdots \\ & (\mu_1, \dots, \mu_{n-\ell}) = (0, 0, \dots, 0, 1) \quad \text{bestimme Lösung } y_{n-\ell}. \end{array}$$

So ist $\{y_1, y_2, \dots, y_k\}$ eine Basis in dem Lösungsraum $\text{Kern } T_A$.

Begründung. 1) Wir haben gesehen, dass A durch elementare Zeilenumformungen auf Zeilenstufenform gebracht werden kann. Dies entspricht also Multiplikation mit einer invertierbaren Matrix S von links. Daher ist $Ax = 0$ äquivalent zu $SAx = 0$, wobei SA jetzt Zeilenstufenform hat. Es genügt also nur solche Matrizen zu studieren.

2) Wie schon erwähnt bedeutet dieser Schritt nur eine Ummummerierung der Variablen. Vertauschen von *Spalten* entspricht Multiplikation mit einer invertierbaren Matrix R von *rechts*. Daher ist $SAx = 0$ äquivalent zu $SARR^{-1}x = 0$. Also genügt es, $SARz = 0$ zu lösen mit $SAR = \tilde{A}$ von der obigen Form. Dann setzt man $x = Rz$, so ist x eine Lösung von $SAx = 0$.

3) Offenbar ist $\dim \text{Kern } T_A = n - \text{Rang } A = n - \ell$. Das bedeutet aber, dass es maximal $n - \ell$ linear unabhängige Lösungen existieren können. Die Lösungen $y_1, \dots, y_{n-\ell}$ die wir oben konstruiert haben, sind aber linear unabhängig, also bilden eine Basis in $\text{Kern } T_A$. ■

Beispiel. Betrachte das Gleichungssystem:

$$\begin{array}{rcl} x_1 - 3x_2 + x_3 & & = 0 \\ x_1 - 3x_2 & - x_4 & = 0 \\ & 2x_3 + x_4 & = 0 \end{array}$$

Die Koeffizientenmatrix sieht folgenderweise aus:

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

Wir bringen sie auf Zeilenstufenform:

$$\begin{pmatrix} 1 & -3 & 1 & 0 \\ 1 & -3 & 0 & -1 \\ 0 & 0 & 2 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -3 & 1 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 2 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -3 & 1 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Nach Vertauschen der Spalten:

$$\begin{pmatrix} 1 & 1 & 0 & -3 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

Wir haben also freie Wahl für die letzte Koordinate hier (vor dem Vertauschen dies war die 2-te). Setze also $x_2 = 1$, so liest man leicht die Lösung ab: $x_1 = 3$, $x_3 = 0$, $x_4 = 0$. Der Vektor $x = (x_1, x_2, x_3, x_4)$ bildet eine Basis im Lösungsraum, also die Lösungen des homogenen Problems sind die Vektoren:

$$\text{Kern } T_A = \left\{ \begin{pmatrix} 3\lambda \\ \lambda \\ 0 \\ 0 \end{pmatrix} : \lambda \in \mathbb{R} \right\},$$

ist also ein Gerade durch 0. (Was ist der Rang der Matrix A ?)

Beispiel. Betrachte das Gleichungssystem:

$$\begin{array}{rcl} x_1 - 2x_2 + x_3 & & = 0 \\ x_1 - 2x_2 & - x_4 & = 0 \\ & x_3 + x_4 & = 0 \end{array}$$

Die Koeffizientenmatrix sieht folgenderweise aus:

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

Wir bringen sie auf Zeilenstufenform:

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

Nach Vertauschen der Spalten:

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

Wir haben also freie Wahl für die letzten zwei Koordinaten hier (vor dem Vertauschen waren diese die 2-te und die 4-te). Setze also $x_2 = \lambda$ und $x_4 = \mu$, so liest man leicht die Lösung ab: $-x_3 - x_4 = 0$, also $x_3 = -\mu$; $x_1 + x_3 - 2x_2 = 0$, also $x_1 = 2\lambda + \mu$. Daher ist der Lösungsraum:

$$\text{Kern } T_A = \left\{ \begin{pmatrix} 2\lambda + \mu \\ \lambda \\ -\mu \\ \mu \end{pmatrix} : \lambda, \mu \in \mathbb{R} \right\}.$$

(Was ist diese Menge geometrisch? Was ist die Dimension von $\text{Kern } T_A$? Was ist der Rang von A ?)

Bemerkung. Die Vektorräume $\mathbb{K}^{n-\ell}$ und $\text{Kern } T_A$ sind isomorph! Ein Isomorphismus ist gegeben durch

$$(\mu_1, \dots, \mu_{n-\ell}) \mapsto (x_1, \dots, x_\ell, \mu_1, \dots, \mu_{n-\ell}) = (x_1, x_2, \dots, x_\ell, x_{\ell+1}, \dots, x_n) = x,$$

wobei $x \in \mathbb{K}^n$ die *eindeutige* Lösung von $Ax = 0$ ist mit $x_\ell = \mu_1, \dots, x_n = \mu_{n-\ell}$. Natürlich kann man eine andere Wahl einer Basis in (B) treffen: So ergibt sich wieder ein Isomorphismus und eine andere Basis im Lösungsraum.

4.6. Inhomogene Gleichungssysteme, abstrakt. Betrachte die Gleichung

$$Ax = b.$$

Für allgemeines $b \in \mathbb{K}^m$ heißt diese Gleichung **inhomogen**. Die Lösungen der homogenen Gleichung bilden den Vektorraum $\text{Kern } T_A$, dies stimmt aber für das inhomogene Problem nicht mehr. Was aber offensichtlich ist:

1. Ist $Ax = b$ und $Ax_0 = 0$, so gilt $A(x + x_0) = b$. Also zu beliebiger Lösung des inhomogenen Problems kann man eine Lösung des homogenen Problems addieren, so erhält man wieder eine Lösung des inhomogenen Problems.
2. Ist $Ax_1 = b$ und $Ax_2 = b$, dann gilt $A(x_1 - x_2) = 0$. Also: Die Differenz zweier Lösungen des inhomogenen Gleichungssystem ist eine Lösung des homogenen Problems.

Zusammenfassend können wir das folgende Resultat formulieren.

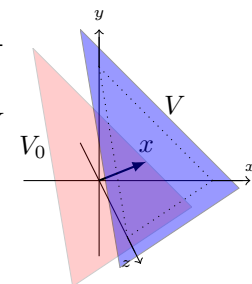
Satz 4.6.1. Die allgemeine Lösungen des inhomogenen linearen Gleichungssystems erhält man durch Addition einer beliebigen, speziellen Lösung des inhomogenen Problems zur allgemeinen Lösung des homogenen Systems.

Diese algebraischen Eigenschaften haben natürlich ihre geometrische Interpretation:

Definition. Sei V ein Vektorraum, $V_0 \subseteq V$ ein linearer Teilraum, $x \in V$ ein fester Vektor. Dann heißt die Menge

$$x + V_0 := \{y \in V : y = x + v_0, v_0 \in V_0\}$$

affiner Teilraum (oder verschobener Unterraum).



Ein Beispiel ist, wie eben festgestellt, die Lösungsmenge eines inhomogenen linearen Gleichungssystems. "Weitere" Beispiele sind allgemeine Geraden, Ebenen, Hyperebenen (siehe LAI).

Nun können wir den folgenden Satz über Lösbarkeit linearer Gleichungssysteme beweisen: Betrachte wieder die Gleichung

$$(*) \quad Ax = b.$$

Satz 4.6.2. 1. Die Gleichung (*) ist genau dann lösbar, wenn $b \in \text{Bild } T_A$

2. (*) ist genau dann *universell lösbar* (d.h. für jede rechte Seite $b \in \mathbb{K}^m$ lösbar), wenn T_A surjektiv ist. Dies ist weiter äquivalent zu $\text{Rang } A = m$.

3. Für jedes b hat (*) genau dann *höchstens eine* Lösung, wenn $\text{Kern } T_A = \{0\}$. Dies ist weiter äquivalent zu $\text{Rang } A = \text{Rang } T_A = n$.
4. (*) ist genau dann für jedes b *eindeutig lösbar*, wenn T_A bijektiv. Dies ist weiter äquivalent zu der Invertierbarkeit von A und zu $\text{Rang } A = m = n$.

Beweis. 1. Die Aussage ist trivial, denn $T_A(x) = Ax$.

2. Die erste Aussage folgt aus 1. Wir wollen noch zeigen, dass die Surjektivität von T_A zu der Aussage über $\text{Rang}(A)$ äquivalent ist. Die Surjektivität von T_A ist zu $\text{Dim Bild } T_A = m$ äquivalent. Wir wissen, dass

$$\text{Bild}(T_A) = \text{Lin}\{\text{Spaltenvektoren von } A\} \text{ gilt.}$$

Daher ist die Surjektivität zu $\text{Rang}(A) = m$ äquivalent.

3. Aus dem obigen Satz folgt, dass wenn zwei Lösungen des inhomogenen Problems existieren, so existiert auch eine nicht 0 Lösung des homogenen Problems. Dies beweist 3.

4. Diese Aussage folgt aus 2. und 3. ■

Aufgabe. Löse das inhomogene Gleichungssystem $Ax = b$. Wir definieren

$$(A, b) := \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} & b_1 \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} & b_2 \\ \vdots & & & \vdots & \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} & b_m \end{pmatrix},$$

die so genannte **erweiterte Koeffizientenmatrix**.

1. Stelle fest ob (*) lösbar ist: Bringe dazu (A, b) auf Zeilenstufenform (\tilde{A}, \tilde{b}) . So ist $Ax = b$ genau dann lösbar, wenn (\tilde{A}, \tilde{b}) die folgende Form hat:

$$(\tilde{A}, \tilde{b}) = \left(\begin{array}{cccc|c} 0 & & & * & * \\ \dots & \dots & 0 & * & * \\ & \dots & \dots & 0 & * \\ 0 & \dots & & \dots & 0 & 0 \end{array} \right).$$

$\underbrace{\hspace{10em}}_{\tilde{A}} \quad \underbrace{\hspace{2em}}_{\tilde{b}}$

Dies passiert aber genau dann, wenn $\text{Rang}(A) = \text{Rang}((\tilde{A}, \tilde{b}))$.

Begründung. Die Gleichung $Ax = b$ ist genau dann lösbar, falls $b \in \text{Bild}(T_A)$. Der (Spalten)Rang von $S\text{-Rang}((A, b))$ ist mindestens so groß wie der Rang von A , und er ist auch echt größer, wenn b linear unabhängig von den Spalten von A ist. Deswegen ist $S\text{-Rang}(A) = S\text{-Rang}((A, b))$ äquivalent zu

$$b \in \text{Lin}\{\text{Spaltenvektoren von } A\} = \text{Bild}(T_A).$$

Beispiel. Nun zu dem vorherigen Beispiel:

$$Ax = \begin{pmatrix} 1 & -2 & 1 & 0 \\ 1 & -2 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = b.$$

Die erweiterte Koeffizientenmatrix sieht so aus:

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

Wir bringen sie auf Zeilenstufenform:

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

Diese Matrix hat Rang 2, gleich wie die Matrix A . Wir stellen also fest, dass die Gleichung lösbar ist. Wenn man eine andere rechte Seite $b \in \mathbb{R}^3$ wählt, wird die Gleichung nicht lösbar. Zum Beispiel setze $b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. So bekommt man das Folgende:

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

Diese Matrix hat Rang 3, echt größer als die Matrix A . Wir stellen also fest, dass die Gleichung nicht lösbar ist.

2. Bestimme die allgemeine Lösung, falls (*) lösbar ist. Bringe (A, b) auf Zeilenstufenform und dann durch Vertauschen von Spalten auf die Form

$$(\tilde{A}, \tilde{b}) = \left(\begin{array}{ccc|cc} \tilde{\alpha}_{11} & & * & * & \tilde{b}_1 \\ \vdots & \ddots & & * & \vdots \\ 0 & \dots & \tilde{\alpha}_{\ell\ell} & * & \tilde{b}_\ell \\ \hline 0 & & \dots & 0 & 0 \end{array} \right),$$

wobei $\tilde{\alpha}_{ii} \neq 0$ für $i = 1, \dots, \ell$. Hier ist der Block, den wir mit * bezeichnet haben, uninteressant. Betrachte also das Gleichungssystem (ein Teilsystem)

$$\begin{pmatrix} \tilde{\alpha}_{11} & & * \\ \vdots & \ddots & \\ 0 & \dots & \tilde{\alpha}_{\ell\ell} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_\ell \end{pmatrix} = \begin{pmatrix} \tilde{b}_1 \\ \vdots \\ \tilde{b}_\ell \end{pmatrix},$$

das einfach zu lösen ist: beginne von unten und setze ein. Eine spezielle Lösung des kompletten Gleichungssystems (*) ergibt sich zum Beispiel, wenn man $x_{\ell+1} = \dots = x_n = 0$ setzt. Nun bestimme die allgemeine Lösung des *homogenen* Problems, wie es in 4.5 gezeigt wurde. Die allgemeine Lösung der *inhomogenen* Gleichung bekommt man mithilfe des ersten Satzes in 4.6.

Begründung. Die Blockmatrix (\tilde{A}, \tilde{b}) habe Zeilenstufenform, d.h. entsteht als $(\tilde{A}, \tilde{b}) = S(A, b)$ mit S invertierbare $m \times m$ -Matrix. Genauer $SA = \tilde{A}$ und $Sb = \tilde{b}$. Daher ist $Ax = b$ zu $SAX = Sb$, also zu $\tilde{A}x = \tilde{b}$ äquivalent. ■

Beispiel. Auf unserem Beispiel sieht es folgenderweise aus:

$$(A, b) = \begin{pmatrix} 1 & -2 & 1 & 0 & 1 \\ 1 & -2 & 0 & -1 & 2 \\ 0 & 0 & 1 & 1 & -1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -2 & 1 & 0 & 1 \\ 0 & 0 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = (\tilde{A}, \tilde{b}).$$

(Wir vertauschen jetzt die Spalten nicht, aber wir könnten es tun.) Wir wählen $x_2 = x_4 = 0$, und müssen nun die Gleichung

$$\begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

lösen. Durch Anschauen bekommt man $x_3 = -1$ und $x_1 = 2$. Wir haben früher auch alle Lösungen des homogenen Problems bestimmt. So ist die Lösungsmenge der Gleichung der folgende affine Teilraum

$$\left\{ \begin{pmatrix} 2 \\ 0 \\ -1 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ \lambda \\ 0 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix} : \lambda, \mu \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} 2+2\lambda+\mu \\ \lambda \\ -1-\mu \\ \mu \end{pmatrix} : \lambda, \mu \in \mathbb{R} \right\}.$$

4.7. Determinanten. Für gegebene Vektoren $a_1, \dots, a_n \in \mathbb{K}^n$, betrachte die Matrix deren Zeilenvektoren a_1, \dots, a_n sind, also

$$A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

Ab sofort benutzen wir diese bequeme Schreibweise.

Definition. Sei

$$M : \underbrace{\mathbb{K}^n \times \cdots \times \mathbb{K}^n}_n \rightarrow \mathbb{K}$$

eine n -lineare antisymmetrische Abbildung (n -Linearform), so dass $M(e_1, \dots, e_n) = 1$ (e_1, \dots, e_n ist die kanonische Basis). So heißt die Abbildung

$$\text{Det} : M_{n,n}(\mathbb{K}) \rightarrow \mathbb{K}$$

definiert für $A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ durch

$$\text{Det}(A) = M(a_1, a_2, \dots, a_n)$$

eine **Determinante**.

Wir werden später zeigen, dass es eine (und nur eine) Determinante gibt. Zunächst die Eigenschaften einer Determinante nochmal und ganz explizit:

Bemerkung. Die Abbildung $\text{Det} : M_{n,n}(\mathbb{K}) \rightarrow \mathbb{K}$ ist eine Determinante, falls gelten:

(D1) Für $\lambda, \mu \in \mathbb{K}$, $A \in M_{n,n}(\mathbb{K})$, $b \in \mathbb{K}^n$ gilt

$$\text{Det} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ \lambda a_i + \mu b \\ \vdots \\ a_n \end{pmatrix} = \lambda \text{Det} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} + \mu \text{Det} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ b \\ \vdots \\ a_n \end{pmatrix}$$

Also: Det ist linear in den Zeilenvektoren. Insbesondere ist eine Zeile der Matrix A komplett 0, so gilt $\text{Det}(A) = 0$. (Diese Eigenschaft entspricht der n -Linearität von M .)

(D2) Wenn man die i -te und die j -te Zeilen von A vertauscht, wird die Determinante mit -1 multipliziert. D.h.

$$\text{Det} \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} = -\text{Det} \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \end{pmatrix}$$

(Dies entspricht der Antisymmetrie von M .) Man sagt auch, dass die Determinante **alternierend** ist.

(D3) Für die $n \times n$ -Einheitsmatrix gilt

$$\text{Det} \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & & \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} = 1$$

(Diese Eigenschaft, die **Normiertheit**, entspricht der Forderung $M(e_1, \dots, e_n) = 1$.)

Beispiele haben wir schon gesehen:

Beispiel. Für $n = 2$, das Spatprodukt in \mathbb{R}^2 :

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

4.8. Eigenschaften der Determinate.

Satz 4.8.1. (D4) Sind zwei Zeilen von A gleich, z.B. $a_i = a_j$ ($i \neq j$), so gilt $\text{Det}(A) = 0$.

(D5) Bei elementaren Zeilenumformungen von Typ II ändert sich die Determinante nicht.

Beweis. (D4): Beim Vertauschen der i -ten und der j -ten Zeilen bleibt A unverändert, dabei muss aber die Determinante die Gegenzahl werden. Das ist nur möglich wenn $\text{Det}(A) = 0$. Genauer: mit

$$A = \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix}, \quad \text{und} \quad \tilde{A} = \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \end{pmatrix}$$

gelten $\text{Det}(A) = \text{Det}(\tilde{A})$ (denn $A = \tilde{A}$) und $\text{Det}(A) = -\text{Det}(\tilde{A})$ (wegen (D2)).

(D5): Addieren wir λ -mal die i -te Zeile zu der j -ten:

$$\text{Det} \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j + \lambda a_i \\ \vdots \end{pmatrix} \stackrel{(D1)}{=} \text{Det} \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} + \underbrace{\lambda \text{Det} \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_i \\ \vdots \end{pmatrix}}_{=0} = \text{Det} \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix},$$

wobei, wegen (D4), der zweite Term 0 ist. ■

Besonders wichtig sind die Determinante der Matrizen, die elementaren Umformungen entsprechen.

Beispiel 4.8.2. Betrachte die Matrizen

$$S_I := \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \lambda & \ddots & \\ & & & \ddots & 0 \\ 0 & \dots & & 0 & 1 \end{pmatrix} \leftarrow i\text{-te Zeile} \quad S_{II} := \begin{pmatrix} \ddots & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & & 0 \\ 0 & \ddots & \ddots & \ddots & \\ & \lambda & & 1 & 0 \\ 0 & \dots & & 0 & \ddots \end{pmatrix} \leftarrow \begin{array}{l} j\text{-te Zeile} \\ i\text{-te Zeile} \end{array}$$

So gilt $\text{Det}(S_I) = \lambda$ und $\text{Det}(S_{II}) = 1$. Betrachte nun die Matrix S_{III} die aus $I_{n \times n}$ durch Vertauschen der i -ten und j -ten Zeilen entsteht, d.h.

$$S_{III} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 1 & 0 \\ 0 & \ddots & \ddots & \ddots & \\ & 1 & & 0 & 0 \\ 0 & \dots & & 0 & 1 \end{pmatrix} \leftarrow \begin{array}{l} j\text{-te Zeile} \\ i\text{-te Zeile} \end{array}$$

So vertauscht die Multiplikation mit S_{III} von links die Zeilen einer Matrix A (und ebenso die Multiplikation von rechts vertauscht die Spalten). Es gilt $\text{Det}(S_{III}) = -1$.

In manchen Fällen ist die Determinante sehr leicht auszurechnen:

Satz 4.8.3. Sei $A \in M_{n,n}(\mathbb{K})$ eine $n \times n$ -Matrix, welche in Zeilenstufenform obere Dreiecksform hat:

$$\begin{pmatrix} \alpha_{11} & & & * \\ 0 & \ddots & & \\ \vdots & & \ddots & \\ 0 & \dots & 0 & \alpha_{nn} \end{pmatrix},$$

(eventuell mit $\alpha_{ii} = 0$ möglich). So gilt

$$\text{Det}(A) = \alpha_{11} \cdot \alpha_{22} \cdots \alpha_{nn}.$$

Beweis. Ist $\alpha_{ii} = 0$ so ist auch $\alpha_{jj} = 0$ für $j \geq i$. Insbesondere ist $\alpha_{nn} = 0$, also die letzte Zeile ist 0. Somit $\text{Det}(A) = 0$ wegen (D1), die Aussage stimmt in diesem Fall. Nehme an, das $\alpha_{ii} \neq 0$ für alle $i = 1, \dots, n$. Wir können durch elementare Zeilenumformungen von Typ II die Matrix auf Diagonalform bringen:

$$B = \begin{pmatrix} \alpha_{11} & 0 & \dots & 0 \\ 0 & \ddots & & 0 \\ \vdots & & \alpha_{(n-1)(n-1)} & \vdots \\ 0 & \dots & 0 & \alpha_{nn} \end{pmatrix}.$$

Es gilt $\text{Det}(B) = \text{Det}(A)$, und $\text{Det}(B)$ ist leicht auszurechnen:

$$\text{Det} \begin{pmatrix} \alpha_{11} & 0 & \dots & 0 \\ 0 & \ddots & & 0 \\ \vdots & & \alpha_{(n-1)(n-1)} & \vdots \\ 0 & \dots & 0 & \alpha_{nn} \end{pmatrix} = \alpha_{11} \cdot \alpha_{22} \cdots \alpha_{nn} \text{Det}(I_{n \times n}) = \alpha_{11} \cdot \alpha_{22} \cdots \alpha_{nn}.$$

■

Wir haben auch das folgende Resultat teilweise bewiesen:

Satz 4.8.4. Für eine $n \times n$ -Matrix sind die folgenden Aussagen äquivalent:

- (i) $\text{Det}(A) = 0$.
- (ii) A ist *nicht* invertierbar.
- (iii) $\text{Rang}(A) < n$.
- (iv) die Zeilen (bzw. die Spalten) von A sind *linear abhängig*.

Beweis. Bringe A auf Zeilenstufenform. Dadurch ändert sich die Determinante nicht, und die neue Matrix ist genau dann invertierbar, wenn A invertierbar ist. Dies beweist die Äquivalenz “(i) \Leftrightarrow (ii)”. Die restliche Implikationen wissen wir schon. ■

Ein weitere Korollar ist das Folgende:

Satz 4.8.5. Es gibt höchstens eine Determinante.

Beweis. Bringe A auf Zeilenstufenform, dann ist die Determinante nach Satz 4.8.3 eindeutig auszurechnen. ■

Der folgende Satz ist sehr wichtig:

Satz 4.8.6 (Multiplikationssatz). Für $A, B \in M_{n,n}(\mathbb{K})$ gilt

$$(D6) \quad \text{Det}(AB) = \text{Det}(A) \cdot \text{Det}(B).$$

Insbesondere gelten die folgenden Aussagen:

- (i) $\text{Det}(A^N) = \text{Det}(A)^N$.
- (ii) $\text{Det}(AB) = \text{Det}(A) \text{Det}(B) = \text{Det}(B) \text{Det}(A) = \text{Det}(BA)$.
- (iii) $\text{Det}(A^{-1}) = (\text{Det}(A))^{-1}$, falls A invertierbar ist.
- (iv) $\text{Det}(SAS^{-1}) = \text{Det}(A)$, falls S invertierbar ist (ähnliche Matrizen haben dieselbe Determinante).

Beweis. Die Aussagen (i)–(iv) sind leicht, sobald man die Gleichheit $\text{Det}(AB) = \text{Det}(A) \text{Det}(B)$ bewiesen hat.

Ist A nicht invertierbar, so ist $\text{Rang}(A) < n$ und somit $\text{Rang}(AB) < n$, also auch AB ist nicht invertierbar. In diesem Fall gilt

$$\text{Det}(AB) = 0 = 0 \cdot \text{Det}(B) = \text{Det}(A) \text{Det}(B),$$

also die Aussage. Wir können daher annehmen, dass A invertierbar ist. Wir haben aber gesehen, dass dann A durch elementare Zeilenumformungen auf Diagonalform gebracht werden kann

$$A \underset{\text{EZU}_{\text{II}}}{\sim} \tilde{A} = \begin{pmatrix} \alpha_{11} & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \alpha_{nn} \end{pmatrix},$$

mit $\alpha_{ii} \neq 0$. Wir wissen auch, dass $\tilde{A} = SA$ mit einer invertierbaren Matrix S . Führe die gleichen elementaren Zeilenumformungen auf AB durch, so bekommen wir $S(AB)$. Daraus folgt:

$$\text{Det}(AB) = \text{Det}(SAB) = \text{Det}(\tilde{A}B).$$

Nun rechnen wir $\text{Det}(\tilde{A}B)$ aus:

$$\begin{aligned} \text{Det}(\tilde{A}B) &= \text{Det} \begin{pmatrix} \tilde{\alpha}_{11} & & & 0 \\ 0 & \ddots & & \\ \vdots & & \ddots & \\ 0 & \dots & 0 & \tilde{\alpha}_{nn} \end{pmatrix} \begin{pmatrix} \beta_{11} & & & \beta_{1n} \\ & \ddots & & \\ \vdots & & \ddots & \\ \beta_{n1} & \dots & & \beta_{nn} \end{pmatrix} = \text{Det} \begin{pmatrix} \tilde{\alpha}_{11}\beta_{11} & & & \tilde{\alpha}_{11}\beta_{1n} \\ & \ddots & & \\ \vdots & & \ddots & \\ \tilde{\alpha}_{nn}\beta_{n1} & \dots & & \tilde{\alpha}_{nn}\beta_{nn} \end{pmatrix} \\ &\stackrel{\text{(D1)}}{=} \tilde{\alpha}_{11} \cdot \tilde{\alpha}_{22} \cdots \tilde{\alpha}_{nn} \text{Det}(B) = \text{Det}(\tilde{A}) \text{Det}(B) = \text{Det}(A) \text{Det}(B). \end{aligned}$$

■

Bemerkung. Sei V ein (n -dimensionaler) Vektorraum über \mathbb{K} , und sei $T : V \rightarrow V$ linear. Sind \mathcal{L} und \mathcal{L}' Basen in V , so ist Beziehung zwischen der Matrizen von T bzgl. \mathcal{L} und \mathcal{L}' das Folgende:

$$M_T^{\mathcal{L}', \mathcal{L}'} = M_I^{\mathcal{L}, \mathcal{L}'} M_T^{\mathcal{L}, \mathcal{L}} M_I^{\mathcal{L}', \mathcal{L}}.$$

Daraus folgt aber

$$\text{Det}(M_T^{\mathcal{L}', \mathcal{L}'}) = \text{Det}(M_T^{\mathcal{L}, \mathcal{L}}),$$

d.h. die Determinante ist *unabhängig* von der Wahl der Basen. Wir können also von der Determinante der linearen Abbildung reden.

Satz 4.8.7. Sei $A \in M_{n,n}(\mathbb{K})$.

- Sei B eine Matrix die aus A durch vertauschen zweier Spalten entstanden ist. Es gilt $\text{Det}(A) = -\text{Det}(B)$.
- Es gilt $\text{Det}(A^\top) = \text{Det}(A)$.

Beweis. a) Es gilt $B = AS$ mit S eine Matrix die Vertauschen der zwei Spalten beschreibt. Wir wissen bereits $\text{Det}(S) = -1$ (siehe Beispiel 4.8.2). Wegen des Multiplikationssatzes gilt

$$\text{Det}(B) = \text{Det}(AS) = \text{Det}(A) \text{Det}(S) = -\text{Det}(A).$$

b) Ist $\text{Det}(A) = 0$, so ist $\text{S-Rang}(A) < n$ und somit $\text{Z-Rang}(A^\top) < n$. In diesem Fall gilt also $\text{Det}(A^\top) = 0$. Nehme an, dass $\text{Det}(A) \neq 0$, d.h., A ist invertierbar. Bringe A durch elementare Zeilenumformungen auf Diagonalform B , d.h., $B = SA$. Es gilt dann

$$\text{Det}(A) = \text{Det}(B) = \text{Det}(B^\top) = \text{Det}((SA)^\top) = \text{Det}(A^\top S^\top) = \text{Det}(A^\top) \text{Det}(S^\top) = \text{Det}(A^\top).$$

■

4.9. Berechnung der Determinante. Sei $A \in M_{n,n}(\mathbb{K})$ eine $n \times n$ -Matrix. Das Berechnen der Determinante werden wir auf die Bestimmung von $(n-1) \times (n-1)$ Determinanten zurückführen. Dazu brauchen wir die folgenden Konstruktionen:

Definition. Für $A \in M_{n,n}(\mathbb{K})$ und festes $1 \leq i, j \leq n$ führe die folgenden Notationen ein:

$$A_{ij} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1(j-1)} & \alpha_{1j} & \alpha_{1(j+1)} & \cdots & \alpha_{1n} \\ \vdots & & & & & & \vdots \\ \alpha_{i1} & \cdots & \alpha_{i(j-1)} & \alpha_{ij} & \alpha_{i(j+1)} & \cdots & \alpha_{in} \\ \vdots & & & & & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{n(j-1)} & \alpha_{nj} & \alpha_{n(j+1)} & \cdots & \alpha_{nn} \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1(j-1)} & \alpha_{1(j+1)} & \cdots & \alpha_{1n} \\ \vdots & & & & & \vdots \\ \vdots & & & & & \vdots \\ \vdots & & & & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{n(j-1)} & \alpha_{n(j+1)} & \cdots & \alpha_{nn} \end{pmatrix}$$

d.h., streiche die i -te Zeile und die j -te Spalte aus A ; Also $A_{ij} \in M_{n-1,n-1}(\mathbb{K})$.

$$\tilde{A}_{ij} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1(j-1)} & 0 & \alpha_{1(j+1)} & \cdots & \alpha_{1n} \\ \vdots & & & \vdots & & & \vdots \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & \vdots & & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{n(j-1)} & 0 & \alpha_{n(j+1)} & \cdots & \alpha_{nn} \end{pmatrix}$$

d.h., ersetze α_{ij} durch 1 und sonst i -te Zeile und die j -te Spalte komplett durch 0; Also $\tilde{A}_{ij} \in M_{n,n}(\mathbb{K})$.

$$\hat{A}_{ij} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1(j-1)} & \alpha_{1j} & \alpha_{1(j+1)} & \cdots & \alpha_{1n} \\ \vdots & & & \vdots & & & \vdots \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & \vdots & & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{n(j-1)} & \alpha_{nj} & \alpha_{n(j+1)} & \cdots & \alpha_{nn} \end{pmatrix}$$

d.h., ersetze α_{ij} durch 1 und sonst i -te Zeile komplett durch 0; Also $\hat{A}_{ij} \in M_{n,n}(\mathbb{K})$.

Hilfssatz 4.9.1. Mit den obigen Notationen gilt:

$$\text{Det}(\hat{A}_{ij}) = \text{Det}(\tilde{A}_{ij}) = (-1)^{i+j} \text{Det}(A_{ij}).$$

Beweis. 1) Die Matrix \hat{A}_{ij} entsteht aus \tilde{A}_{ij} durch elementare Zeilenumformungen. Daher haben diese Matrizen die gleiche Determinante.

2) Durch $n-i$ Zeilenvertauschungen und $n-j$ Spaltenvertauschungen bringen wir die Matrix \tilde{A}_{ij} auf die folgende Form:

$$B = \begin{pmatrix} A_{ij} & 0 \\ 0 & 1 \end{pmatrix},$$

dabei ändert sich die Determinante durch einen Faktor $(-1)^{n-i+n-j} = (-1)^{i+j}$, es gilt also $\text{Det}(B) = (-1)^{i+j} \text{Det}(A_{ij})$. Um $\text{Det}(B)$ zu bestimmen bringen wir B auf Zeilenstufenform, dabei bleibt die letzte Zeile stehen:

$$\tilde{B} = \begin{pmatrix} B_{ij} & 0 \\ 0 & 1 \end{pmatrix},$$

wobei B_{ij} eine Zeilenstufenform von A_{ij} ist. Aus Satz 4.8.4.8.3 folgt $\text{Det}(B) = \text{Det}(B_{ij}) = \text{Det}(A_{ij})$, und somit ist die komplette Behauptung bewiesen. ■

Satz 4.9.2 (Laplace'scher Entwicklungssatz). Sei $A \in M_{n,n}(\mathbb{K})$. Für jedes i mit $1 \leq i \leq n$ gilt:

$$\text{Det}(A) = \sum_{j=1}^n (-1)^{i+j} \alpha_{ij} \text{Det}(A_{ij}).$$

Beweis. Es gilt

$$\text{Det}(A) = \sum_{j=1}^n \alpha_{ij} \text{Det}(\widehat{A}_{ij}),$$

denn die Determinante ist linear in der i -ten Zeile (siehe (D1)). Zusammen mit dem obigen Hilfssatz folgt die Behauptung. ■

4.10. Die komplementäre Matrix.

Satz 4.10.1. Sei $A \in M_{n,n}(\mathbb{K})$, und seien $1 \leq i, i' \leq n$, $i \neq i'$. So gilt:

$$\sum_{j=1}^n (-1)^{i+j} \alpha_{ij} \text{Det}(A_{i'j}) = 0.$$

Beweis. Sei B die Matrix, die aus A dadurch entsteht, dass die i' -te Zeile von A durch i -te Zeile ersetzt wird. So hat B zwei gleiche Zeilen also $\text{Det}(B) = 0$. Außerdem gilt

$$\text{Det}(B) = \sum_{j=1}^n (-1)^{i+j} \alpha_{ij} \text{Det}(A_{i'j}) = 0$$

wegen dem Satz von Laplace. ■

Definition. Für $A \in M_{n,n}(\mathbb{K})$ und $1 \leq i, j \leq n$ definiere

$$\tilde{\alpha}_{ij} := (-1)^{i+j} \text{Det}(A_{ji}) \quad \text{und} \quad \tilde{A} := \begin{pmatrix} \tilde{\alpha}_{11} & \dots & \tilde{\alpha}_{1n} \\ \vdots & & \vdots \\ \tilde{\alpha}_{n1} & \dots & \tilde{\alpha}_{nn} \end{pmatrix}.$$

So heißt \tilde{A} die zu A **komplementäre Matrix** (oder auch die **Adjunkt** von A).

Satz 4.10.2. Für $A \in M_{n,n}(\mathbb{K})$ gilt

$$A\tilde{A} = \text{Det}(A) \cdot I_{n \times n}.$$

Falls A invertierbar ist (d.h. $\text{Det}(A) \neq 0$), so gilt

$$A^{-1} = \frac{1}{\text{Det}(A)} \tilde{A}.$$

Beweis. Wir rechnen einfach das Produkt $A\tilde{A}$ aus. Seien $1 \leq i, j \leq n$ gegeben. So gilt

$$(A\tilde{A})_{ij} = \sum_{k=1}^n \alpha_{ik} \tilde{\alpha}_{kj} = \sum_{k=1}^n \alpha_{ik} (-1)^{k+j} \text{Det}(A_{jk}) = \begin{cases} \text{Det}(A) & \text{falls } i = j \\ 0 & \text{sonst.} \end{cases}$$

Die Aussage über die Inverse folgt unmittelbar aus dieser Gleichung. ■

Beispiel. Sei $A \in M_{2,2}(\mathbb{K})$,

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Die Determinante von A ist $ac - bd$, ist sie nicht 0 so ist die Inverse von A durch

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

gegeben.

4.11. Berechnung und Interpretation von Determinanten.

1. In Dimension $n = 1$: $A = (\alpha)$ und somit $\text{Det}(A) = \alpha$.

2. In Dimension $n = 2$: Für $A \in M_{2,2}(\mathbb{K})$ gilt

$$\text{Det}(A) = \begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix} = \alpha_{11} \text{Det}(\alpha_{22}) - \alpha_{12} \text{Det}(\alpha_{21}) = \alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21}$$

Interpretation: Für $a_1 = \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \end{pmatrix}$ und $a_2 = \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \end{pmatrix}$, gilt

$\text{Det}(A) =$ orientierte Fläche des von a_1, a_2 aufgespannten Parallelogramms.

3. In Dimension $n = 3$: Für $A \in M_{3,2}(\mathbb{K})$ gilt (Entwicklung nach erster Zeile)

$$\begin{aligned} \text{Det}(A) &= \begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix} = \alpha_{11} \begin{vmatrix} \alpha_{22} & \alpha_{23} \\ \alpha_{32} & \alpha_{33} \end{vmatrix} - \alpha_{12} \begin{vmatrix} \alpha_{21} & \alpha_{23} \\ \alpha_{31} & \alpha_{33} \end{vmatrix} + \alpha_{13} \begin{vmatrix} \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{vmatrix} \\ &= \alpha_{11}(\alpha_{22}\alpha_{33} - \alpha_{23}\alpha_{32}) + \alpha_{12}(\alpha_{23}\alpha_{31} - \alpha_{21}\alpha_{33}) + \alpha_{13}(\alpha_{21}\alpha_{32} - \alpha_{22}\alpha_{31}). \end{aligned}$$

Sarrus'sche Regel: Rechts-unten mit +, links -untenmit – Vorzeichen. A C H T U N G : nur für $n = 3!$

Interpretation: Ist $A = (a_1, a_2, a_3)$ (Spaltenvektoren) so gilt

$\text{Det}(A) = \langle a_1 \times a_2, a_3 \rangle =$ orientiertes Volumen des von a_1, a_2, a_3 aufgespannten Parallelepipeds.

4. Um die Determinante in allgemeinem Fall berechnen zu können benutze 1) Entwicklungssatz oder 2) bringe auf Zeilenstufenform. Hier ist das Vorzeichnenschema leicht zu merken, es folgt einen Schachbrettmuster:

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ \vdots & \vdots & & & \end{pmatrix}$$

Die Interpretation ist dieselbe nur in mehrerer Dimension: Für $A = (a_1, \dots, a_n)$ (Spaltenvektoren) ist

$\text{Det}(A) =$ orientiertes Volumen des a_1, \dots, a_n aufgespannten (n -dim.) Parallelepipeds.

4.12. Permutationen.

Definition. Eine **Permutation** der Elementen $\{1, \dots, n\}$ ist eine bijektive Abbildung

$$\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}.$$

Es ist leicht zu sehen, dass die **Hintereinanderführung** zweier Permutationen ergibt wieder eine Permutation. Man definiert dadurch das **Produkt** zweier Permutationen σ, π

$$\sigma\pi = \pi \circ \sigma, \quad \text{d.h.} \quad \sigma\pi(k) = \pi \circ \sigma(k) = \pi(\sigma(k)).$$

Beachte hier die umgekehrte Reihenfolge! Man kann Permutationen so vorstellen, wie Umordnung der Elementen $\{1, \dots, n\}$. Es ist bequem also eine Permutation als

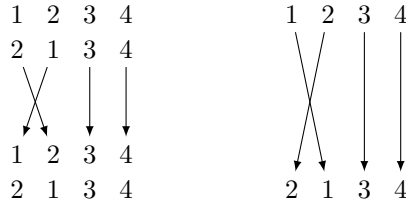
$$\begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$$

zu notieren.

Beispiel. 1. Die Permutation σ

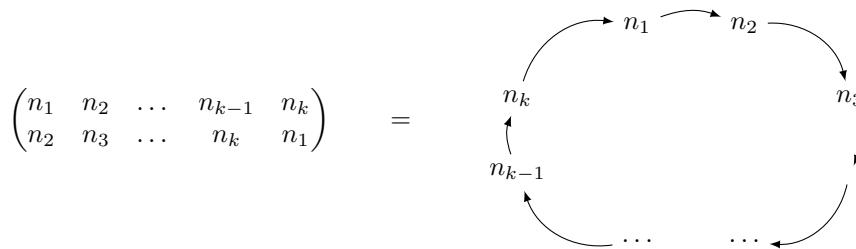
$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix}$$

vertauscht 1 und 2, und lässt 3 und 4 fest. Also $\sigma(1) = 2, \sigma(2) = 1, \sigma(3) = 3, \sigma(4) = 4$. Das Produkt rechnet man so aus:

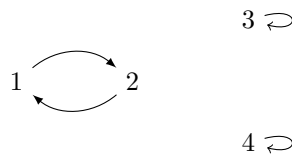


Also $\sigma\sigma =$ die **identische Permutation**.

2. Bezeichne mit S_n die Menge aller Permutationen auf n Elementen. Versehen mit dem obigen Produkt wird S_n zu einer Gruppe, der so genannten **Symmetrischen Gruppe**. Die Anzahl der Elemente in S_n ist $n! = 1 \cdot 2 \cdot \dots \cdot n$. Die Inverse einer Permutation ist die Inverse der bijektiven Abbildung σ . Das neutrale Element $\mathbf{1} \in S_n$ ist die identische Abbildung, d.h. $\mathbf{1}(k) = k$ für jedes $k = 1, \dots, n$. Für $\pi \in S_n$ ist $\{\pi\sigma : \sigma \in S_n\} = S_n = \{\sigma\pi : \sigma \in S_n\}$.
3. Es ist möglich, dass eine Permutation eine Teilmenge $M \subseteq \{1, \dots, n\}$ invariant lässt. Wenn M invariant ist und keine weitere, nichtleere **invariante Mengen** erhält, nennen wir es eine **Zyklus**. Anders gesagt ist ein Zyklus in einer Permutation eine Folge von unterschiedlichen Elementen n_1, \dots, n_k mit $\sigma(n_i) = n_{i+1}$ für $i \leq k - 1$ und $\sigma(n_k) = n_1$. Einen solchen Zyklus können wir als $(n_1 n_2 \dots n_k)$ notieren. Wir nennen die Permutation



auch einen Zyklus. Hier heißt k die Länge des Zyklus. Ein Zyklus von Länge 2 heißt auch **Transposition**. Die Inverse einer Transposition ist dieselbe Transposition (warum?). In dem obigen Beispiel ist also (12) ein Zyklus (sogar eine Transposition), sowie (3) und (4) sind auch Zyklen. Eine bequeme Notation ist also für das obige σ : $(12)(3)(4)$. Ein Zyklus der Länge 1 ist ein Element, das auf sich abgebildet wird, also ein **Fixpunkt** der Permutation. Im obigen Beispiel sind 3 und 4 Fixpunkte.



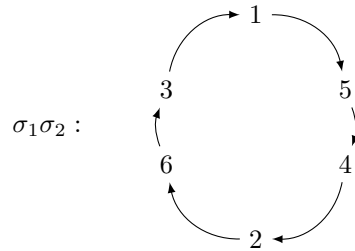
4. Sei σ_1 die Permutation von 6 Elementen (in Zyklus-Notation) $(1324)(56)$ und sei σ_2 die Permutation $(12)(35)(46)$. Wir sehen also, σ_1 und σ_2 keinen Fixpunkt haben.



Das Produkt von σ_1 und σ_2 ist

$$\sigma_1\sigma_2 = (1324)(56)(12)(35)(46) = (154263),$$

also $\sigma_1\sigma_2$ besteht aus nur einer Zyklus, ist also eine **zyklische Permutation**.



Satz 4.12.1. Jede Permutation lässt sich eindeutig als Produkt von disjunkten Zyklen schreiben.

Beweis. Fange mit 1 an:

$$1 \rightarrow \sigma(1) \rightarrow \sigma(\sigma(1)) \rightarrow \dots$$

so müssen wir irgendwann zurück zu dem Element 1 in dieser Kette kommen (es gibt nämlich endlich viele Elemente): so haben wir eine Zyklus gefunden. Wiederhole die ganze Prozedere mit den restlichen Elementen (solange es noch welche gibt). ■

Satz 4.12.2. Jede Permutation $\sigma \in S_n$ lässt sich als Produkt von (*nicht notwendigerweise disjunkten*) Transpositionen schreiben.

Beweis. Es reicht ein Zyklus als Produkt von Transpositionen schreiben (wegen Satz 4.12.1; schreibe σ als Produkt von disjunkten Zyklen). Sei also $\sigma \in S_n$ und

$$(n_1 n_2 \dots n_k) \text{ eine Zyklus von } \sigma.$$

Es gilt

$$(n_1 n_2)(n_1 n_3) \dots (1 n_k) = (n_1 n_2 \dots n_k).$$

Somit ist der Beweis fertig. ■

Beispiel. 1. A C H T U N G : Ein typischer Fehler: $(12)(23) \neq (123)$ sondern $(12)(23) = (132)$!

2. Die Transposition-Zerlegung ist nicht eindeutig:

$$(1342) = (13)(14)(12) \quad \text{und} \quad (1342) = (13)(34)(34)(14)(12).$$

Wir merken aber, dass in diesem Beispiel die Anzahl der Transpositionen in beiden Zerlegungen *ungerade* ist.

Das ist keine Ausnahme. Um dies zu sehen führen wir den folgenden Begriff ein:

Definition. Sei $\sigma \in S_n$. Eine **Inversion** in σ , ist ein Paar $i, j \in \{1, \dots, n\}$ mit $i < j$ und $\sigma(i) > \sigma(j)$. (In diesem Fall sagt man auch, dass die Elemente i und j in σ in falscher (inverser) Reihenfolge, oder in Inversion stehen.)

Beispiel. 1. Für die identische Permutation

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

ist diese Anzahl 0.

2. Für $\sigma = (1342)$, also für

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

gilt: $\sigma(1) > \sigma(2)$, $\sigma(1) > \sigma(4)$, $\sigma(3) > \sigma(4)$. Also diese Anzahl ist 3, ungerade.

Definition. Sei N die Anzahl der Inversionen in σ . So heißt

$$\text{Sign}(\sigma) = (-1)^N,$$

das **Signum** (oder **Vorzeichen**) von σ .

Satz 4.12.3. a) Sind $\pi, \sigma \in S_n$, so gilt $\text{Sign}(\pi \circ \sigma) = \text{Sign}(\pi) \cdot \text{Sign}(\sigma)$.

- b) Ist $\sigma \in S_n$, so ist die Parität der Anzahl der Transpositionen in einer Transposition-Zerlegung von σ gleich die Parität der Anzahl der Inversionen in σ . (Also sie ist immer dieselbe, egal welche Transposition-Zerlegung wir betrachten.)

Beweis. Sei $\sigma \in S_n$. Es reicht zu zeigen, dass, wenn wir σ mit einer Transposition τ multiplizieren, das Vorzeichen mit -1 multipliziert wird: $\text{Sign}(\sigma) = -\text{Sign}(\sigma\tau)$. Sei also $\tau = (i'j')$ eine Transposition mit $i' = \sigma(i)$, $j' = \sigma(j)$, und

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & i & \dots & j & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(i) & \dots & \sigma(j) & \dots & \sigma(n) \end{pmatrix}.$$

Damit ist das Produkt

$$\sigma\tau = \begin{pmatrix} 1 & 2 & \dots & i & \dots & j & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(j) & \dots & \sigma(i) & \dots & \sigma(n) \end{pmatrix}.$$

Wir notieren die Änderung in der Anzahl der Inversionen. War (i, j) eine Inversion von σ , so ist sie keine mehr von $\sigma\tau$; waren sie keine Inversion, so werden sie es. Dies ändert die Anzahl der Inversionen um ± 1 . Ab jetzt OBdA $i < j$. Ist $k < i$ oder $k > j$, so sind i und k (j und k) genau dann in inverser Position in $\sigma\tau$, wenn sie in inverser Position in σ stehen. Sei also $i < k < j$, so gilt

$$\sigma = \begin{pmatrix} 1 & \dots & i & \dots & k & \dots & j & \dots & n \\ \sigma(1) & \dots & \sigma(i) & \dots & \sigma(k) & \dots & \sigma(j) & \dots & \sigma(n) \end{pmatrix}$$

und
$$\sigma\tau = \begin{pmatrix} 1 & \dots & i & \dots & k & \dots & j & \dots & n \\ \sigma(1) & \dots & \sigma(j) & \dots & \sigma(k) & \dots & \sigma(i) & \dots & \sigma(n) \end{pmatrix}.$$

Es gibt drei Möglichkeiten für σ : 1) k steht mit i und j in Inversion 2) k steht mit keinem der beiden (i oder j) in Inversion 3) steht mit *einer* der beiden (i oder j) in Inversion. Fall 1) bedeutet $\sigma(k) < \sigma(i)$, $\sigma(j) < \sigma(k)$, und somit $\sigma\tau(k) = \sigma(k) > \sigma(j) = \sigma\tau(i)$ und $\sigma\tau(k) = \sigma(k) < \sigma(i) = \sigma\tau(j)$. Dabei ändert sich die Anzahl der Inversionen um 2. Fall 2) geht ähnlich, und die Änderung in der Anzahl der Inversionen ist wieder um 2. Fall 3) bedeutet $\sigma(k) < \sigma(i)$, $\sigma(k) < \sigma(j)$ (oder $\sigma(i) < \sigma(k)$, $\sigma(j) < \sigma(k)$), und somit $\sigma\tau(k) = \sigma(k) < \sigma(i) = \sigma\tau(j)$, $\sigma\tau(k) = \sigma(k) < \sigma(j) = \sigma\tau(i)$ (oder $\sigma\tau(k) = \sigma(k) > \sigma(i) = \sigma\tau(j)$, $\sigma\tau(k) = \sigma(k) > \sigma(j) = \sigma\tau(i)$). Dabei ändert sich die Anzahl der Inversionen nicht. ■

Bemerkung. Sei $\sigma \in S_n$ gegeben und geschrieben als Produkt von Zyklen, von denen N gerade Länge haben. So gilt

$$\text{Sign}(\sigma) = (-1)^N.$$

4.13. Permutationsmatrizen und die Determinante.

Sei $\mathcal{E} = \{e_1, \dots, e_n\}$ die kanonische Basis in \mathbb{K}^n , und sei σ eine Permutation. Die Vektoren $e_{\sigma(1)}, \dots, e_{\sigma(n)}$ bilden wieder eine Basis in \mathbb{K}^n , und es gibt eine bijektive lineare Abbildung T_σ mit $T_\sigma(e_j) = e_{\sigma(j)}$. Sei P_σ die Matrix von T_σ bezüglich \mathcal{E} , d.h. $P_\sigma = M_{T_\sigma}^{\mathcal{E}, \mathcal{E}}$. Für das obige Beispiel $\sigma = (12)(3)(4)$ sieht sie so aus:

$$P_\sigma = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Die Matrix P_σ heißt eine **Permutationsmatrix**.

Satz 4.13.1. 1. Eine Matrix P ist genau dann eine Permutationsmatrix, wenn in jeder Spalte und in jeder Zeile genau eine 1 steht, sonst sind die Einträge 0.

2. $S_n \ni \sigma \mapsto P_\sigma \in M_{n,n}$ ist ein Antihomomorphismus von Gruppen, d.h. für $\sigma_1, \sigma_2 \in S_n$ gilt

$$P_{\sigma_2 \circ \sigma_1} = P_{\sigma_1 \sigma_2} = P_{\sigma_2} P_{\sigma_1}.$$

3. Multiplikation mit P_σ^\top von links führt zu Vertauschung der Zeilen von A . Genauer (in Zeilenvektornotation):

$$P_\sigma^\top \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} a_{\sigma(1)} \\ a_{\sigma(2)} \\ \vdots \\ a_{\sigma(n)} \end{pmatrix}.$$

4. Multiplikation mit P_σ von rechts führt zu Vertauschung der Spalten von A . Genauer (in Spaltenvektornotation):

$$\begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix} P_\sigma = \begin{pmatrix} a_{\sigma(1)} & a_{\sigma(2)} & \dots & a_{\sigma(n)} \end{pmatrix}.$$

Beweis. 1. Eine Permutationsmatrix hat offensichtlich in jeder Spalte und Zeile genau eine 1. Sei also P eine Matrix, so dass in jeder Spalte und in jeder Zeile genau eine 1 steht sonst seien die Einträge 0. Jede kanonische Basisvektor e_1, \dots, e_n also taucht als Spaltenvektor von P auf (und genau einmal), in welcher Reihenfolge dies geschieht definiert σ .

2. Für $1 \leq j \leq n$ gilt

$$T_{\sigma_2 \circ \sigma_1} e_j = e_{\sigma_2 \circ \sigma_1(j)} = T_{\sigma_2} e_{\sigma_1(j)} = T_{\sigma_2} T_{\sigma_1} e_j.$$

Daraus folgt $T_{\sigma_2 \circ \sigma_1} = T_{\sigma_2} T_{\sigma_1}$, und somit die Behauptung.

3., 4.: Siehe elementare Umformungen von Typ III. ■

Satz 4.13.2. P_σ lässt sich durch Zeilenvertauschungen auf die Einheitsmatrix bringen. Für eine Permutation σ ist $\text{Det}(P_\sigma) = \text{Sign}(\sigma)$.

Beweis. Klar. ■

Bis jetzt haben wir angenommen, dass die Determinante definiert werden kann, und haben aus den definierenden Eigenschaften ganz viele Information hergeleitet. Nun können wir die Determinante eigentlich definieren.

Definition (Leibniz). Für $A \in M_{n,n}(\mathbb{K})$ setze

$$\text{Det}(A) = \sum_{\sigma \in S_n} \text{Sign}(\sigma) \alpha_{1\sigma(1)} \alpha_{2\sigma(2)} \cdots \alpha_{n\sigma(n)} = \sum_{\sigma \in S_n} \text{Sign}(\sigma) \prod_{i=1}^n \alpha_{i\sigma(i)}.$$

Satz 4.13.3. Es gilt

$$\text{Det}(A) = \sum_{\sigma \in S_n} \text{Sign}(\sigma) \alpha_{1\sigma(1)} \alpha_{2\sigma(2)} \cdots \alpha_{n\sigma(n)} = \sum_{\pi \in S_n} \text{Sign}(\pi) \alpha_{\pi(1)1} \alpha_{\pi(2)2} \cdots \alpha_{\pi(n)n}.$$

Die so definierte Funktion Det hat die Eigenschaften (D1), (D2) und (D3).

Beweis. Es gilt

$$\begin{aligned} \text{Det}(A) &= \sum_{\sigma \in S_n} \text{Sign}(\sigma) \prod_{i=1}^n \alpha_{i\sigma(i)} = \sum_{\sigma \in S_n} \text{Sign}(\sigma) \prod_{i=1}^n \alpha_{\sigma^{-1}\sigma(i)\sigma(i)} \\ &= \sum_{\sigma^{-1} \in S_n} \text{Sign}(\sigma^{-1}) \prod_{i=1}^n \alpha_{\sigma^{-1}(i)i} = \sum_{\sigma \in S_n} \text{Sign}(\sigma) \prod_{i=1}^n \alpha_{\sigma(i)i}. \end{aligned}$$

(D3) ist trivial (setze ein!). Um (D2) zu sehen, sei $\tau \in S_n$ eine Transposition, so ist

$$\begin{aligned} \text{Det}(P_\tau A) &= \sum_{\sigma \in S_n} \text{Sign}(\sigma) \prod_{i=1}^n \alpha_{\sigma(\tau(i))i} = \sum_{\sigma \in S_n} \text{Sign}(\tau^{-1}\sigma) \prod_{i=1}^n \alpha_{\sigma(i)i} \\ &= - \sum_{\sigma \in S_n} \text{Sign}(\sigma) \prod_{i=1}^n \alpha_{\sigma(i)i} = -\text{Det}(A). \end{aligned}$$

(D1) folgt aus Nachrechnen. ■

4.14. Wieder Gleichungssysteme.

Satz 4.14.1 (Die Cramersche Regel). Sei $A \in M_{n,n}(\mathbb{K})$ eine invertierbare Matrix (also $\text{Det}(A) \neq 0$). Für $b \in \mathbb{K}^n$ sei $x \in \mathbb{K}^n$ eine Lösung der Gleichung

$$Ax = b.$$

So gilt

$$x_i = \frac{\text{Det}(A_i)}{\text{Det}(A)},$$

wobei für $i = 1, \dots, n$ ist

$$A_i = \left(\begin{array}{ccc|c|ccc} a_{11} & \dots & a_{1(i-1)} & b_1 & a_{1(i+1)} & \dots & a_{1n} \\ a_{21} & \dots & a_{2(i-1)} & b_2 & a_{2(i+1)} & \dots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n(i-1)} & b_n & a_{n(i+1)} & \dots & a_{nn} \end{array} \right),$$

↑
 i -te Spalte

d.h. $A_i = A$ mit der i -te Spalte durch b ersetzt.

Beweis. Sei x die eindeutige Lösung $x = A^{-1}b$, und betrachte die Matrix

$$B_i = \left(\begin{array}{ccc|c|ccc} 1 & \dots & 0 & x_1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots & \vdots & & \vdots \\ \vdots & & 1 & x_{i-1} & 0 & & \vdots \\ \vdots & & 0 & x_i & 0 & & \vdots \\ \vdots & & 0 & x_{i+1} & 1 & & \vdots \\ \vdots & & \vdots & \vdots & & \ddots & 0 \\ 0 & \dots & 0 & x_n & 0 & \dots & 1 \end{array} \right).$$

↑
 i -te Spalte

Dann gilt

$$AB_i = A_i,$$

und somit

$$\text{Det}(AB_i) = \text{Det}(A) \text{Det}(B_i) = \text{Det}(A_i).$$

Die Determinante von B_i berechnet man mithilfe des Laplace'schen Entwicklungssatzes (nach i -ter Zeile):

$$\text{Det}(B_i) = x_i.$$

Die Behauptung folgt daraus. ■

Diese Methode, um die Lösung zu finden, ist sehr zeitaufwendig, vor Allem wenn man mit der gleichen Matrix A aber mit *vielen* unterschiedlichen rechten Seiten b die Gleichung lösen möchte. Hier ist eine andere Methode, welche genau für solchen Situationen geeignet ist:

Die LU-Zerlegung.

Wir wissen bereits, dass Gleichungen $Ax = b$ mit A obere oder untere Dreiecksmatrizen leicht zu lösen sind: einfach einsetzen!. Die Idee ist also, eine allgemeine Matrix A als Produkt solcher Matrizen zu schreiben: $A = LU$ mit L untere und U obere Dreiecksmatrix. Dies ist leider nicht immer möglich, aber wenn doch, dann kann man $Ax = b$ dadurch lösen, dass man die "leichten" Gleichungen

$$Ly = b \quad \text{und} \quad Ux = y$$

löst.

Beispiel. Sei

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

Wir wollen A als Produkt $A = LU$ zerlegen mit L unterer und U oberer Dreiecksmatrix. und bringen dazu A auf Zeilenstufenform durch elementare Umformungen vom Typ II so, dass vielfach einer Zeile immer nur zu *darunter*liegenden Zeilen addiert werden — sofern dies es machbar ist!. Diese elementaren Umformungen entsprechen Multiplikation von links mit Matrizen von Typ S_{II} . Sei also U die so entstandene Zeilenstufenform von A und $S_N \cdots S_1 A = U$. An diesem Beispiel sieht es so aus:

$$\begin{aligned} S_1 &= \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \rightsquigarrow \text{addiere } -1 \times \text{erste Zeile zu der zweiten;} \\ S_2 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \rightsquigarrow \text{addiere } -1 \times \text{erste Zeile zu der dritten;} \\ S_2 S_1 A &= \begin{pmatrix} 1 & 2 & 3 \\ 0 & -2 & 1 \\ 0 & -1 & 1 \end{pmatrix} \\ S_3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{2} & 1 \end{pmatrix} \rightsquigarrow \text{addiere } -\frac{1}{2} \times \text{zweite Zeile zu der dritten;} \\ S_3 S_2 S_1 A &= \begin{pmatrix} 1 & 2 & 3 \\ 0 & -2 & 1 \\ 0 & 0 & \frac{1}{2} \end{pmatrix} = U \end{aligned}$$

Die Matrix L bekommt man aus $S_3 S_2 S_1 A = U$, also $L = (S_3 S_2 S_1)^{-1}$:

$$L = S_1^{-1} S_2^{-1} S_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1}{2} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & \frac{1}{2} & 1 \end{pmatrix}$$

Wir überprüfen diese Zerlegung:

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

also stimmt!

Satz 4.14.2. a) Sei $A \in M_{n,n}$ eine invertierbare $n \times n$ -Matrix. Für $k = 1, \dots, n$ bezeichne durch A_k die $k \times k$ -Matrix, welche die Elemente der ersten k -Zeilen und der k -Spalten enthält, d.h.

$$A_k = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1k} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \alpha_{k1} & \alpha_{k2} & \dots & \alpha_{kk} \end{pmatrix}$$

(die so genannte **Hauptminoren**). Die Matrix A besitzt genau dann eine LU -Zerlegung, wenn

$$\text{Det}(A_k) \neq 0 \quad \text{für alle } k = 1, \dots, n.$$

b) Sei $A \in M_{n,n}$ eine beliebiger $n \times n$ -Matrix. Dann existieren L, U untere bzw. obere Dreiecksmatrizen und P eine Permutationsmatrix, so dass

$$LU = PA.$$

5. SPEKTRALTHEORIE

Für $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ heisst

$$\text{diag}(\lambda_1, \dots, \lambda_n) := \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

Diagonalmatrix.

Bemerkung. Mit Diagonalmatrizen ist sehr leicht zu rechnen. Zum Beispiel:

$$M_{T^k}^{\mathcal{L}, \mathcal{L}} = (M_T^{\mathcal{L}, \mathcal{L}})^k = \begin{pmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n^k \end{pmatrix}.$$

Falls T invertierbar ist (also $\lambda_i \neq 0$ für $i = 1, \dots, n$)

$$M_{T^{-1}}^{\mathcal{L}, \mathcal{L}} = (M_T^{\mathcal{L}, \mathcal{L}})^{-1} = \begin{pmatrix} \lambda_1^{-1} & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n^{-1} \end{pmatrix}$$

Die Determinante ist auch recht einfach zu bestimmen (siehe Satz 4.8.3):

$$\text{Det}(T) = \text{Det}(M_T^{\mathcal{L}, \mathcal{L}}) = \lambda_1 \lambda_2 \cdots \lambda_n.$$

Beispiel.

$$A := \begin{pmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 10 \end{pmatrix} \quad \text{versus} \quad B := \begin{pmatrix} 6 & -4 & 2 \\ -4 & 6 & 2 \\ 2 & 2 & 0 \end{pmatrix} \quad \text{in der Basis } b_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, b_2 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, b_3 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

Sei V ein endlichdimensionaler Vektorraum über \mathbb{K} ($=\mathbb{R}$ oder \mathbb{C}) und sei $T : V \rightarrow V$ eine lineare Abbildung ($\text{Dim}(V) = n$).

Problem. Finde eine Basis \mathcal{L} so dass

$$A = M_T^{\mathcal{L}, \mathcal{L}}$$

möglichst einfach wird. Oder: ist diese Matrix A bzgl. einer gegebenen Matrix nicht so einfach, finde eine andere Basis \mathcal{L}' , so dass der Wechsel auf \mathcal{L}' eine einfache Matrix liefert:

$$B = S^{-1} M_T^{\mathcal{L}, \mathcal{L}} S, \quad S = M_{I_{n \times n}}^{\mathcal{L}', \mathcal{L}}.$$

5.1. Eigenwerte und Eigenvektoren.

Definition. Sei $T : V \rightarrow V$ linear.

- a) Ein $\lambda \in \mathbb{K}$ heisst **Eigenwert** von T , falls es $x \in V$, $x \neq 0$ mit $Tx = \lambda x$ gibt. Das x heisst dann **Eigenvektor** zum Eigenwert λ .
- b) Für $\lambda \in \mathbb{K}$ setze

$$E_\lambda^T := \{x \in V : Tx = \lambda x\} = \{x \in V : Tx - \lambda x = 0\} = \text{Kern}(T - \lambda I).$$

Ist λ ein Eigenwert, so heisst E_λ^T der **Eigenraum** zum Eigenwert λ . Die Zahl $\text{Dim } E_\lambda^T$ heisst die **geometrische Vielfachheit** des Eigenwertes λ . Bemerke schon, dass E_λ^T ein Unterraum ist.

- c) Das **Spektrum** von T ist die Menge aller Eigenwerte von T

$$\sigma(T) := \{\lambda \in \mathbb{K} : \lambda \text{ ist ein Eigenwert von } T\}.$$

Der nächste Satz soll es erleuchten, warum es sich lohnt Eigenvektoren zu suchen.

Satz 5.1.1. Sei $T : V \rightarrow V$, $\mathcal{L} = \{b_1, \dots, b_n\}$ eine Basis von V . Die folgenden Aussagen sind Äquivalent:

- (i) Die Basis besteht aus Eigenvektoren.
- (ii) Die Matrix von T bzgl. \mathcal{L} hat **Diagonalgestalt** $M_T^{\mathcal{L}, \mathcal{L}} = \text{diag}(\lambda_1, \dots, \lambda_n)$.

In diesem Fall gilt auch $Tb_i = \lambda_i b_i$, d.h. b_i ist Eigenvektor zum Eigenwert λ_i .

Beweis. ■

Definition. Eine lineare Abbildung T heißt **diagonalisierbar**, falls es eine Basis für V aus Eigenvektoren von T gibt.

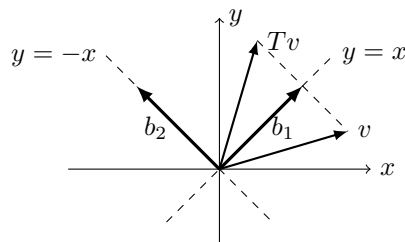
5.2. Beispiele und Gegenbeispiele.

Beispiel. Sei $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ die Spiegelung auf die Gerade $y = x$. So gilt:

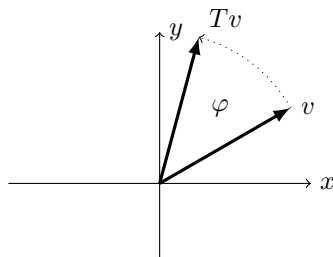
$$\begin{aligned} T \begin{pmatrix} 1 \\ 1 \end{pmatrix} &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{also } b_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ ist Eigenvektor zum Eigenwert } 1, \\ T \begin{pmatrix} 1 \\ -1 \end{pmatrix} &= -1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{also } b_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ ist Eigenvektor zum Eigenwert } -1. \end{aligned}$$

Offensichtlich ist $\mathcal{L} = \{b_1, b_2\}$ eine Basis in \mathbb{R}^2 , also T ist diagonalisierbar und

$$M_T^{\mathcal{L}, \mathcal{L}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$



Beispiel. Sei $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ die Drehung um Winkel φ ($0 < \varphi < \pi$). Diese Abbildung hat keinen Eigenvektor, also keinen Eigenwert.



Beispiel. Betrachte die Abbildung $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, die bzgl. der kanonischen Basis die Matrix

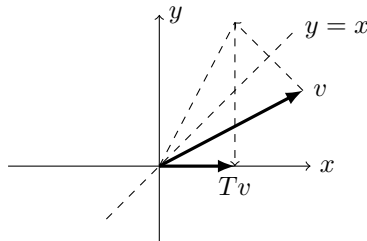
$$M_T^{\mathcal{E}, \mathcal{E}} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{hat.}$$

Wir suchen die Eigenwerte, also die Zahlen $\lambda \in \mathbb{R}$, für die die Gleichung $Tx = \lambda x$ nichttriviale Lösungen hat. Dies bedeutet:

$$\begin{pmatrix} x_2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \end{pmatrix}.$$

Daraus bekommen wir, dass für $\lambda \neq 0$ x_2 und somit x_1 Null sein müssen; und für $\lambda = 0$ muss $x_2 = 0$ gelten und x_1 ist beliebig wählbar. Es gibt keine Basis aus Eigenvektoren. Wir stellen fest: $\sigma(T) = \{0\}$

und T ist nicht diagonalisierbar.



Beispiel. Sei \mathcal{P}_n der Vektorraum der Polynome von Grad kleiner gleich n . Betrachte die lineare Abbildung $T : \mathcal{P}_n \rightarrow \mathcal{P}_n$ mit $Tp = p'$ (die Ableitung). Wir bestimmen nun das Spektrum von T . Sei $\lambda \in \mathbb{C}$ ein Eigenwert mit Eigenvektor $p \in \mathcal{P}_n$. Ist $\lambda \neq 0$, so folgt

$$\text{Grad}(p) = \text{Grad}(\lambda p) = \text{Grad}(Tp) = \text{Grad}(p') < \text{Grad}(p) :$$

ein Widerspruch. Ist $\lambda = 0$, dann bedeutet $p' = Tp = 0p = 0$, dass p ein konstantes Polynom ist. Es gilt also $\sigma(T) = \{0\}$ mit $E_0^T = \{p = a_0 : a_0 \in \mathbb{C}\}$. Die Polynome $1, x, \dots, \frac{x^n}{n!}$ bilden eine Basis \mathcal{L} in \mathcal{P}_n . Bezüglich dieser Basis ist die Matrix von T wie folgt:

$$M_T^{\mathcal{L}, \mathcal{L}} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & 0 & 1 \\ 0 & \dots & 0 & 0 & 0 \end{pmatrix}.$$

5.3. Eigenvektoren.

Satz 5.3.1. Seien $\lambda_1, \dots, \lambda_k$ *verschiedene* Eigenwerte von T mit jeweiligen Eigenvektoren x_1, \dots, x_k . Dann sind die Vektoren x_1, \dots, x_k linear unabhängig.

Beweis. Der Beweis geht mit vollständiger Induktion nach k . Der Fall $k = 1$ ist klar, denn $x_1 \neq 0$. Sei also die Aussage richtig für $k \geq 1$. Falls gilt

$$\alpha_1 x_1 + \dots + \alpha_{k+1} x_{k+1} = 0,$$

dann multipliziere mit λ_{k+1} : $\lambda_{k+1} \alpha_1 x_1 + \dots + \lambda_{k+1} \alpha_{k+1} x_{k+1} = 0.$ (i)

Außerdem $T(\alpha_1 x_1 + \dots + \alpha_{k+1} x_{k+1}) = 0$

und somit $\lambda_1 \alpha_1 x_1 + \dots + \lambda_{k+1} \alpha_{k+1} x_{k+1} = 0.$ (ii)

Wir subtrahieren (i) aus (ii):

$$(\lambda_1 - \lambda_{k+1})\alpha_1 x_1 + \dots + (\lambda_{k+1} - \lambda_{k+1})\alpha_{k+1} x_{k+1} = 0$$

also $(\lambda_1 - \lambda_{k+1})\alpha_1 x_1 + \dots + (\lambda_k - \lambda_{k+1})\alpha_k x_k = 0.$

Die Vektoren x_1, \dots, x_k sind aber linear unabhängig, d.h. die Koeffizienten hier müssen alle 0 sein: $(\lambda_i - \lambda_{k+1})\alpha_i = 0$ für $i = 1, \dots, k$. Da $\lambda_i \neq \lambda_{k+1}$, wir folgern $\alpha_i = 0$ für $i = 1, \dots, k$, und weil $x_{k+1} \neq 0$, folgt auch $\alpha_{k+1} = 0$. Das war zu beweisen. ■

Hilfssatz 5.3.2. Sei $T : V \rightarrow V$ linear mit $\sigma(T) = \{\lambda_1, \dots, \lambda_k\}$. Seien ferner

$$\mathcal{L}_i = \{x_1^i, \dots, x_{n_i}^i\} \subseteq E_{\lambda_i}^T \quad \text{Basen} = \text{Kern}(T - \lambda_i I).$$

Dann sind

$$x_1^1, \dots, x_{n_1}^1, x_1^2, \dots, x_{n_2}^2, \dots, x_1^k, \dots, x_{n_k}^k$$

alle linear unabhängig in V .

Beweis. Nehme an:

$$\sum_{i=1}^k \sum_{j=1}^{n_i} \alpha_j^i x_j^i = 0.$$

Zu zeigen ist $\alpha_j^i = 0$ für alle $i = 1, \dots, n_j$, $j = 1, \dots, k$. Da

$$y_i := \sum_{j=1}^{n_i} \alpha_j^i x_j^i \in E_{\lambda_i}^T,$$

y_i ist ein Eigenvektor, falls $y_i \neq 0$. Wegen Satz 5.3.1 sind die nicht Null y_i linear unanständig, aber

$$\sum_{i=1}^k y_i = 0$$

wäre dann eine nichttriviale Linearkombination solcher λ_i , ein Widerspruch. Somit gilt $y_i = 0$ für alle $i = 1, \dots, n$. Dies impliziert $\alpha_j^i = 0$ für alle $i = 1, \dots, n_j$, $j = 1, \dots, k$. ■

Satz 5.3.3. Für eine lineare Abbildung $T : V \rightarrow V$ mit $\sigma(T) = \{\lambda_1, \dots, \lambda_k\}$ und $n_i = \text{Dim}(E_{\lambda_i}^T)$ sind äquivalent:

- (i) T ist diagonalisierbar.
- (ii) $n = n_1 + \dots + n_k$ ($\text{Dim}(V) = n$).

Beweis. (ii) \Rightarrow (i) Definiere die Vektoren x_j^i wie in Hilfsatz 5.3.2. Wegen der Gültigkeit von $n = n_1 + \dots + n_k$ bilden diese Vektoren eine Basis in V .

(i) \Rightarrow (ii) Hausaufgabe. ■

5.4. Das Charakteristische Polynom. Um Eigenwerte zu bestimmen, gilt es zu untersuchen, für welche $\lambda \in \mathbb{K}$ der Unterraum

$$E_{\lambda}^T = \text{Kern}(T - \lambda I)$$

nicht-trivial ist, d.h. $\neq \{0\}$. Wie in Satz 4.8.6iv) erläutert, hängt die Determinante einer linearen Abbildung nicht von der Wahl der Basis ab.

Satz 5.4.1. Sei $T : V \rightarrow V$ eine lineare Abbildung, $\text{Dim}(V) = n$.

- a) $\lambda \in \mathbb{K}$ ist genau dann ein Eigenwert, wenn $\text{Det}(T - \lambda I) = 0$.
- b) $p_T(\lambda) = \text{Det}(T - \lambda I)$ ist ein Polynom n -ten Grades. Ist A Matrix von T , so gilt

$$p_T(\lambda) = (-1)^n \lambda^n + (-1)^{n-1} \text{Spur}(A) \cdot \lambda^{n-1} + \dots + \text{Det}(A).$$

- c) Für das Spektrum gilt:

$$\sigma(T) = \{\lambda : p_T(\lambda) = 0\}.$$

Beweis. a) $\lambda \in \mathbb{K}$ ist genau dann ein Eigenwert, wenn $E_{\lambda}^T \neq \{0\}$, also wenn $T - \lambda I$ nicht invertierbar ist. Dies ist weiter zu $\text{Det}(T - \lambda I) = 0$ äquivalent (siehe Satz 4.8.4).

- c) Aus der Definition folgt:

$$p_A(\lambda) = \text{Det}(A - \lambda I) = \sum_{\sigma \in S_n} \text{Sign}(\sigma) \beta_{1\sigma(1)} \beta_{2\sigma(2)} \dots \beta_{n\sigma(n)},$$

wobei $B = (\beta_{ij}) = A - \lambda I$. Ein bisschen Überlegen ergibt, dass p_T tatsächlich ein Polynom ist, und wenn man die Koeffizienten von λ^n , λ^{n-1} und den konstanten Term ausrechnet, bekommt man die restlichen Aussagen. ■

Definition. a) Das Polynom p_T heißt das **charakteristische Polynom** von T (oder von A).

b) Ist λ_0 eine Nullstelle k -ter Ordnung von p_T , d.h., $p_T(\lambda) = (\lambda - \lambda_0)^k q(\lambda)$ mit $q(\lambda_0) \neq 0$, so heißt k die **algebraische Vielfachheit** der Eigenwert λ_0 .

c) Die **geometrische Vielfachheit** eines Eigenwertes λ_0 ist

$$\text{Dim}(\text{Kern}(T - \lambda_0 I)).$$

d) Ist $A \in M_{n,n}(\mathbb{K})$, so heißt

$$\text{Spur}(A) = \alpha_{11} + \alpha_{22} + \dots = \alpha_{nn}$$

die Spur von A .

Ist T linear und sind A, B Matrizen von T , so gilt $\text{Spur}(A) = \text{Spur}(B)$. D.h. $\text{Spur}(T) := \text{Spur}(A)$ ist wohldefiniert.

Korollar 5.4.2. a) Ist $\mathbb{K} = \mathbb{C}$, so gibt es mindestens einen Eigenwert. Genauer zerfällt p_T in Linearfaktoren:

$$p_T(\lambda) = (\lambda_1 - \lambda)^{m_1} \dots (\lambda_k - \lambda)^{m_k},$$

mit $\sigma(T) = \{\lambda_1, \dots, \lambda_k\}$ und m_i die algebraische Vielfachheit von λ_i .

b) Falls p_T n verschiedene Nullstellen hat, so ist T diagonalisierbar.

Beweis. a) Ein Polynom n -ten Grades ($n \geq 1$) mit komplexen Koeffizienten hat mindestens eine komplexe Nullstelle. Diese gar nicht triviale Aussage ist der *Hauptsatz der Algebra*.

b) Folgt aus dem Satz 5.3.1. ■

Beispiel. Sei $T = I$ und somit

$$M_{T-\lambda I}^{\mathcal{L}, \mathcal{L}} = \begin{pmatrix} 1-\lambda & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 1-\lambda \end{pmatrix}.$$

Das charakteristische Polynom ist also $p_T(\lambda) = (1-\lambda)^n$. Wir sehen, dass 1 eine n -fache Nullstelle von p_T ist, also von algebraischer Vielfachheit n . Die geometrische Vielfachheit ist auch n . Denn:

$$\text{Dim}(\text{Kern}(T - I)) = \text{Dim}(V) = n.$$

Beispiel. Für

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

ist das charakteristische Polynom

$$p_A(\lambda) = \text{Det} \begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 - 1.$$

Die Nullstellen von p_A sind $+1$ und -1 . Zwei verschiedene Eigenwerte liefern die Diagonalisierbarkeit ($\text{Dim}(\mathbb{R}^2) = 2$).

Beispiel. Für

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

ist das charakteristische Polynom

$$p_A(\lambda) = \text{Det} \begin{pmatrix} -\lambda & 1 \\ 0 & -\lambda \end{pmatrix} = \lambda^2.$$

Das charakteristische Polynom hat nur eine Nullstelle, die Null. Die algebraische Vielfachheit ist 2, die geometrische Vielfachheit ist 1.

Beispiel. Die planare Rotation um den Winkel φ

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

besitzt das charakteristische Polynom

$$p_A(\lambda) = \text{Det} \begin{pmatrix} \cos \varphi - \lambda & \sin \varphi \\ -\sin \varphi & \cos \varphi - \lambda \end{pmatrix} = (\cos \varphi - \lambda)^2 + \sin^2 \varphi = \lambda^2 - 2\lambda \cos \varphi + \cos^2 \varphi + \sin^2 \varphi = \lambda^2 - 2\lambda \cos \varphi + 1.$$

Die Nullstellen sind leicht zu bestimmen:

$$\lambda_{1,2} = \frac{2 \cos \varphi \pm \sqrt{4 \cos^2 \varphi - 4}}{2} = \cos \varphi \pm \sqrt{-\sin^2 \varphi}.$$

Wir sehen also: Im Falle $\mathbb{K} = \mathbb{R}$, für $\varphi \neq k\pi$, $k \in \mathbb{Z}$ hat A keinen Eigenwert (unter $\sqrt{\quad}$ würde eine negative Zahl stehen). Im Falle $\mathbb{K} = \mathbb{C}$, für $\varphi \neq k\pi$, $k \in \mathbb{Z}$, gibt es zwei unterschiedliche Eigenwerte

$$\lambda_{1,2} = \cos \varphi \pm i \sin \varphi.$$

Im letzten Fall ist also A diagonalisierbar über \mathbb{C} (A C H T U N G: Aber nicht über \mathbb{R} !)

Satz 5.4.3. Sei λ_0 ein Eigenwert von $T : V \rightarrow V$. Dann ist die geometrische Vielfachheit von λ_0 kleiner gleich der algebraischen Vielfachheit von λ_0

Beweis. Sei $\mathcal{L} = \{b_1, \dots, b_k\}$ eine Basis in $E_{\lambda_0}^T$ (also die geometrische Vielfachheit ist k). Ergänze diese Basis zu einer Basis \mathcal{M} in V . Die Matrix von T bzgl. \mathcal{M} hat die Form

$$M_T^{\mathcal{M}, \mathcal{M}} = \left(\begin{array}{ccc|c} \lambda_0 & \dots & 0 & * \\ \vdots & \ddots & \vdots & * \\ 0 & \dots & \lambda_0 & * \\ \hline 0 & \dots & 0 & \tilde{A} \end{array} \right),$$

mit \tilde{A} eine $(n-k) \times (n-k)$ -Matrix. Daraus rechnet man das charakteristische Polynom leicht aus:

$$p_T(\lambda) = \text{Det} \left(\begin{array}{ccc|c} \lambda_0 - \lambda & \dots & 0 & * \\ \vdots & \ddots & \vdots & * \\ 0 & \dots & \lambda_0 - \lambda & * \\ \hline 0 & \dots & 0 & \tilde{A} - \lambda I_{n-k \times n-k} \end{array} \right) = (\lambda_0 - \lambda)^k \text{Det}(\tilde{A} - \lambda I_{n-k \times n-k}),$$

daher ist die algebraische Vielfachheit mindestens k . ■

5.5. Diagonalisieren.

Hauptsatz. Für $T : V \rightarrow V$ sind die folgende Aussagen äquivalent.

- (i) T ist diagonalisierbar.
- (ii) Das charakteristische Polynom zerfällt in Linearfaktoren (hat also n Nullstellen), und für jede solche Nullstelle λ_0 (also einen Eigenwert von T) gilt, dass die algebraische und die geometrische Vielfachheit von λ_0 miteinander übereinstimmen.

Beweis. (i) \Rightarrow (ii) Sei \mathcal{L} eine Basis in V aus Eigenvektoren von T . So wird die Matrix

$$M_T^{\mathcal{L}, \mathcal{L}} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix},$$

diagonal, und somit zerfällt das charakteristische Polynom in Linearfaktoren

$$p_T(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda).$$

Sei λ_0 eine Nullstelle von p_T . Taucht λ_0 unter $\lambda_1, \dots, \lambda_n$ genau m -mal auf, so ist die algebraische Vielfachheit von λ_0 eben m . Außerdem, wegen der Diagonalform ist es leicht zu sehen, dass

$$\text{Dim}(\text{Kern}(T - \lambda_0 I)) = m.$$

(ii) \Rightarrow (i) Sei $\sigma(T) = \{\lambda_1, \dots, \lambda_k\}$ und seien n_i bzw. m_i die geometrische und die algebraische Vielfachheiten von λ_i für $i = 1, \dots, k$. Nach Voraussetzung gilt $n_i = m_i$. Außerdem gilt

$$m_1 + m_2 + \dots + m_k = n,$$

denn p_T zerfällt in Linearfaktoren. Daraus

$$n_1 + n_2 + \cdots + n_k = n,$$

und Satz 5.3.3 liefert die Diagonalisierbarkeit von T . ■

Aufgabe. Diagonalisiere eine gegebene lineare Abbildung $T : V \rightarrow V$, $\dim(V) = n$.

Lösung. **1.** Bestimme das charakteristische Polynom p_T .

2. Bestimme die Nullstellen von p_T in \mathbb{K} . item Bestimme für jede Nullstelle λ_0 die geometrische Vielfachheit $\dim E_{\lambda_0}^T = \dim \text{Kern}(T - \lambda_0 I)$.

Falls diese nicht mit der algebraischen Vielfachheit übereinstimmt, so ist T gar nicht diagonalisierbar; andernfalls:

3. Bestimme für jeden Eigenwert $\lambda_0 \in \sigma(T)$ eine Basis in $E_{\lambda_0}^T$. Die Vektoren in dieser Basen ergeben zusammen eine Basis in V bzgl. derer die Matrix von T diagonal wird. ■

Beispiel. Betrachte die Matrix

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

mit Eigenwerten (über \mathbb{C})

$$\lambda_1 = \cos \varphi + i \sin \varphi, \quad \lambda_2 = \cos \varphi - i \sin \varphi.$$

Um einen Eigenvektor zum Eigenwert λ_1 zu bestimmen betrachte

$$A - \lambda_1 I = \begin{pmatrix} -i \sin \varphi & \sin \varphi \\ -\sin \varphi & -i \sin \varphi \end{pmatrix} = \sin \varphi \begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix}.$$

Die Lösung von

$$\sin \varphi \begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

ergibt einen Eigenvektor, zum Beispiel $b_1 = \begin{pmatrix} 1 \\ i \end{pmatrix}$. Einen Eigenvektor zum λ_2 bestimmt man analog: $b_2 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$. ■

5.6. Die Jordan-Normalform.

Definition. Für $\lambda_0 \in \mathbb{K}$ heißt die $k \times k$ -Matrix

$$J_{\lambda_0}^k = \begin{pmatrix} \lambda_0 & 1 & 0 & \dots & 0 \\ 0 & \lambda_0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \lambda_0 & 1 \\ 0 & \dots & \dots & 0 & \lambda_0 \end{pmatrix}$$

eine **Jordan-Matrix**.

Beobachtung. Das Charakteristische Polynom von $J_{\lambda_0}^k$ ist

$$p_{J_{\lambda_0}^k}(\lambda) = \text{Det} \begin{pmatrix} \lambda_0 - \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda_0 - \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \lambda_0 - \lambda & 1 \\ 0 & \dots & \dots & 0 & \lambda_0 - \lambda \end{pmatrix} = (\lambda_0 - \lambda)^k.$$

So sehen wir

$$\sigma(J_{\lambda_0}^k) = \{\lambda_0\},$$

also, λ_0 ist der einzige Eigenwert von $J_{\lambda_0}^k$, und die algebraische Vielfachheit von λ_0 ist k . Wir bestimmen nun $\text{Kern}(J_{\lambda_0}^k - \lambda_0 I)$:

$$J_{\lambda_0}^k - \lambda_0 I = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 0 & 1 \\ 0 & \dots & \dots & 0 & 0 \end{pmatrix}$$

Daher gilt für die Lösungen x der Gleichung

$$\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 0 & 1 \\ 0 & \dots & \dots & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{k-1} \\ x_k \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \\ 0 \end{pmatrix},$$

dass $x_2 = x_3 = \dots = x_k = 0$. Also

$$\text{Kern}(J_{\lambda_0}^k - \lambda_0 I) = \{x \in \mathbb{K}^n : x_2 = x_3 = \dots = x_k = 0\}.$$

Somit ist die geometrische Vielfachheit von λ_0 1. Insbesondere ist $J_{\lambda_0}^k$ nicht diagonalisierbar für $k \geq 2$.

Satz 5.6.1 (Jordansche Normalform). Sei $T : V \rightarrow V$ lineare Abbildung. Falls das Charakteristische Polynom p_T über \mathbb{K} in Linearfaktoren zerfällt, d.h.

$$p_T(\lambda) = (\lambda_1 - \lambda)^{n_1} (\lambda_2 - \lambda)^{n_2} \dots (\lambda_\ell - \lambda)^{n_\ell},$$

so gibt es eine Basis von V , so daß

$$M_T^{\mathcal{L}, \mathcal{L}} = \begin{pmatrix} J_1 & & & & 0 \\ & J_2 & & & \\ & & \ddots & & \\ & & & J_r & \\ 0 & & & & \end{pmatrix},$$

mit J_i , Jordan-Matrizen.

Diese obige Form einer Matrix nennt man **Jordansche Normalform**.

Bemerkung. Die Abbildung T ist genau dann diagonalisierbar, falls alle Jordan-Matrizen in der Jordan-Normalform 1-dimensional sind.

Beispiel. 1. Die folgenden Matrizen sind bereits in Jordan-Normalform:

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \pi \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}$$

2. A', B' sind Jordan-Normalformen der folgenden Matrizen A, B :

$$A := \begin{pmatrix} 0 & -4 & -4 \\ 0 & 0 & -1 \\ 2 & 4 & 6 \end{pmatrix} \quad A' = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \quad B := \begin{pmatrix} 2 & -4 & -2 \\ 0 & 0 & -1 \\ 0 & 4 & 4 \end{pmatrix} \quad B' = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

6. NORMALE ABBILDUNGEN

6.1. Erinnerung. Sei V ein n -dimensionaler prä-Hilbertraum, also ein n -dimensionaler Vektorraum über \mathbb{K} ($=\mathbb{R}$ oder \mathbb{C}) versehen auch mit einem **Skalarprodukt**

$$\langle \cdot, \cdot \rangle \rightarrow \mathbb{K}.$$

Die **euklidische Norm** ist definiert durch $\|x\| = \sqrt{\langle x, x \rangle}$. Zwei Vektoren $x, y \in V$ heißen **orthogonal**, falls $\langle x, y \rangle = 0$ ($x \perp y$). Man sagt, dass die Vektoren e_1, \dots, e_n eine **Orthonormalbasis** in V bilden falls

$$\langle e_i, e_j \rangle = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{sonst} \end{cases}, \quad \text{also falls } \|e_i\| = 1, e_i \perp e_j \quad \text{für } i \neq j.$$

Das **Standardskalarprodukt** in \mathbb{C}^n (oder \mathbb{R}^n) ist definiert durch:

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

Eine Orthonormalbasis ist tatsächlich eine Basis. Die Koordinaten bzgl. einer Orthonormalbasis sind leicht zu bestimmen! Den folgenden Satz haben wir auch schon gesehen, jetzt aber nochmal als Wiederholung:

Satz 6.1.1. Sei e_1, \dots, e_n eine Orthonormalbasis in V .

a) Ist $x \in V$ mit $x = \sum_{j=1}^n x_j e_j$, dann gilt

$$x_i = \langle x, e_i \rangle \quad \text{also} \quad x = \sum_{j=1}^n \langle x, e_j \rangle e_j.$$

b) Zwei Vektoren $x, y \in V$ sind genau dann gleich, wenn für alle $z \in V$ die Gleichheit

$$\langle x, z \rangle = \langle y, z \rangle \quad \text{gilt.}$$

Dies ist weiter äquivalent zu

$$\langle x, e_i \rangle = \langle y, e_i \rangle \quad \text{für alle } i = 1, \dots, n.$$

c) Ist $L : V \rightarrow \mathbb{K}$ eine Linearform, dann existiert genau ein $z \in V$ mit

$$L(x) = \langle x, z \rangle \quad \text{für alle } x \in V.$$

Beweis. a) Es gilt:

$$\langle x, e_i \rangle = \sum_{j=1}^n \langle x_j e_j, e_i \rangle = x_i.$$

b) Seien $x, y \in V$. Es reicht $x = y$ zu zeigen: falls $\langle x, e_i \rangle = \langle y, e_i \rangle$ gilt für alle $i = 1, \dots, n$. Dies folgt aber aus a): Schreibe

$$x = \sum_{j=1}^n x_j e_j \quad \text{und} \quad y = \sum_{j=1}^n y_j e_j.$$

Wegen a) gilt $x_j = \langle x, e_j \rangle$ und $y_j = \langle y, e_j \rangle$, also nach der Voraussetzung folgt $x = y$ auch.

c) Setze $y_i := \overline{L(e_i)}$. Dann gilt

$$L(x) = \sum_{i=1}^n x_i L(e_i) = \sum_{i=1}^n x_i \bar{y}_i = \left\langle \sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j \right\rangle = \langle x, y \rangle. \quad \blacksquare$$

Satz 6.1.2. Sei $T : U \rightarrow V$ eine lineare Abbildung und $\mathcal{L} = \{e_1, \dots, e_n\}$, $\mathcal{L}' = \{f_1, \dots, f_m\}$ Orthonormalbasen in U bzw. in V .

a) Ist $A = M_T^{\mathcal{L}, \mathcal{L}'}$, $A = (\alpha_{ij})$, dann gilt

$$\alpha_{ij} = \langle T e_j, f_i \rangle.$$

b) Für $S : U \rightarrow V$ linear gilt $S = T$ genau dann, wenn

$$\langle Sx, y \rangle = \langle Tx, y \rangle \quad \text{für alle } x, y \in V.$$

Beweis. a) Folgt einfach aus den Definitionen: in der j -ten Spalte von $M_T^{\mathcal{L}, \mathcal{L}'}$ steht der Koordinatenvektor von Te_j . Diese Koordinaten sind eben

$$\alpha_{ij} = \langle Te_i, f_j \rangle.$$

b) Folgt aus dem obigen Satz 6.1.1.b). ■

6.2. Adjungierte einer linearen Abbildung. Der Vektorraum aller linearen Abbildungen von U nach V wird mit $\mathcal{L}(U, V)$ bezeichnet.

Satz 6.2.1. Sei $T \in \mathcal{L}(U, V)$. Es gibt dann genau ein $T^* \in \mathcal{L}(V, U)$ (also eine lineare Abbildung $T^* : V \rightarrow U$) mit

$$\langle Tx, y \rangle = \langle x, T^*y \rangle \quad \text{für alle } x \in U, y \in V.$$

Ist $A = M_T^{\mathcal{L}, \mathcal{L}'}$ mit \mathcal{L} bzw. \mathcal{L}' Orthonormalbasen in U bzw. V , so gilt

$$M_{T^*}^{\mathcal{L}', \mathcal{L}} = \overline{A}^\top =: A^*.$$

Die Abbildung T^* heißt die **adjungierte Abbildung** von T , und die Matrix A^* ist die **Adjungierte der Matrix** A .

Beweis. Für beliebiges $y \in V$ definiere die Linearform $L_y : U \rightarrow \mathbb{K}$, $L_y(x) := \langle Tx, y \rangle$. Wegen Satz 6.1.1c) existiert ein eindeutiger Vektor $z \in U$ mit

$$L_y(x) = \langle x, z \rangle.$$

Setze $T^*y = z$. Aus der Eindeutigkeit bekommt man, dass T^* linear sein muss. Nämlich: sind $y_1, y_2 \in V$ und $\lambda \in \mathbb{K}$, so ist

$$L_{y_1 + \lambda y_2}(x) = \langle Tx, y_1 + \lambda y_2 \rangle = \langle Tx, y_1 \rangle + \lambda \langle Tx, y_2 \rangle = L_{y_1}(x) + \lambda L_{y_2}(x) = \langle x, T^*y_1 \rangle + \langle x, \lambda T^*y_2 \rangle. \quad \blacksquare$$

Beispiel. 1. Sei

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad \text{So gilt } A^* = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = A.$$

2. Sei

$$A = \begin{pmatrix} 0 & i \\ 0 & 0 \end{pmatrix}. \quad \text{So gilt } A^* = \begin{pmatrix} 0 & 0 \\ -i & 0 \end{pmatrix}.$$

3. Sei

$$A = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}. \quad \text{So gilt } A^* = \begin{pmatrix} \cos(-\varphi) & -\sin(-\varphi) \\ \sin(-\varphi) & \cos(-\varphi) \end{pmatrix} = A^{-1}.$$

4. Sei T gegeben durch der Diagonalmatrix

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

(bzgl. einer Orthonormalbasis). Dann ist die Matrix von T^*

$$A^* = \begin{pmatrix} \overline{\lambda_1} & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \overline{\lambda_n} \end{pmatrix}.$$

6.3. Eigenschaften der adjungierten Abbildung und der adjungierten Matrix. Im Folgenden werden wir Satz 6.1.2b) oft benutzen.

0. Für die Identität $I : U \rightarrow U$ gilt

$$I^* = I.$$

Für die Nullabbildung $0 : U \rightarrow V$ gilt

$$0^* : V \rightarrow U \quad \text{die Nullabbildung.}$$

Beweis. Es gelten:

$$\text{und} \quad \langle x, I^*y \rangle = \langle Ix, y \rangle = \langle x, y \rangle, \langle x, 0^*y \rangle = \langle 0x, y \rangle = 0 = \langle x, 0y \rangle.$$

Daraus folgt die Behauptung. ■

1. Für $T : V \rightarrow V$ gilt $T^{**} = T$.

Beweis. Seien $u, v \in V$. Dann gilt:

$$\langle Tu, v \rangle = \langle u, T^*v \rangle = \overline{\langle T^*v, u \rangle} = \overline{\langle v, T^{**}u \rangle} = \langle T^{**}u, v \rangle. \quad \blacksquare$$

2. Für $S, T \in \mathcal{L}(U, V)$, $\lambda \in \mathbb{K}$ gilt

$$(S + \lambda T)^* = S^* + \bar{\lambda}T^*.$$

Beweis. Seien $u, v \in V$. Dann gilt:

$$\langle (S + \lambda T)u, v \rangle = \langle Su + \lambda Tu, v \rangle = \langle u, S^*v \rangle + \lambda \langle u, T^*v \rangle = \langle u, (S^* + \bar{\lambda}T^*)v \rangle. \quad \blacksquare$$

3. Für $S \in \mathcal{L}(U, V)$, $T \in \mathcal{L}(V, W)$ gilt

$$(TS)^* = S^*T^*.$$

Beweis. Seien $u \in U$, $w \in W$. Dann gilt:

$$\langle TSu, w \rangle = \langle Su, T^*w \rangle = \langle u, S^*T^*w \rangle. \quad \blacksquare$$

4. Für $T \in \mathcal{L}(U, V)$ gilt

$$\text{Kern } T = (\text{Bild } T^*)^\perp.$$

Beweis. Sei $u \in \text{Kern } T$, und sei $v \in V$. Dann gilt

$$\langle Tu, v \rangle = \langle u, T^*v \rangle = 0,$$

also $u \in (\text{Bild } T^*)^\perp$. Ist umgekehrt $u \in (\text{Bild } T^*)^\perp$, so gilt

$$0 = \langle u, T^*v \rangle \quad \text{und somit} \quad \langle Tu, v \rangle = 0 \quad \text{für alle } v \in V.$$

Daraus folgt $Tu = 0$, d.h. $u \in \text{Kern } T$. ■

5. Es gilt

$$\text{Rang}(T^*) = \text{Rang}(T).$$

Beweis. Sei A eine Matrix von T bzgl. einer Orthonormalbasis. So ist A^* eine Matrix von T^* . Daraus folgt die Behauptung, denn:

$$\text{Rang}(T) = \text{Rang}(A) = \text{Rang}(A^\top) = \text{Rang}(\overline{A^\top}) = \text{Rang}(T^*). \quad \blacksquare$$

6. Für $T : V \rightarrow V$ gilt

$$\text{Det}(T) = \overline{\text{Det}(T^*)}$$

Beweis. Bemerke zunächst, dass für eine beliebige quadratische Matrix gilt

$$\text{Det}(\overline{B}) = \overline{\text{Det}(B)}.$$

Diese folgt aus der Definition der Determinante (siehe 4.13). Sei also A eine Matrix von T bzgl. einer Orthonormalbasis. Dann gilt

$$\text{Det}(T) = \text{Det}(A) = \text{Det}(A^\top) = \overline{\text{Det}(\overline{A^\top})} = \overline{\text{Det}(A^*)} = \overline{\text{Det}(T^*)}. \quad \blacksquare$$

7. Für $T : V \rightarrow V$ und $\lambda \in \mathbb{K}$ gilt die Äquivalenz:

$$\lambda \in \sigma(T) \iff \bar{\lambda} \in \sigma(T^*).$$

Beweis. Die Behauptung folgt aus den folgenden Äquivalenzen:

$$\lambda \in \sigma(T) \iff \text{Det}(T - \lambda I) = 0 \iff \text{Det}(T^* - \bar{\lambda} I) = 0 \iff \bar{\lambda} \in \sigma(T^*). \quad \blacksquare$$

6.4. Normale Abbildungen.

Definition. Eine Abbildung $T : V \rightarrow V$ heißt **normal**, falls

$$T^*T = TT^* \quad \text{gilt.}$$

Beispiel. 1. Die Abbildung T gegeben durch die Matrix

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

ist *nicht normal*.

2. Die Abbildung T gegeben durch die Matrix

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

ist *normal*.

3. Ist $T^* = T$, oder $T^* = -T$, oder $T^* = T^{-1}$ so ist T normal.

Bemerkung. Sei $A = M_T^{\mathcal{L}, \mathcal{L}}$ die Matrix von T bzgl. einer Orthonormalbasis. Dann ist $M_{T^*}^{\mathcal{L}, \mathcal{L}} = A^*$, und somit ist T genau dann normal, falls

$$A^*A = AA^*$$

gilt. Matrizen mit dieser Eigenschaft heißen auch **normal**.

Satz 6.4.1. Sei $T : V \rightarrow V$ eine lineare Abbildung, so dass eine Orthonormalbasis aus Eigenvektoren von T in V existiert. Dann ist T normal.

Beweis. Sei \mathcal{L} eine Orthonormalbasis in V aus Eigenvektoren von T , dann

$$A := M_T^{\mathcal{L}, \mathcal{L}} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}, \quad \text{und} \quad M_{T^*}^{\mathcal{L}, \mathcal{L}} = \begin{pmatrix} \bar{\lambda}_1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \bar{\lambda}_n \end{pmatrix} = A^*.$$

Daraus folgt

$$AA^* = A^*A \quad \text{und somit} \quad T^*T = TT^*. \quad \blacksquare$$

Satz 6.4.2. Ist T normal und $x \in V$, so gilt

$$\|Tx\| = \|T^*x\|.$$

Beweis. Es gilt

$$\|Tx\|^2 = \langle Tx, Tx \rangle = \langle x, T^*Tx \rangle = \langle x, TT^*x \rangle = \langle T^*x, T^*x \rangle = \|T^*x\|^2. \quad \blacksquare$$

Satz 6.4.3. Sei T normal. Ist $e \in V$ und $\lambda \in \mathbb{K}$ mit

$$Te = \lambda e$$

so gilt

$$T^*e = \bar{\lambda}e,$$

d.h.

$$\text{Kern}(T - \lambda I) = \text{Kern}(T^* - \bar{\lambda} I).$$

Beweis. Bemerke zunächst das Folgende: Ist T normal, so ist für jedes $\lambda \in \mathbb{K}$ auch die Abbildung $T - \lambda I$ normal. Denn es gilt

$$(T - \lambda I)^*(T - \lambda I) = (T^* - \bar{\lambda}I)(T - \lambda I) = T^*T - \lambda T^* - \bar{\lambda}T + \lambda \bar{\lambda}I = (T - \lambda I)(T^* - \bar{\lambda}I) = (T - \lambda I)(T - \lambda I)^*.$$

Für e und λ wie in der Behauptung gilt nun wegen Satz 6.4.2:

$$0 = \|(T - \lambda I)e\| = \|(T^* - \bar{\lambda}I)e\|,$$

also $T^*e = \bar{\lambda}e$. ■

Satz 6.4.4. Sei T normal, und seien $e \in V$, $\lambda \in \mathbb{K}$ mit

$$Te = \lambda e.$$

Setze

$$W := \{e\}^\perp = \{x \in V : \langle x, e \rangle = 0\}.$$

Dann ist W invariant unter T und T^* , d.h. es gelten:

$$T(W) \subseteq W \quad \text{und} \quad T^*(W) \subseteq W.$$

Beweis. Sei $w \in W$, also $\langle w, e \rangle = 0$. Es gilt

$$\langle e, Tw \rangle = \langle T^*e, w \rangle = \bar{\lambda} \langle e, w \rangle = 0, \quad \text{d.h. } Tw \in W.$$

$$\langle e, T^*w \rangle = \langle Te, w \rangle = \lambda \langle e, w \rangle = 0, \quad \text{d.h. } T^*w \in W,$$

wir haben hier Satz 6.4.3 benutzt. ■

Hauptsatz 6.4.5. Sei V ein *komplexer*, endlichdimensionaler (prä-)Hilbertraum. Für $T : V \rightarrow V$ linear sind die folgenden Aussagen äquivalent:

- (i) Es gibt eine Orthonormalbasis in V die aus Eigenvektoren von T besteht.
- (ii) Die Abbildung T ist normal, d.h. $T^*T = T^*T$.

Beweis. Die Implikation “(i) \Rightarrow (ii)” haben wir im Satz 6.4.1 gesehen.

Nun zur Implikation “(ii) \Rightarrow (i)”: Wir beweisen mit vollständiger Induktion nach $n = \text{Dim}(V)$. Der Fall $n = 1$ ist klar: jede Abbildung ist normal und (i) ist auch trivialerweise erfüllt. Sei also die Aussage für $n \geq 1$ für jede normale Abbildung auf W mit Dimension $\text{Dim}(W) \leq n$ richtig. Sei e_{n+1} ein Eigenvektor für $T : V \rightarrow V$, welcher wegen $\mathbb{K} = \mathbb{C}$ existiert (siehe Korollar 5.4.2). Wir können natürlich e_{n+1} umnormieren und somit $\|e_{n+1}\| = 1$ erreichen. Sei $W = \{e_{n+1}\}^\perp$ und setze $S := T|_W$ (T eingeschränkt auf W). So ist, wegen Satz 6.4.4, auch die Abbildung

$$S : W \rightarrow W$$

normal. Da $\text{Dim}(W) = n$ existiert nach Induktionsvoraussetzung eine Orthonormalbasis e_1, \dots, e_n in W aus Eigenvektoren von S (also von T). Diese zusammen mit e_{n+1} haben die gewünschten Eigenschaften. ■

6.5. Unitäre und orthogonale lineare Abbildungen.

Definition. Sei $T : V \rightarrow V$ linear, das das Skalarprodukt erhält, also mit

$$\langle Tu, Tv \rangle = \langle u, v \rangle \quad \text{für alle } u, v \in V.$$

So heißt T

- a) (im Falle $\mathbb{K} = \mathbb{C}$) **unitär**;
- b) (im Falle $\mathbb{K} = \mathbb{R}$) **orthogonal**.

Beispiel. Rotationen und Spiegelungen sind orthogonal, Diagonalmatrizen mit unimodularen Einträgen sind unitär:

$$\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \begin{pmatrix} \exp(i\mu_1) & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \exp(i\mu_n) \end{pmatrix}, \quad \mu_1, \dots, \mu_n \in \mathbb{R}$$

Entscheidend wird der folgende Trick, genannt **Polarisierung**, der das Skalarprodukt mithilfe der Norm (zum Quadrat) ausdrückt:

$$\text{Für } \mathbb{K} = \mathbb{R} \text{ gilt:} \quad \langle x, y \rangle = \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2)$$

$$\text{Für } \mathbb{K} = \mathbb{C} \text{ gilt:} \quad \langle x, y \rangle = \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2 - i\|x - iy\|^2 + i\|x + iy\|^2)$$

(Dies haben wir schon gesehen, siehe 2.10).

Hauptsatz. Sei V komplexer, n -dimensionaler (prä-)Hilbertraum. Für $T \in \mathcal{L}(V, V)$ sind die folgenden Aussagen äquivalent:

- (i) T ist unitär.
- (ii) T ist **isometrisch**, d.h. erhält die Längen von Vektoren, genauer: $\|Tv\| = \|v\|$ gilt für alle $v \in V$.
- (iii) T ist invertierbar mit $T^{-1} = T^*$.
- (iv) T ist normal mit

$$\sigma(T) \subseteq \{\lambda : |\lambda| = 1\}.$$

- (v) T transformiert Orthonormalbasen in Orthonormalbasen.
- (vi) Für *jede* Orthonormalbasis \mathcal{L} in V , bilden die Spalten der Matrix $M_T^{\mathcal{L}, \mathcal{L}}$ eine Orthonormalbasis in \mathbb{C}^n .
- (vii) *Es gibt* ein Orthonormalbasis \mathcal{L} in V , so dass die Spalten der Matrix $M_T^{\mathcal{L}, \mathcal{L}}$ eine Orthonormalbasis in \mathbb{C}^n bilden.

Beweis. (i) \Rightarrow (ii): Falls T das Skalarprodukt erhält, ist es auch isometrisch:

$$\|Tx\|^2 = \langle Tx, Tx \rangle = \langle x, x \rangle = \|x\|^2.$$

(ii) \Rightarrow (iii): Ist T isometrisch, so ist T injektiv, denn $Tx = 0$ impliziert $\|x\| = \|Tx\| = 0$, also $x = 0$. Wegen der Dimensionsformel ist T auch surjektiv, also T ist invertierbar. Wegen Polarisation gilt

$$\langle x, y \rangle = \langle Tx, Ty \rangle = \langle x, T^*Ty \rangle.$$

Aus Satz 6.1.2b) folgt $x = T^*Tx$ für jedes $x \in V$, also $T^* = T^{-1}$.

(iii) \Rightarrow (iv): Es gilt $TT^* = TT^{-1} = I = T^{-1}T = T^*T$, also T ist normal. Nach Satz 6.4.5 gibt es eine Orthonormalbasis \mathcal{L} , so dass $M_T^{\mathcal{L}, \mathcal{L}} = A$ diagonal wird; auf der Diagonale stehen die Eigenwerte $\lambda_1, \dots, \lambda_n$. Die Abbildung T^* hat einerseits die Matrix

$$M_{T^*}^{\mathcal{L}, \mathcal{L}} = A^* = \overline{A}^\top = \overline{A},$$

andererseits

$$M_{T^*}^{\mathcal{L}, \mathcal{L}} = M_{T^{-1}}^{\mathcal{L}, \mathcal{L}} = A^{-1}.$$

Daraus folgt $\overline{\lambda_i} = \lambda_i^{-1}$, also $|\lambda_i|^2 = 1$ und somit die Behauptung.

(iv) \Rightarrow (i): Sei $\mathcal{L} = \{e_1, \dots, e_n\}$ eine Orthonormalbasis aus Eigenvektoren von T . So ist, wegen Satz 6.4.3,

$$\begin{aligned} \langle Tx, Ty \rangle &= \sum_{i=1}^n \langle Tx, e_i \rangle \overline{\langle Ty, e_i \rangle} = \sum_{i=1}^n \langle x, T^*e_i \rangle \overline{\langle y, T^*e_i \rangle} \\ &= \sum_{i=1}^n \lambda_i \langle x, e_i \rangle \cdot \overline{\lambda_i \langle y, e_i \rangle} = \sum_{i=1}^n \langle x, e_i \rangle \overline{\langle y, e_i \rangle} = \langle x, y \rangle. \end{aligned}$$

(i) \Rightarrow (v): Sei $\mathcal{L} = \{e_1, \dots, e_n\}$ eine Orthonormalbasis in V . Dann gilt

$$\langle Te_i, Te_j \rangle = \langle e_i, e_j \rangle = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j, \end{cases}$$

also Te_1, \dots, Te_n ist eine Orthonormalbasis in V .

(v) \Rightarrow (vi): Sei $\mathcal{L} = \{e_1, \dots, e_n\}$ eine Orthonormalbasis in V . Nach Voraussetzung bilden Te_1, \dots, Te_n eine Orthonormalbasis in V , deren Koordinatenvektoren in den Spalten von $M_T^{\mathcal{L}, \mathcal{L}}$ stehen.

(vi) \Rightarrow (vii) ist klar.

(vii) \Rightarrow (iii): Sei \mathcal{L} eine Orthonormalbasis in V , so dass die Spalten von $M_T^{\mathcal{L},\mathcal{L}}$ eine Orthonormalbasis f_1, \dots, f_n in \mathbb{C}^n bilden. In Spaltenvektornotation ist also

$$M_T^{\mathcal{L},\mathcal{L}} = (f_1, f_2, \dots, f_n).$$

Somit gilt in Zeilenvektornotation:

$$M_{T^*}^{\mathcal{L},\mathcal{L}} = \begin{pmatrix} \overline{f_1}^\top \\ \vdots \\ \overline{f_n}^\top \end{pmatrix}.$$

Daraus bekommen wir

$$M_{T^*}^{\mathcal{L},\mathcal{L}} M_T^{\mathcal{L},\mathcal{L}} = \begin{pmatrix} \overline{f_1}^\top \\ \vdots \\ \overline{f_n}^\top \end{pmatrix} (f_1, f_2, \dots, f_n) = \begin{pmatrix} \langle f_1, f_1 \rangle & \langle f_1, f_2 \rangle & \dots & \langle f_1, f_n \rangle \\ \vdots & \vdots & \dots & \vdots \\ \langle f_n, f_1 \rangle & \langle f_n, f_2 \rangle & \dots & \langle f_n, f_n \rangle \end{pmatrix} = I_{n \times n}.$$

D.h. $T^*T = I$ gilt, also (iii) folgt. ■

Bemerkung. Im Falle $\mathbb{K} = \mathbb{R}$ betrachte die Drehung T_φ um Winkel φ in \mathbb{R}^2 mit Matrix

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

Betrachte die folgende Aussage:

(iv') Es gibt eine Orthonormalbasis \mathcal{L} in V , so dass die Matrix von T die folgende Block-Form hat:

$$M_T^{\mathcal{L},\mathcal{L}} = \begin{pmatrix} 1 & \dots & 0 & & & & & & & 0 \\ \vdots & \ddots & \vdots & & & & & & & \vdots \\ 0 & \dots & 1 & & & & & & & \vdots \\ & & & -1 & \dots & 0 & & & & \vdots \\ & & & \vdots & \ddots & \vdots & & & & \vdots \\ & & & 0 & \dots & -1 & & & & \vdots \\ & & & & & & \boxed{T_{\varphi_1}} & \dots & & 0 \\ & & & & & & \vdots & \ddots & & \vdots \\ 0 & & & & & & 0 & \dots & \boxed{T_{\varphi_k}} & \vdots \end{pmatrix},$$

Dann gilt:

T orthogonal \Leftrightarrow (i) \Leftrightarrow (ii) \Leftrightarrow (iii) \Leftrightarrow (iv') \Leftrightarrow (v) \Leftrightarrow (vi) \Leftrightarrow (vii),
wobei $\varphi_i \neq m\pi$ für $m \in \mathbb{Z}$, und wobei in (vi) und in (vii) \mathbb{C}^n durch \mathbb{R}^n ersetzt werden muss.

6.6. Selbstadjungierte Abbildungen.

Definition. Eine lineare Abbildung $T : V \rightarrow V$ heißt

- a) **selbstadjungiert**, falls $T = T^*$.
- b) **schiefadjungiert**, falls $T = -T^*$.

Bemerkung. Nach 6.2.1 ist T genau selbstadjungiert, wenn bzgl. einer Orthonormalbasis \mathcal{L} die Matrix $A := M_T^{\mathcal{L},\mathcal{L}}$ **selbstadjungiert** ist, d.h. sie erfüllt

$$A^* = A.$$

Ist $A^* = -A$, so heißt A **schiefadjungiert**.

Beispiel. 1. Ist $T \in \mathcal{L}(V, V)$ beliebig, so ist $T + T^*$ selbstadjungiert.

2. Ist $T \in \mathcal{L}(V, V)$ beliebig, so ist $T - T^*$ schiefadjungiert.

3. Ist $T \in \mathcal{L}(V, V)$ selbstadjungiert, so ist iT schiefadjungiert.

4. Ist $T \in \mathcal{L}(V, V)$ beliebig, dann gilt

$$T = \frac{T + T^*}{2} + \frac{T - T^*}{2}$$

wobei

$$\frac{T + T^*}{2} \text{ ist selbstadjungiert und } \frac{T - T^*}{2} \text{ ist schiefadjungiert.}$$

5. Ist $T \in \mathcal{L}(V, V)$ beliebig, so sind T^*T und TT^* selbstadjungiert.

Hilfssatz 6.6.1 (Polarisierung). a) Sei $\mathbb{K} = \mathbb{C}$ und seien $S, T \in \mathcal{L}(V, V)$ lineare Abbildungen. Dann gilt:

$$S = T \iff \langle Sx, x \rangle = \langle Tx, x \rangle \text{ für alle } x \in V.$$

b) Ist $\mathbb{K} = \mathbb{R}$ und sind $S, T \in \mathcal{L}(V, V)$ selbstadjungierte lineare Abbildungen, so gilt

$$S = T \iff \langle Sx, x \rangle = \langle Tx, x \rangle \text{ für alle } x \in V.$$

Beweis. a) Die Implikation “ \Rightarrow ” ist trivial. Um “ \Leftarrow ” zu beweisen, reicht es wegen Satz 6.1.2.b) zu zeigen, dass für alle $x, y \in V$

$$\langle Sx, y \rangle = \langle Tx, y \rangle \text{ gilt.}$$

Seien $x, y \in V$ beliebig. Wie bei der Polarisierung können wir schreiben:

$$\begin{aligned} & \frac{1}{4} \left(\langle S(x+y), x+y \rangle - \langle S(x-y), x-y \rangle + i \langle S(x+iy), x+iy \rangle - i \langle S(x-iy), x-iy \rangle \right) \\ &= \frac{1}{4} \left(\langle Sx, x \rangle + \langle Sy, x \rangle + \langle Sx, y \rangle + \langle Sy, y \rangle \right. \\ & \quad - \langle Sx, x \rangle + \langle Sy, x \rangle + \langle Sx, y \rangle - \langle Sy, y \rangle \\ & \quad + i \langle Sx, x \rangle - \langle Sy, x \rangle + \langle Sx, y \rangle + i \langle Sy, y \rangle \\ & \quad \left. - i \langle Sx, x \rangle - \langle Sy, x \rangle + \langle Sx, y \rangle - i \langle Sy, y \rangle \right) = \langle Sx, y \rangle. \end{aligned}$$

Dies ist natürlich auch richtig, wenn S durch T ersetzt wird. Daraus folgt, dass $\langle Tx, y \rangle = \langle Sx, y \rangle$ gilt, falls für alle $x \in V$ die Gleichheit $\langle Tx, x \rangle = \langle Sx, x \rangle$ gilt.

b) geht analog wie a), nur muss man die folgende Identität verwenden:

$$\begin{aligned} & \frac{1}{4} \left(\langle S(x+y), x+y \rangle - \langle S(x-y), x-y \rangle \right) \\ &= \frac{1}{4} \left(\langle Sx, x \rangle + \langle Sx, y \rangle + \langle Sy, x \rangle + \langle Sy, y \rangle - \langle Sx, x \rangle + \langle Sx, y \rangle + \langle Sy, x \rangle - \langle Sy, y \rangle \right) \end{aligned}$$

und, da S selbstadjungiert ist:

$$= \frac{1}{4} \left(\langle Sx, y \rangle + \langle Sx, y \rangle + \langle Sx, y \rangle + \langle Sx, y \rangle \right) = \langle Sx, y \rangle. \quad \blacksquare$$

Bemerkung. Der obige Satz bleibt nicht richtig für $\mathbb{K} = \mathbb{R}$, wenn S oder T nicht selbstadjungiert ist. Betrachte die lineare Abbildungen $T_A, T_B, T_C : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ gegeben durch die Matrizen

$$A := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad B := \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad C := \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}.$$

Dann gelten

$$\begin{aligned} \langle T_A \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \rangle &= \left\langle \begin{pmatrix} x+y \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle = x^2 + xy + y^2, \\ \langle T_B \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \rangle &= \left\langle \begin{pmatrix} x \\ x+y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle = x^2 + xy + y^2, \\ \langle T_C \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \rangle &= \left\langle \begin{pmatrix} x+\frac{1}{2}y \\ y+\frac{1}{2}x \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle = x^2 + xy + y^2. \end{aligned}$$

Die Abbildungen T_A, T_B, T_C sind aber alle unterschiedlich, und T_C ist sogar selbstadjungiert.

Hauptsatz 6.6.2. Sei V ein komplexer Vektorraum. Für $T \in \mathcal{L}(V, V)$ sind äquivalent:

(i) T ist selbstadjungiert.

(ii) T ist normal und

$$\sigma(T) \subseteq \mathbb{R}.$$

(iii) $\langle Tx, x \rangle \in \mathbb{R}$ gilt für alle $x \in V$.

Beweis. (i) \Rightarrow (ii): Ist T selbstadjungiert, so ist es natürlich auch normal: $T^*T = TT = TT^*$. Wegen Satz 6.4.5 gibt es eine Orthonormalbasis in V aus Eigenvektoren von T . Somit gilt

$$A := M_T^{\mathcal{L}, \mathcal{L}} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}, \quad \text{und} \quad A^* = M_{T^*}^{\mathcal{L}, \mathcal{L}} = \begin{pmatrix} \overline{\lambda_1} & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \overline{\lambda_n} \end{pmatrix}.$$

Da $A = A^*$, es gilt $\lambda_i = \overline{\lambda_i}$, d.h. $\lambda_i \in \mathbb{R}$.

(ii) \Rightarrow (i): Falls T normal ist, finden wir eine Orthonormalbasis aus Eigenvektoren von T , so dass $A = M_T^{\mathcal{L}, \mathcal{L}}$ diagonal wird. Da $\sigma(T) \subseteq \mathbb{R}$, gilt $A = A^*$, und somit $T = T^*$.

(i) \Rightarrow (iii): Für $x \in V$ gilt

$$\langle Tx, x \rangle = \langle x, T^*x \rangle = \langle x, Tx \rangle = \overline{\langle Tx, x \rangle},$$

und somit $\langle Tx, x \rangle \in \mathbb{R}$.

(iii) \Rightarrow (i): Nach Voraussetzung gilt

$$\langle x, Tx \rangle = \overline{\langle x, Tx \rangle} = \langle Tx, x \rangle = \langle x, T^*x \rangle,$$

und somit wegen Polarisierung (Lemma 6.6.1a) $T = T^*$. ■

Bemerkung. Der Fall $\mathbb{K} = \mathbb{R}$. Sei \mathcal{L} irgendeine Orthonormalbasis und sei $A = M_T^{\mathcal{L}, \mathcal{L}}$ eine reelle Matrix. Die Abbildung T ist genau dann selbstadjungiert, wenn

$$A = A^\top = A^* \quad \text{also wenn } A \text{ symmetrisch ist.}$$

Nach 6.4.5 existiert eine Orthonormalbasis aus Eigenvektoren von T . Diese Eigenvektoren sind aber nur in \mathbb{C}^n garantiert! Wir wissen bereits, dass die Eigenwerte von T alle reell sind. Daraus bekommen wir für $\lambda \in \sigma(T)$, dass für $x \in \mathbb{C}^n$ mit

$$Ax = \lambda x, \quad \text{gilt auch} \quad A\bar{x} = \overline{Ax} = \overline{\lambda x} = \lambda \bar{x},$$

wobei man benutzt dass A und λ beide reell sind. Daher können wir Eigenvektoren in \mathbb{R}^n finden:

$$A\left(\frac{x + \bar{x}}{2}\right) = \lambda \frac{x + \bar{x}}{2} \quad \text{und} \quad A\left(\frac{x - \bar{x}}{2i}\right) = \lambda \frac{x - \bar{x}}{2i}.$$

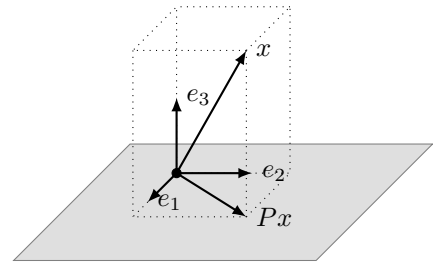
Also falls $x \in \mathbb{C}^n$ ein Eigenvektor zum Eigenwert λ ist, so ist $\operatorname{Re} x \in \mathbb{R}^n$ oder $\operatorname{Im} x \in \mathbb{R}^n$ auch Eigenvektor zum λ (einer der beiden ist nicht 0). Daraus kann man induktiv eine Orthonormalbasis aus Eigenvektoren von T in \mathbb{R}^n finden.

Wir haben also den folgenden Satz bewiesen.

Satz. Ist V ein endlichdimensionaler \mathbb{R} -Vektorraum und $T \in \mathcal{L}(V, V)$ selbstadjungiert, so ist T auch über \mathbb{R} diagonalisierbar.

Beispiel. Sei $U \subseteq V$ ein Unterraum, und sei e_1, \dots, e_m Orthonormalbasis in U , ergänze sie zu einer Orthonormalbasis e_1, \dots, e_n in V . Die **orthogonale Projektion** P von V auf U ist definiert durch

$$P : V \rightarrow V, \quad Px = \sum_{i=1}^m \langle x, e_i \rangle e_i \in U.$$



Es gilt

$$\text{Bild}(P) = U \quad \text{und} \quad \text{Kern}(P) = U^\perp = \{x \in V : x \perp U\}.$$

Die Abbildung P ist selbstadjungiert. Denn:

$$\langle Px, x \rangle = \left\langle \sum_{i=1}^m \langle x, e_i \rangle e_i, \sum_{j=1}^n \langle x, e_j \rangle e_j \right\rangle = \sum_{i=1}^m \langle x, e_i \rangle \overline{\langle x, e_i \rangle} = \sum_{i=1}^m |\langle x, e_i \rangle|^2 \geq 0.$$

6.7. Positiv semidefinite Abbildungen. Bei reellen Zahlen unterscheidet man positive, nicht-negative, negative etc. Im Komplexen gelingt dies nicht immer: Sollen wir $1 - i$ als positiv oder als negativ ansehen? Eine (reelle) Zahl t ist nicht-negativ genau dann, wenn $t = \bar{s} \cdot s$ gilt für ein $s \in \mathbb{C}$. Eine ähnliche Eigenschaft gilt auch für lineare Abbildungen.

Beispiel. Welche der folgenden linearen Abbildungen (auf \mathbb{R}^2 bezüglich der Standardbasis) würde man sinnvollerweise *positiv* nennen?

$$A = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \sim \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \quad B = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}$$

Satz. Sei V komplexer Vektorraum. Für $T : V \rightarrow V$ sind äquivalent.

- (i) $\langle Tx, x \rangle \geq 0$ [oder > 0] für alle $x \in V, x \neq 0$.
- (ii) T ist selbstadjungiert und

$$\sigma(T) \subseteq \mathbb{R}_+ \quad [\text{oder } \mathbb{R}_+ \setminus \{0\}].$$

- (iii) Es gibt $S : V \rightarrow V$ mit $T = S^*S$ [und S ist invertierbar].

Beweis. (i) \Rightarrow (ii): Da $\langle Tx, x \rangle \geq 0$ bedeutet auch, dass $\langle Tx, x \rangle \in \mathbb{R}$. Satz 6.6.2 liefert, dass T selbstadjungiert ist. Sei $\lambda \in \sigma(T)$ mit einem Eigenvektor $x \in V, \|x\| = 1$. Dann gilt

$$\lambda = \lambda \|x\|^2 = \langle \lambda x, x \rangle = \langle Tx, x \rangle \geq 0,$$

d.h. $\lambda \geq 0$ muss gelten. Dies zeigt die gewünschte Implikation. Gilt in (i) zusätzlich > 0 für $x \neq 0$. So kann $\lambda = 0$ kein Eigenwert sein.

(ii) \Rightarrow (iii): Sei \mathcal{L} eine Orthonormalbasis in V aus Eigenvektoren von T . Dann gilt

$$A = M_T^{\mathcal{L}, \mathcal{L}} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \quad \text{mit } \lambda_i \geq 0.$$

Setze

$$B := \begin{pmatrix} \sqrt{\lambda_1} & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \sqrt{\lambda_n} \end{pmatrix}.$$

Dann gilt $B = B^*$ und $B^2 = BB = A$. Die Abbildung S gegeben durch diese Matrix bzgl. der Basis \mathcal{L} ist selbstadjungiert und erfüllt

$$S^*S = SS = T.$$

Der Zusatz: Ist $\sigma(T) \subseteq \mathbb{R} \setminus \{0\}$, so gilt $\lambda_i > 0$, also S und somit B ist invertierbar.

(iii) \Rightarrow (i): Für $x \in V$ gilt

$$\langle Tx, x \rangle = \langle S^* Sx, x \rangle = \langle Sx, Sx \rangle \geq 0.$$

Dies beweist "(iii) \Rightarrow (i)". Ist zusätzlich S invertierbar, so kann hier $\|Sx\|^2 = \langle Sx, Sx \rangle = 0$ nur dann stehen, wenn $x = 0$. ■

Definition. Eine lineare Abbildung $T : V \rightarrow V$ heißt

- a) **positiv semidefinit**, falls $\langle Tx, x \rangle \geq 0$ für alle $x \in V$.
- b) **positiv definit**, falls $\langle Tx, x \rangle > 0$ für alle $x \in V, x \neq 0$.
- c) **negativ (semi)definit**, falls $-T$ positiv (semi)definit ist.
- d) **indefinit**, falls keiner der obigen gilt.

Wir sagen, dass eine Matrix $A \in M_{n,n}$ eine dieser Eigenschaften hat, falls die lineare Abbildung $T_A : \mathbb{C}^n \rightarrow \mathbb{C}^n, T_A x = Ax$ die jeweilige Eigenschaft besitzt.

Satz 6.7.1 (Sylvester). Sei A eine $n \times n$ -Matrix. Bezeichne durch A_k die $k \times k$ -Matrix, welche die Elemente der ersten k -Zeilen und der k -Spalten enthält, d.h.

$$A_k = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1k} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \alpha_{k1} & \alpha_{k2} & \dots & \alpha_{kk} \end{pmatrix}$$

(die so genannte **Hauptminoren**). Genau dann ist A positiv definit, wenn

$$\text{Det}(A_k) > 0 \quad \text{für alle } k = 1, \dots, n \text{ gilt.}$$

Beweis. Ohne Beweis. ■

