Lineare Algebra 1

Anton Deitmar, WS 2024/25

Inhaltsverzeichnis

1 Grundlagen

1.1 Mengen und Abbildungen

Definition 1.1.1. Eine **Menge** ist eine (gedachte) Zusammenfassung von (wirklichen oder gedachten) Objekten, die die **Elemente** der Menge genannt werden.

Beispiele 1.1.2. (a) Die Leere Menge $\{\} = \emptyset$.

(b) Die Menge $\{1,2,3\}$ der Zahlen 1,2,3.

(c) Die Menge $\mathbb{N} = \{1,2,3,\dots\}$ der **natürlichen Zahlen**.

Definition 1.1.3. Zwei Menge heißen **gleich**, falls sie die gleichen Elemente haben:

$$N = M \quad \Leftrightarrow \quad (x \in N \;\Leftrightarrow\; x \in M).$$

Eine Menge A ist **Teilmenge** einer Menge B, falls jedes Element von A schon in B liegt, also

$$A \subset B \quad \Leftrightarrow \quad (x \in A \;\Rightarrow\; x \in B).$$

Definition 1.1.4. Seien A und B Mengen. Der **Durchschnitt** $A \cap B$ von A und B ist die Menge der gemeinsamen Elemente:

$$x \in A \cap B \quad \Leftrightarrow \quad \big(x \in A \text{ und } x \in B\big),$$

oder, was das Gleiche bedeutet:

$$A \cap B = \big\{x \in A : x \in B\big\}.$$

Zwei Menge A und B heißen **disjunkt**, falls $A \cap B = \emptyset$, sie also keine gemeinsamen Elemente haben.

Definition 1.1.5. Die **Vereinigung** zweier Mengen $A \cup B$ ist die Menge aller x, die in mindestens einer der beiden Mengen liegen, also

$$x \in (A \cup B) \quad \Leftrightarrow \quad \big(x \in A \text{ oder } x \in B\big).$$

Definition 1.1.6. Die Menge der **natürlichen Zahlen**

$$\mathbb{N} = \big\{1,2,3,\dots\big\}$$

erweitert man zur Menge der **ganzen Zahlen**

$$\mathbb{Z} = \big\{\dots,-2,-1,0,1,2,\dots\big\}$$

und diese zur Menge der **rationalen Zahlen**

$$\mathbb{Q} = \left\{ \frac{p}{q} : p, q \in \mathbb{Z},\ q > 0 \right\}.$$

Wir haben also

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \underbrace{\mathbb{R}}_{\text{die reellen Zahlen}} \subset \underbrace{\mathbb{C}}_{\text{die komplexen Zahlen}}$$

Definition 1.1.7. Sind X und Y Mengen, so schreibt man

$$X \smallsetminus Y = \left\{ x \in X : x \notin Y \right\}.$$

Beispiele 1.1.8. (a) $\mathbb{Z} \smallsetminus \mathbb{N} = \left\{ \dots, -2, -1, 0 \right\}$.

(b) $\left\{1,2\right\} \smallsetminus \left\{2,3\right\} = \left\{1\right\}$.

Definition 1.1.9. Das **Produkt** oder auch das **kartesische Produkt** $X \times Y$ zweier Mengen X und Y ist definiert als die Menge aller Paare (x, y) wobei $x \in X$ in $y \in Y$ ist.

Beispiele 1.1.10. (a) $\left\{1,2\right\} \times \left\{3,4,5\right\} = \left\{(1,3),(1,4),(1,5),(2,3),(2,4),(2,5)\right\}$.

(b) Die Menge $\mathbb{R} \times \mathbb{R}$ kann als die Ebene verstanden werden.

Definition 1.1.11. Sind X, Y Mengen, so ist eine **Abbildung** $f : X \to Y$ eine Zuordnung, die jedem Element x von X genau ein Element $y = f(x)$ von Y zuordnet.

Beispiele 1.1.12. (a) $f : \left\{1,2,3\right\} \to \left\{4,5\right\}$

$$f(1) = 4, \quad f(2) = 4 \quad f(3) = 5.$$

(b) $f : \mathbb{R} \to \mathbb{R}$

$$f(x) = x^2.$$

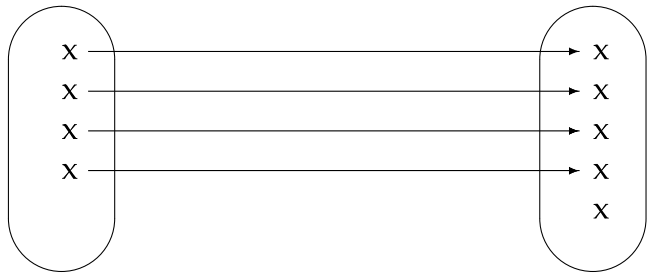
(c) Auf jeder gegebenen Menge X gibt es eine kanonische Abbildung $X \to X$, nämlich die **Identität**

$$\mathrm{Id}_X : X \to X; \quad x \mapsto x.$$

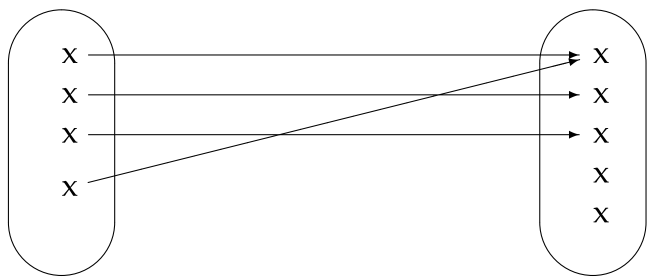
Definition 1.1.13. Eine Abbildung $f : X \to Y$ heißt **injektiv**, falls **verschiedene Elemente verschiedene Bilder** haben, also wenn für alle $x, x' \in X$ gilt

$$x \neq x' \ \Rightarrow\ f(x) \neq f(x').$$

injektiv:



nicht injektiv:

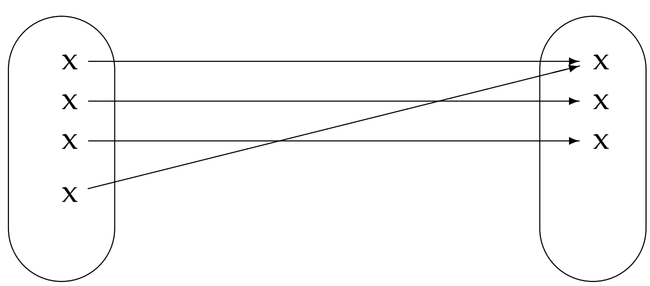


Beispiele 1.1.14. (a) Die Abbildung $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^2$ ist nicht injektiv, denn $f(1) = f(-1) = 1$.

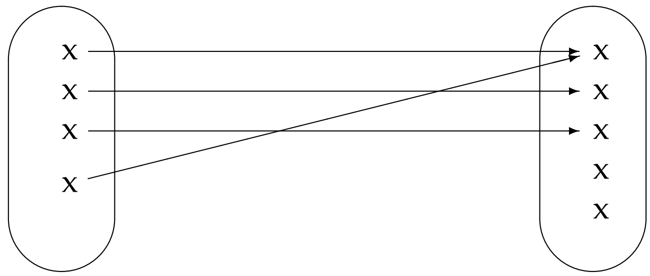
(b) Die Abbildung $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^3$ ist injektiv.

Definition 1.1.15. Eine Abbildung $f : X \to Y$ heißt **surjektiv**, falls es zu jedem $y \in Y$ ein $x \in X$ gibt mit $f(x) = y$, wenn also jedes $y \in Y$ getroffen wird.

surjektiv:



nicht surjektiv:



Beispiele 1.1.16. (a) Die Abbildung $F : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^2$ ist nicht surjektiv, da es kein $x \in \mathbb{R}$ gibt mit $f(x) = -1$.

(b) Die Abbildung $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^3$ ist surjektiv (Analysis).

Definition 1.1.17. Eine Abbildung $f : X \to Y$ die sowohl injektiv als auch surjektiv ist, heißt **bijektiv**.

Satz 1.1.18. *Ist X eine* **endliche** *Menge und ist $f : X \to X$ eine Selbstabbildung, so sind äquivalent:*

(a) *f ist injektiv,*

(b) *f ist surjektiv,*

(c) *f ist bijektiv.*

Beweis. Sei $X = \{x_1, \dots, x_n\}$ eine Menge mit n Elementen.

(a)⇒(b) Da $f(x_1), \dots, f(x_n)$ alle verschieden sind, sind es wieder n Elemente, also $\{f(x_1), \dots, f(x_n)\} = X$, damit ist f surjektiv.

(b)⇒(c) Sei f surjektiv. Wir haben zu zeigen, dass f injektiv ist. Wegen $\{f(x_1), \dots, f(x_n)\} = \{x_1, \dots, x_n\}$ müssen die $f(x_1), \dots, f(x_n)$ paarweise verschieden sein, also ist f injektiv.

Die Richtung (c)⇒(a) ist trivial. □

Definition 1.1.19. Eine **Permutation** auf einer Menge M ist eine bijektive Abbildung $M \to M$. Mit $\mathrm{Per}(M)$ bezeichnen wir die Menge aller Permutationen auf M. Ist $n = 1, 2, 3, \dots$ eine natürliche Zahl, so bezeichnet $\mathrm{Per}(n)$ die Menge der Permutationen auf der Menge $\{1, 2, 3, \dots, n\}$.

Die Menge Per(1) hat ein Element. Die Menge Per(2) hat zwei Elemente, die Menge Per(3) hat sechs Elemente:

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

Die Schreibweise bedeutet, dass jedes Element der ersten Zeile auf das darunterliegende abgebildet wird.

Definition 1.1.20. Seien $f : X \to Y$ und $g : Y \to Z$ Abbildungen, so kann man ihre **Komposition** $g \circ f$ (lies: "geh nach eff") bilden, dies ist eine Abbildung von X nach Z definiert durch

$$g \circ f(x) = g(f(x)).$$

Man visualisiert dies durch ein Diagramm:

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[dr, "g\circ f"'] & Y \arrow[d, "g"] \\ & Z. \end{tikzcd}$$

Beispiel 1.1.21. Ist $f : X \to Y$, dann gilt $f \circ \mathrm{Id}_X = f$ und $\mathrm{Id}_Y \circ f = f$.

Lemma 1.1.22 (Assoziativität der Komposition). *Seien $f : W \to X$, $g : X \to Y$ und $h : Y \to Z$ Abbildungen. Dann gilt*

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

Beweis. Sei $w \in W$, dann gilt

$$h \circ (g \circ f)(w) = h\big(g \circ f(w)\big) = h\big(g(f(w))\big) = (h \circ g)\big(f(w)\big) = (h \circ g) \circ f(w). \qquad \square$$

Proposition 1.1.23. *Sind g und f beide injektiv, so ist $g \circ f$ injektiv. Ebenso für surjektiv.*

Beweis. Es seien beide injektiv und $x \neq x'$ in X. Da f injektiv ist, folgt $f(x) \neq f(x')$. Da g injektiv ist, folgt hieraus $g \circ f(x) = g(f(x)) \neq g(f(x')) = g \circ f(x')$. Also ist $g \circ f$ injektiv.

Es seien beide surjektiv. Sei $z \in Z$. Da g surjektiv ist, gibt es ein $y \in Y$ mit $g(y) = z$. Da f surjektiv ist, gibt es ein $x \in X$ mit $f(x) = y$. Dann folgt $g \circ f(x) = g(f(x)) = g(y) = z$. Also ist $g \circ f$ surjektiv. $\square$

Insbesondere ist die Komposition bijektiver Abbildungen bijektiv.

Definition 1.1.24. Eine Abbildung $f : X \to Y$ heißt **invertierbar**, wenn es eine Abbildung $g : Y \to X$ gibt mit

$$g \circ f = \mathrm{Id}_X \quad \text{und} \quad f \circ g = \mathrm{Id}_Y.$$

Bemerkung 1.1.25. (a) Sei f eine invertierbare Abbildung. Dann ist die Abbildung g mit $f \circ g = \mathrm{Id}_Y$ und $g \circ f = \mathrm{Id}_X$ eindeutig bestimmt. Wir nennen sie die **inverse Abbildung** oder **Umkehrabbildung** und schreiben $g = f^{-1}$.

(b) Eine Abbildung $f : X \to Y$ ist genau dann invertierbar, wenn sie bijektiv ist. (Uebungsaufgabe)

1.2 Vollständige Induktion

Ist $A(n)$ eine Aussage, die für eine natürliche Zahl n Sinn macht, so gilt

$$\left.\begin{array}{c} A(1) \text{ gilt,} \\ \\ A(n) \Rightarrow A(n+1) \\ \text{für jedes } n \in \mathbb{N} \end{array}\right\} \Rightarrow A(n) \text{ gilt für jedes } n \in \mathbb{N}.$$

Beispiel 1.2.1. Für jede natürliche Zahl n gilt

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}.$$

Beweis. Induktionsanfang: Für $n = 1$ ist die Behauptung wahr, denn die linke Seite ist 1 und die rechte Seite ist $\frac{1(1+1)}{2} = \frac{2}{2} = 1$.

Induktionsschritt: $n \to n+1$ Wir nehmen an, die Behauptung $A(n)$ gilt für n. Wir rechnen dann

$$\begin{aligned} \underbrace{1 + 2 + \cdots + n}_{=\frac{n(n+1)}{2}} + (n+1) &= \frac{n(n+1)}{2} + n + 1 \\ &= \frac{n^2 + n + 2n + 2}{2} = \frac{(n+1)(n+2)}{2}. \end{aligned}$$

Also haben wir gezeigt, dass aus der Behauptung für n folgt:

$$1 + 2 + \cdots + n + (n+1) = \frac{(n+1)(n+2)}{2}.$$

Dies ist aber gerade $A(n+1)$. □

Es kann auch passieren, dass eine Aussage $A(n)$ erst ab einer bestimmten Zahl richtig ist. Wir zeigen dass an folgendem

Beispiel 1.2.2. Für jede natürliche Zahl $n \geq 5$ gilt $2^n > n^2$.

Beweis. Induktionsanfang: $n = 5$. In diesem Fall rechnen wir

$$2^n = 2^5 = 32 > 25 = n^2.$$

Induktionsschritt: $n \to n+1$. Es gelte $2^n > n^2$ und es sei $n \geq 5$. Wir rechnen

$$2^{n+1} = 2 \cdot 2^n > 2 \cdot n^2 = n^2 + n^2 = n^2 + \underbrace{(n-1)}_{\geq 2} n + n \geq n^2 + 2n + 1 = (n+1)^2,$$

dies ist gerade die Behauptung für $n+1$. □

Es kommt auch vor, dass man im Induktionsschritt nicht nur $A(n)$ als Voraussetzung braucht, sondern dass der Induktionsschritt lautet: $A(1), \dots A(n) \Rightarrow A(n+1)$.

Beispiel 1.2.3. Jede natürliche Zahl $n \geq 2$ ist ein Produkt von Primzahlen, lässt sich also schreiben als

$$n = p_1 \cdot p_2 \cdots p_k,$$

wobei die Zahlen $p_1, \dots, p_k$ Primzahlen sind.

Beweis. Induktionsanfang: $n = 2$. Diese Zahl ist allerdings selbst eine Primzahl, also insbesondere ein Produkt von solchen.

Induktionsschritt: $2, \dots, n \to n+1$. Die Behauptung sei wahr für $2, \dots, n$. Ist nun $n+1$ selbst eine Primzahl, so gilt die Behauptung auch für $n+1$. Andernfalls hat n einen echten Teiler, sagen wir d. Dann gilt $d, \frac{n+1}{d} \in \{2, 3, \dots, n\}$, damit gilt also die Behauptung für d und für $\frac{n+1}{d}$, welche damit beide Produkte von Primzahlen sind, also ist auch

$$n + 1 = d \cdot \frac{n+1}{d}$$

ein Produkt von Primzahlen. □

Proposition 1.2.4. *Die Anzahl aller möglichen Anordnungen einer n-elementigen Menge ist* $n!$.

Beispiele: Alle Anordnungen der 2-elementigen Menge $\{1, 2\}$ sind

$$(1,2), \quad (2,1)$$

also 2 Stück. Alle Anordnungen der 3-elementigen Menge $\{1,2,3\}$ sind

$$(1,2,3),(1,3,2),(2,1,3),(3,1,2),(2,3,1),(3,2,1),$$

also $6 = 3 \cdot 2 \cdot 1 = 3!$.

Beweis. Wir wollen n-Elemente anordnen. Für die erste Position haben wir die freie Wahl unter allen n-Elementen. Haben wir die erste Position besetzt, so bleibt uns noch die Wahl unter $n-1$ Elementen, für die ersten beiden Positionen haben wir also $n \cdot (n-1)$ verschiedenen Wahlen. Für die dritte Position bleiben $(n-2)$ viele Wahlmöglichkeiten und so weiter, bis am Ende die letzte Position vorgegeben ist, so dass wir im Endeffekt $n \cdot (n-1) \cdot (n-2) \cdots 1 = n!$ viele Wahlmöglichkeiten haben. □

Es seien zwei ganze Zahlen $k, n \geq 0$ gegeben. Wir definieren den **Binomialkoeffizienten** $\binom{n}{k}$ durch

$$\binom{n}{k} = \begin{array}{c}\text{Anzahl der } k\text{-elementigen Teilmengen} \\ \text{einer } n\text{-elementigen Menge.}\end{array}$$

Offensichtlich ist dies gleich Null, falls $n < k$.

Beispiele 1.2.5. (a) Ist $k = 0$ und n beliebig, so ist $\binom{n}{k} = 1$, denn man fischt immer nur die leere Menge.

(b) Ist $k = 1$, und $n \geq 1$ beliebig, so ist $\binom{n}{k} = n$, denn die Menge $\{1, \dots, n\}$ hat genau die 1-elementigen Teilmengen $\{1\}, \{2\}, \dots, \{n\}$.

(c) Es gilt $\binom{n}{2} = \frac{n(n-1)}{2}$, denn, suchen wir eine 2-elementige Teilmenge, so haben wir für das erste Element n-Wahlmöglichkeiten und für das zweite $n-1$. Dann haben wir aber jede Teilmenge doppelt gezählt, weil ja beide mögliche Reihenfolgen auftreten.

Proposition 1.2.6. (a) *Für $0 \leq k \leq n$ gilt*

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n \cdot (n-1) \cdots (n-k+1)}{k!}.$$

(b) *Es gilt $\binom{n}{0} = \binom{n}{n} = 1$ und*

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}.$$

Beweis. (a) Wir wählen k Elemente aus n aus. Wählen wir zunächst mit Reihenfolge. Für das erste Element haben wir n Wahlmöglichkeiten, für das zweite $n-1$ und so weiter, so dass wir insgesamt $n \cdot (n-1) \cdots (n-k+1)$ Möglichkeiten haben. Jetzt treten aber alle möglichen Reihenfolgen der k Elemente auf, also müssen wir das Ergebnis noch durch $k!$ dividieren.

(b) Sei I eine $(n+1)$-elementige Menge und sei $x_0 \in I$ fest gewählt. Unter den k-elementigen Mengen gibt es $\binom{n}{k-1}$ viele, die x_0 enthalten und $\binom{n}{k}$ viele, die x_0 nicht enthalten. □

Satz 1.2.7 (Binomischer Lehrsatz)**.** *Für jedes $n \geq 0$ gilt*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

Insbesondere für $y = 1$ erhalten wir

$$(x+1)^n = \sum_{k=0}^{n} \binom{n}{k} x^k.$$

Dieser Satz ist der Grund, warum die Zahlen $\binom{n}{k}$ die **Binomialkoeffizienten** heißen.

Beweis. Induktion nach n. Für $n = 0$ ist nichts zu zeigen.

Induktionsschritt $n \to n+1$: Wir rechnen

$$\begin{aligned}
(x+y)^{n+1} = (x+y)(x+y)^n &= (x+y) \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} \\
&= \sum_{k=0}^{n} \binom{n}{k} x^{k+1} y^{n-k} + \sum_{k=0}^{n} \binom{n}{k} x^k y^{n+1-k} \\
&= \sum_{k=1}^{n+1} \binom{n}{k-1} x^k y^{n+1-k} + \sum_{k=0}^{n} \binom{n}{k} x^k y^{n+1-k} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} x^k y^{n+1-k}.
\end{aligned}$$

Damit folgt die Behauptung. □

1.3 Gruppen

Sei M eine Menge. Eine **Verknüpfung** auf M ist eine Abbildung:

$$M \times M \to M.$$

Beispiele 1.3.1. (a) Addition und Multiplikation auf $\mathbb{N}$ sind Verknüpfungen:

$$\mathbb{N} \times \mathbb{N} \to \mathbb{N}, \qquad (m,n) \Mapsto m+n;$$

sowie

$$\mathbb{N} \times \mathbb{N} \to \mathbb{N}, \qquad (m,n) \mapsto mn.$$

(b) Hintereinanderausführung von Permutationen:

$$\mathrm{Per}(M) \times \mathrm{Per}(M) \to \mathrm{Per}(M), \quad (\alpha,\beta) \mapsto \alpha \circ \beta$$

ist eine Verknüpfung.

(c) Sei $M = \mathbb{Q}, \mathbb{R}$, so definiert das **arithmetische Mittel**

$$a * b = \frac{a+b}{2}$$

eine Verknüpfung auf M.

Definition 1.3.2. Eine **Gruppe** ist eine Menge G mit einer Verknüpfung $G \times G \to G$, geschrieben als $(a,b) \mapsto ab$, so dass für alle $a,b,c \in G$ folgende Axiome erfüllt sind:

- **Assoziativgesetz**

$$a(bc) = (ab)c,$$

- **neutrales Element**: es gibt ein $e \in G$ mit

$$ae = a,$$

- **inverses Element**: zu jedem $a \in G$ gibt es ein $a' \in G$ mit

$$aa' = e.$$

Beispiele 1.3.3. (a) Die Menge $\mathbb{N}$ ist mit der Addition keine Gruppe, denn sie enthält kein neutrales Element.

(b) Die Menge der ganzen Zahlen:

$$\mathbb{Z} = \left\{ \dots, -2, -1, 0, 1, 2, \dots \right\}$$

ist eine Gruppe mit der Addition, denn die Null ist ein neutrales Element und das Inverse zu $k \in \mathbb{Z}$ ist $-k \in \mathbb{Z}$. Ebenso sind $(\mathbb{Q}, +)$ und $(\mathbb{R}, +)$ Gruppen.

(c) Die Menge $\mathbb{Q}^\times = \mathbb{Q} \smallsetminus \{0\}$, so ist G mit der Multiplikation eine Gruppe, ebenso $\mathbb{R}^\times = \mathbb{R} \smallsetminus \{0\}$

(d) Die Menge $\mathbb{R}$ ist mit dem arithmetischen Mittel keine Gruppe, denn diese Verknüpfung ist nicht assoziativ. Es ist

$$(a * b) * c = \left(\frac{a+b}{2}\right) * c = \frac{a+b}{4} + \frac{c}{2} = \frac{a}{4} + \frac{b}{4} + \frac{c}{2}$$

im Allgemeinen verschieden von

$$a * (b * c) = a * \left(\frac{b+c}{2}\right) = \frac{a}{2} + \frac{b}{4} + \frac{c}{4}.$$

(e) Ist M eine Menge, dann ist die Menge $G = \mathrm{Per}(M)$ aller Permutationen auf M mit der Hintereinanderausführung als Verknüpfung eine Gruppe. Das neutrale Element ist die identische Abbildung $\mathrm{Id} : M \to M$ und die Inverse zu $f \in \mathrm{Per}(M)$ ist ihre Umkehrabbildung f^{-1}.

Ist etwa $M = \{1,2,3\}$, dann koennen wir $\mathrm{Per}(M)$ komplett hinschreiben:

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

Die Verknüpfung des zweiten und dritten sind zB:

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

Das Inverse des letzten Elements ist

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

Definition 1.3.4. Eine Gruppe G heißt **kommutativ** oder **abelsche Gruppe**, falls für alle $a, b \in G$ gilt

$$ab = ba.$$

Beispiele 1.3.5. (a) Die Gruppen $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{Q}^\times, \times)$, $(\mathbb{R}^\times, \times)$ sind abelsch.

(b) Ist M eine Menge mit mehr als zwei Elementen, dann ist $\mathrm{Per}(M)$ nicht abelsch. Dies zeigen wir exemplarisch an dem Beispiel der Menge $M = \{1,2,3\}$. Seien

$$\alpha = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

Dann ist
$$\alpha\beta = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \beta\alpha = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

Proposition 1.3.6. *Sei G eine Gruppe.*

(a) *Das Inverse a' zu a erfüllt auch $a'a = e$.*

(b) *Das neutrale Element e erfüllt auch $ea = a$ für jedes $a \in G$.*

(c) *Das neutrale Element e ist eindeutig bestimmt.*

(d) *Zu gegebenem $a \in G$ ist das inverse Element a' eindeutig bestimmt.*

Beweis. Wir beginnen mit (a). Ist a' ein Inverses zu a und ist a'' ein Inverses zu a', dann gilt

$$\begin{aligned} a'a &= (a'a)e && \text{neutrales Element} \\ &= (a'a)(a'a'') && \text{inverses Element} \\ &= (a'(aa'))a'' && \text{Assoziativgesetz, mehrfach} \\ &= (a'e)a'' && \text{inverses Element} \\ &= a'a'' && \text{neutrales Element} \\ &= e && \text{inverses Element} \end{aligned}$$

(b): Es gilt

$$\begin{aligned} ea &= (aa')a && \text{inverses Element} \\ &= a(a'a) && \text{Assoziativgesetz} \\ &= ae && \text{nach (a)} \\ &= a && \text{neutrales Element} \end{aligned}$$

(c): Sei e_2 ein weiteres neutrales Element. Dann gilt

$$\begin{aligned} e_2 &= ee_2 && \text{nach (b)} \\ &= e && \text{da } e_2 \text{ neutral} \end{aligned}$$

(d): Sei also a_2 ein weiteres Inverses zu a. Dann gilt

$$\begin{aligned} a_2 &= ea_2 && \text{nach (b)} \\ &= (a'a)a_2 && \text{nach (d)} \\ &= a'(aa_2) && \text{Assoziativgesetz} \\ &= a'e && \text{da } a_2 \text{ invers zu } a \\ &= a' && \text{neutrales Element} \end{aligned}$$

Damit ist die Proposition vollständig bewiesen. □

Da das inverse Element a' zu a nun eindeutig bestimmt ist, schreiben wir a^{-1} für a'.

Korollar 1.3.7. *Sei G eine Gruppe, dann gilt für alle $a, b, x, y \in G$,*

(a) $(a^{-1})^{-1} = a$,

(b) $(ab)^{-1} = b^{-1}a^{-1}$,

(c) $ax = ay \;\Rightarrow\; x = y$.

Beweis. Es gilt $a^{-1}a = e$, also ist a ein Inverses zu a^{-1} und (a) folgt aus der Eindeutigkeit des Inversen.

(b) Es gilt $(b^{-1}a^{-1})(ab) = b^{-1}(a^{-1}a)b = (b^{-1}e)b = b^{-1}b = e$.

(c) Multipliziere beide Seiten der Gleichung mit a^{-1}. □

Definition 1.3.8. Sei G eine Gruppe. Eine Teilmenge $H \subset G$ heißt **Untergruppe**, falls H mit der Verknüpfung von G selbst wieder eine Gruppe ist.

Lemma 1.3.9. *Sei H eine Untergruppe der Gruppe G.*

(a) *Das neutrale Element e von G liegt in H und ist das neutrale Element von H.*

(b) *Ist $h \in H$, so liegt das Inverse h^{-1} zu h in H und ist das Inverse bzgl H*

(c) *Eine nichtleere Teilmenge $H \subset G$ ist genau dann eine Untergruppe, wenn*

$$a, b \in H \quad \Rightarrow \quad ab^{-1} \in H$$

gilt.

Beweis. Sei e_H das neutrale Element von H und e das von G. Dann gilt in G die Identität $e = e_H e_H^{-1} = (e_H e_H) e_H^{-1} = e_H (e_H e_H^{-1}) = e_H e = e_H$ und (a) ist bewiesen.

(b): Sei $h \in H$ und sei h' das Inverse bezüglich H. Dann gilt $hh' = e_H = e$ und damit ist $h' = h^{-1}$ das Inverse in G.

(c): Ist H eine Untergruppe und sind $a, b \in H$, so folgt nach (b), dass $b^{-1} \in H$ und damit $ab^{-1} \in H$. Sei umgekehrt H eine nichtleere Teilmenge, die der Bedingung des Lemmas genügt. Sei $h \in H$ ein Element, dann ist $e = hh^{-1} \in H$, also enthält H das neutrale Element. Wir zeigen nun, dass H zu jedem Element h auch sein Inverses h^{-1} enthält. Hierzu schreibe $h^{-1} = eh^{-1} \in H$. Seien nun $a, b \in H$, dann liegt $b^{-1} \in H$ und es gilt $ab = a(b^{-1})^{-1} \in H$, also ist H abgeschlossen unter der Multiplikation. Zusammen folgt, dass H eine Untergruppe ist. □

Beispiel 1.3.10. Sei $m \in \mathbb{N}$ eine natürliche Zahl. Auf der Menge

$$\mathbb{Z}/m = \{0, 1, \dots, m-1\}$$

definieren wir eine Addition $\boxplus$ durch

$$a \boxplus b = \text{Rest von } a+b \text{ bei Division nach } m.$$

Dann ist $\mathbb{Z}/m$ eine Gruppe. Das neutrale Element ist die Null und das Inverse zu a ist $m-a$, wenn $a \neq 0$.

Für $m = 1$ ist $\mathbb{Z}/m = \{0\}$ die triviale Gruppe.

Für $m = 2$ ist $\mathbb{Z}/m = \{0, 1\}$ und es gilt $1 \boxplus 1 = 0$.

Für $m = 3$ geben wir die Addition durch folgende Tabelle an:

$\boxplus$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

* * *

1.4 Körper

Definition 1.4.1. Ein **Körper** ist ein Tripel $(K, +, \times)$ bestehend aus einer Menge K und zwei Verknüpfungen $+$ und $\times$, wobei wir $a \times b$ als ab schreiben, so dass die folgenden Axiome erfüllt sind:

- $(K, +)$ ist eine abelsche Gruppe. Wir schreiben das neutrale Element dieser Gruppe als $0 = 0_K$ und nennen es das **Nullelement**.
- $(K \smallsetminus \{0\}, \times)$ ist eine abelsche Gruppe. Wir schreiben das neutrale Element als $1 = 1_K$ und nennen es das **Einselement**.
- Es gilt das **Distributivgesetz**: für alle $a, b, c \in K$ gilt

$$a(b+c) = ab + ac.$$

Beispiele 1.4.2. (a) $(\mathbb{Z}, +, \times)$ ist kein Körper, da nicht jedes Element $x \neq 0$ ein multiplikatives Inverses besitzt, so gibt es zum Beispiel kein Inverses für $x = 2$.

(b) $(\mathbb{Q}, +, \times)$ ist ein Körper und $\mathbb{R}$ ebenso.

(c) Auf der Menge $\mathbb{F}_2 = \{0,1\}$ definieren wir Addition und Multiplikation durch

$+$	0	1
0	0	1
1	1	0

$\times$	0	1
0	0	0
1	0	1

Dann ist $\mathbb{F}_2$ ein Körper.

(d) Allgemeiner sei p eine Primzahl. Auf der Gruppe $\mathbb{F}_p = \mathbb{Z}/p$ erklären wir Addition wie gehabt und die Multiplikation durch

$$a \boxtimes b = \text{Rest von } ab \text{ bei Division durch } p.$$

Dann ist $(\mathbb{F}_p, \boxplus, \boxtimes)$ ein Körper (LinA2). Für $\mathbb{F}_3$ sind die Verknüpfungstabellen:

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

$\times$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Schreibweise: Wir schreiben $-a$ für das additive Inverse zu a und $a - b$ für $a + (-b)$.

Satz 1.4.3. *Sei K ein Körper. Dann gilt für alle $a, b, c \in K$:*

(a) $0 \neq 1,$

(b) $a0 = 0$

(c) $ab = 0 \quad \Rightarrow \quad a = 0$ *oder* $b = 0,$ *Nullteilerfreiheit*

(d) $a(-b) = -ab = (-a)b, \quad (-a)(-b) = ab,$

(e) $ab = ac, a \neq 0 \quad \Rightarrow \quad b = c,$

(f) $a(b-c) = ab - ac.$

Beweis. (a) Das Element 1 ist das neutrale Element von $K^\times = K \setminus \{0\}$, also ist $1 \neq 0$.
(b) Für $a \in K$ gilt $a0 = a(0+0) = a0 + a0$. Nun addiere $-0a$ auf beiden Seiten und erhalte $0 = a0 - a0 = a0 + a0 - a0 = a0$.
(c) Ist $a, b \in K^\times$, so auch ab, denn $(ab)(b^{-1}a^{-1} = a(bb^{-1})a^{-1} = aa^{-1} = 1$.
(d) Es gilt $ab + a(-b) = a(b-b) = a0 = 0$. Also folgt wegen der Eindeutigkeit der Inversen, dass $a(-b) = -ab$. Der Rest geht ähnlich.
(f) $a(b-c) = a(b + (-c)) = ab + a(-c) = ab - ac$.
(e) Es gelte $ab = ac$ mit $a \neq 0$. Dann hat a ein multiplikatives Inverses a^{-1}. Multipliziere

beide Seiten der Gleichung mit a^{-1} und erhalte $b = c$. □

Definition 1.4.4. Sei K ein Körper. Die Zahl

$$\mathrm{char}(K) = \begin{cases} \min\left\{n \in \mathbb{N} : n \cdot 1_K = 0_K\right\}, & \text{falls endlich,} \\ 0 & \text{sonst,} \end{cases}$$

heißt die **Charakteristik** des Körpers K.

Beispiele 1.4.5. (a) Die Charakteristik von $\mathbb{F}_p$ ist p.

(b) Die Charakteristik von $\mathbb{Q}$ oder $\mathbb{R}$ ist 0.

Bemerkung. Ist die Charakteristik des Körpers K gleich Null, dann ist die Abbildung

$$\mathbb{N} \to K, \qquad n \mapsto n \cdot 1_K$$

injektiv. Also kann man in diesem Fall $\mathbb{N}$ als Teilmenge von K auffassen.

* * *

Die Körper der rationalen, reellen und komplexen Zahlen Die Menge der rationalen Zahlen $\mathbb{Q}$ reicht nicht aus, um die Welt zu beschreiben. So hat etwa seit Pythagoras ein quadratischer Tisch der Kantenlänge 1 eine Diagonallänge von $\sqrt{2}$. Diese ist aber keine rationale Zahl.

Proposition 1.4.6. *Es gibt keine rationale Zahl r mit $r^2 = 2$.*

Beweis. *Angenommen* doch, etwa $r = \frac{p}{q}$ mit teilerfremden p und q. Dann folgt $2 = r^2 = \frac{p^2}{q^2}$, also $p^2 = 2q^2$. Damit ist p^2 eine gerade Zahl und folglich ist auch p gerade, etwa $p = 2k$ mit $k \in \mathbb{Z}$. Es folgt $2q^2 = (2k)^2 = 4k^2$, also $q^2 = 2k^2$. Damit ist auch q eine gerade Zahl, was der angenommenen Teilerfremdheit *widerspricht!* □

Da die rationalen Zahlen also zur Erfassung der Welt nicht ausreichen, führt man die reellen Zahlen ein, in denen es unter anderem auch eine Zahl $r = \sqrt{2}$ gibt, deren Quadrat die Zahl 2 ist. Man kann die Menge der reellen Zahlen als die Menge aller Dezimalzahlen mit moeglicherweise unendlich vielen Nachkommastellen, also zum Beispiel

$$\pi = 3,141592653589793...$$

oder

$$\sqrt{2} = 1,414213562373095...$$

einführen. Es gibt da allerdings ein kleines Problem mit der Eindeutigkeit, denn es ist etwa

$$0,999\cdots = 0.\overline{9} = 1.$$

Dieses Problem kann ausgeräumt werden, indem man nur solche Zahlen betrachtet, die nicht in einer 9-er Periode enden.

Definition 1.4.7. Eine **Dezimalzahl** ist ein Ausdruck der Form

$$\pm a_N \dots a_1, b_1 b_2 \dots$$

wobei $N \in \mathbb{N}$ ist und $a_j, b_j \in \{0, \dots, 9\}$ gilt. Ferner soll $a_N \neq 0$ oder $N = 1$ sein. Eine Dezimalzahl heißt **regulär**, wenn sie nicht in einer 9-er Periode endet, d.h., wenn es zu jedem $k \in \mathbb{N}$ ein $j > k$ gibt, so dass $b_j \neq 9$ gilt. Die Menge $\mathbb{R}$ der **reellen Zahlen** wird definiert als die Menge aller regulären Dezimalzahlen. Eine genauere Diskussion der reellen Zahlen wird in der Analysis geliefert.

* * *

Die komplexen Zahlen

Auf der Menge $\mathbb{R}^2$ führen wir folgende Verknüpfungen ein:

Addition:

$$(a,b) + (x,y) = (a+x, b+y)$$

und **Multiplikation:**

$$(a,b)(x,y) = (ax - by, ay + bx).$$

Lemma 1.4.8. *Mit diesen Verknüpfungen ist die Menge $\mathbb{R}^2$ ein Körper. Das Nullelement ist $(0,0)$ und das Einselement ist $(1,0)$. Das additive Inverse zu (a,b) ist $(-a,-b)$ und das multiplikative Inverse zu $(a,b) \neq 0$ ist $\left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right)$.*

Beweis. Das muss man im Einzelnen nachrechnen, was langwierig ist, aber nicht schwierig. Als Beispiel rechnen wir die Assoziativität der Multiplikation. Seien also $(x,y), (a,b), (u,v) \in \mathbb{R}^2$, dann gilt

$$\begin{aligned}\big((a,b)(u,v)\big)(x,y) &= (au - bv, av + bu)(x,y)\\ &= (aux - bvx - avy - buy, auy - bvy + avx + bux)\end{aligned}$$

und

$$(a,b)\big((u,v)(x,y)\big) = (a,b)(ux-vy, uy+vx)$$
$$= (aux - avy - buy - bvx, auy + avx + bux - bvy). \qquad \square$$

Wir nennen die Menge $\mathbb{R}^2$ mit diesen Verknüpfungen die **Menge der komplexen Zahlen**, geschrieben $\mathbb{C}$ und setzen

$$1_{\mathbb{C}} = (1,0), \qquad i = (0,1).$$

Jede komplexe Zahl lässt sich eindeutig schreiben als $z = a1_{\mathbb{C}} + bi$ mit $a,b \in \mathbb{R}$. Die Abbildung $a \mapsto a1_{\mathbb{C}}$ identifiziert $\mathbb{R}$ mit einer Teilmenge von $\mathbb{C}$. Es gilt $i^2 = -1$. Man definiert den **Realteil** einer komplexen Zahl $z = x + yi$ als

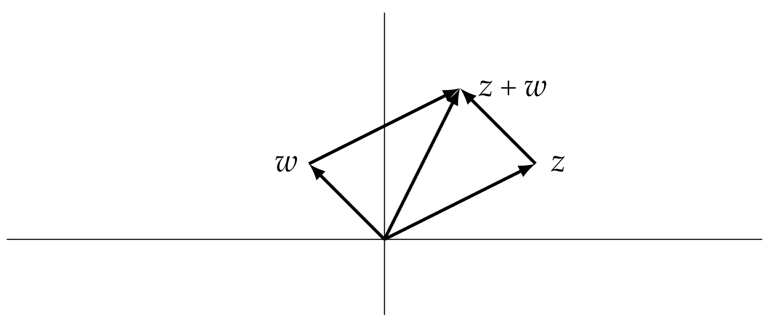
$$\mathrm{Re}(z) = x$$

und den **Imaginärteil** als

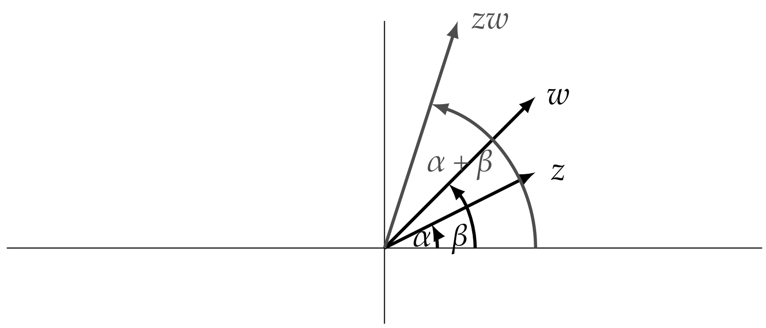
$$\mathrm{Im}(z) = y.$$

Es gilt dann also $z = \mathrm{Re}(z) + \mathrm{Im}(z)i$.

Zeichnet man $\mathbb{R}^2$ als Ebene mit Koordinaten, kann man die Addition als Vektoraddition visualisieren:



Die Multiplikation erhält man wie folgt: Die Längen der beteiligten Vektoren werden multipliziert und die Winkel, die sie mit der positiven x-Achse bilden, addiert:



* * *

2 Vektorräume und lineare Abbildungen

2.1 Räume und Unterräume

Sei K ein Körper. Ein **Vektorraum** über K ist eine abelsche Gruppe $(V,+)$ zusammen mit einer Abbildung

$$\begin{aligned} K \times V &\to V, \\ (\lambda, v) &\mapsto \lambda v, \end{aligned}$$

so dass für alle $v, w \in V$ und alle $\lambda, \mu \in K$ gilt:

(a) $1 \cdot v = v$,

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

(c) $(\lambda + \mu)v = \lambda v + \mu v, \qquad \lambda(v + w) = \lambda v + \lambda w.$

Beispiele 2.1.1. (a) Sei $K^n = K \times \cdots \times K$ die Menge der n-tupel $(x_1, \dots, x_n)$ von Elementen von K. Aus Gründen, die später klar werden, schreiben wir die Elemente von K^n als Spalten:

$$K^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : x_1, \dots, x_n \in K \right\}$$

Die Addition

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

macht K^n zu einer abelschen Gruppe. Die skalare Multiplikation ist erklärt durch

$$\lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}$$

für $\lambda \in K$.

(b) Sei S irgendeine Menge mit $S \neq \emptyset$. Sei $V = \mathrm{Abb}(S, K)$ die Menge aller Abbildungen $f : S \to K$. Durch die Addition

$$(f + g)(s) = f(s) + g(s)$$

wird V eine abelsche Gruppe und die Vorschrift

$$(\lambda f)(s) = \lambda(f(s))$$

macht V zu einem Vektorraum.

Beweis. Die Nullfunktion $f_0(s) = 0$ ist ein neutrales Element für die Addition, denn für jedes $f \in V$ gilt

$$(f + f_0)(s) = f(s) + f_0(s) = f(s) + 0 = f(s).$$

Die anderen Gruppeneigenschaften rechnet man ebenso nach. Wir zeigen noch die Vektorraumaxiome. Es gilt

$$(1 \cdot f)(s) = 1 \cdot f(s) = f(s),$$

für jedes $s \in S$, also $1 \cdot f = f$. Ferner ist für $\lambda, \mu \in K$ und $f \in V$,

$$((\lambda\mu)f)(s) = (\lambda\mu)(f(s)) = \lambda(\mu(f(s)) = \lambda((\mu f)(s)) = (\lambda(\mu f))(s).$$

Schliesslich ist

$$((\lambda + \mu)f)(s) = (\lambda + \mu)(f(s)) = \lambda(f(s)) + \mu(f(s)) = (\lambda f)(s) + (\mu f)(s) = (\lambda f + \mu f)(s).$$

Und ebenso $\lambda(f + g) = \lambda f + \lambda g$. □

(c) Das obige Beispiel (a) ist eigentlich ein Spezialfall von (b), denn K^n kann auch als $\mathrm{Abb}(\underline{n}, K)$ aufgefasst werden, wobei $\underline{n} = \{1, 2, \dots, n\}$ ist. In diesem Fall identifiziert man eine Abbildung $f : \underline{n} \to K$ mit den Spaltenvektor

$$\begin{pmatrix} f(1) \\ f(2 \\ \vdots \\ f(n) \end{pmatrix}.$$

Man schreibt daher auch K^S für die Menge $\mathrm{Abb}(S, K)$ aller Abbildungen von S nach K.

(d) Sei $K[x]$ die Menge aller **Polynome**

$$a_0 + a_1 x + \dots + a_n x^n,$$

$n \in \mathbb{N}$, $a_0, \dots, a_n \in K$. Man beachte, dass man n jederzeit durch ein $m > n$ erstzen kann, indem man die überflüssigen Koeffiziienten alle zu Null setzt. Die Menge $K[x]$ wird ein Vektorraum durch

$$\left(a_0 + a_1 x + \dots + a_m x^m\right) + \left(b_0 + b_1 x + \dots + b_m x^m\right) = (a_0 + b_0) + (a_1 + b_1)x + \dots + (a_m + b_m)x^m$$

und

$$\lambda\Big(a_0 + a_1 x + \cdots + a_m x^m\Big) = \lambda a_0 + \lambda a_1 x + \cdots + \lambda a_m x^m.$$

Definition 2.1.2. Sei V ein Vektorraum. Eine Teilmenge $U \subset V$ heißt **Unterraum** oder **Untervektorraum**, wenn U mit der Addition und der Skalarmultiplikation von V selbst ein Vektorraum ist.

Proposition 2.1.3. (a) *Ist V ein Vektorraum und ist $v \in V$, dann gilt $0 \cdot v = 0$ und $(-1) \cdot v$ ist der Inverse zu v, also $v + (-1)v = 0$.*

(b) *Sei V ein K-Vektorraum. Eine nichtleere Teilmenge $U \subset V$ ist genau dann ein Unterraum, wenn gilt*

$$u, w \in U \quad \lambda, \mu \in V \quad \Rightarrow \quad \lambda u + \mu w \in U. \tag{1}$$

Beweis. (a) Es gilt $(0 \cdot v) + (0 \cdot v) = (0 + 0)v = 0 \cdot v$, woraus nach Subtraktion von $0 \cdot v$ die erste Behauptung folgt. Für die zweite betrachte $v + (-1)v = \big(1 + (-1)\big)v = 0v = 0$.

(b) "$\Rightarrow$" Wenn U ein Unterraum ist, ist die Bedingung offensichtlich erfüllt.

"$\Leftarrow$" Sei U nichtleer und erfülle die Bedingung (1). Dann definiert "+" eine Verknüpfung auf U. Nach Lemma 1.3.9 ist $(U, +)$ eine Untergruppe von $(V, +)$. Die verbleibenden Axiome gelten in U, da sie in V gelten. □

Beispiel 2.1.4. Sei S eine Menge und sei $V = \mathrm{Abb}(S, K)$. Fixiere ein $s_0 \in S$ und setze

$$U = \Big\{ f \in V : f(s_0) = 0 \Big\}.$$

Dann ist U ein Unterraum von V.

Beweis. U ist nichtleer, da die Nullabbildung in U liegt.

(a) Seien $f, g \in U$, so gilt

$$(f + g)(s_0) = f(s_0) + g(s_0) = 0 + 0 = 0,$$

also gilt $f + g \in U$.

(b) Sei $f \in U$ und $\lambda \in K$, so gilt

$$(\lambda f)(s_0) = \lambda(f(s_0)) = \lambda \cdot 0 = 0.$$

Also ist $\lambda f \in U$. □

Definition 2.1.5. Ist I irgendeine Indexmenge und ist für jedes $i \in I$ eine Menge A_i

gegeben, so definieren wir die **Schnittmenge**

$$\bigcap_{i\in I} A_i$$

als die Menge aller x, die in jeder der Mengen A_i liegen. Das heisst:

$$x \in \bigcap_{i\in I} A_i \quad \Leftrightarrow \quad x \in A_i \text{ für jedes } i \in I.$$

Proposition 2.1.6. *Sind U, W Unterräume von V, so ist $U \cap W$ ein Unterraum.*

Allgemeiner gilt: Ist I eine Indexmenge und ist für jedes $i \in I$ ein Unterraum U_i gegeben, so ist

$$U = \bigcap_{i\in I} U_i$$

ein Unterraum von V.

Beweis. Wir zeigen nur die zweite, allgemeinere Aussage. Zunächst ist $U \neq \emptyset$, da die Null in jedem U_i und damit in U liegt. Seien $u, v \in U$ und $\lambda \mu \in K$. Dann gilt $u, v \in U_i$ für jedes $i \in I$. Also ist $\lambda u + \mu v \in U_i$ für jedes $i \in I$ und daher ist $\lambda u + \mu v \in U$. □

* * *

2.2 Summen und direkte Summen

Definition 2.2.1. Sei V ein K-Vektorraum und seien U, $W \subset V$ Unterräume. Die **Summe** von U und W ist definiert als

$$U + W = \{u + w : u \in U, w \in W\}.$$

Beispiel 2.2.2. Sei $K = \mathbb{R}$ und $V = \mathbb{R}^3$, sowie

$$U = \left\{ \begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix} : x \in \mathbb{R} \right\}$$

und

$$W = \left\{ \begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix} : y \in \mathbb{R} \right\}.$$

Dann sind U und W zwei verschiedene Geraden im dreidimensionalen Raum. Es ist dann

$$U + W = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} : x, y \in \mathbb{R} \right\}$$

die Ebene, die von diesen beiden Geraden aufgespannt wird.

Definition 2.2.3. Sind $U_1, \dots U_m \subset V$ Unterräume von V, so definiert man

$$U_1 + \dots + U_m = \Big\{u_1 + \dots + u_m : u_1 \in U_1, \dots, u_m \in U_m\Big\}.$$

Beispiel 2.2.4. Sei $V = K^n$, sowie $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix}$. Für jedes $j \in \big\{1, \dots, n\big\}$ sei

$$U_j = Ke_j = \underbrace{\big\{\lambda e_j : \lambda \in K\big\}}_{\text{die von } e_j \text{ aufgespannte Gerade}}$$

Dann gilt

$$U_1 + \dots + U_n = V.$$

Bemerkung. Für einen Unterraum $U \subset V$ gilt

$$U + U = U.$$

Beweis. "$\subset$" Sei $v \in U + U$. Dann gibt es $u_1, u_2 \in U$ mit $v = u_1 + u_2$. Da U ein Unterraum ist, folgt $v \in U$.

"$\supset$" Sei $u \in U$. Dann ist $u = u + 0 \in U + U$. □

Definition 2.2.5. Seien $U_1, \dots, U_m$ Unterräume von V. Die Summe $U = U_1 + \dots U_m$ ist eine **direkte Summe**, wenn jedes $u \in U$ auf **genau eine Weise** als $u = u_1 + \dots + u_m$, $u_j \in U_j$ geschrieben werden kann. Ist dies der Fall, so schreiben wir

$$U = U_1 \oplus U_2 \oplus \dots \oplus U_m = \bigoplus_{j=1}^{m} U_j.$$

Beispiele 2.2.6. (a) Sei $V = K^3$ und

$$U = \left\{\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} : x, y \in K\right\},$$

sowie

$$W = \left\{\begin{pmatrix} 0 \\ y \\ z \end{pmatrix} : y, z \in K\right\}.$$

Dann ist die Summe $U + W$ nicht direkt, denn die Null kann auf zwei Weisen geschrieben werden:

$$0 = 0 + 0, \qquad 0 = \underbrace{\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}}_{\in U} + \underbrace{\begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}}_{\in W}.$$

(b) Sei $V = K^3$, $U_1 = Ke_1$ und $U_2 = Ke_2$. Dann ist die Summe $U = U_1 + U_2$ direkt

Beweis. Sei $u \in U$ mit $u = u_1 + u_2 = u_1' + u_2'$, wobei $u_j, u_j' \in U_j$ gilt. Wir müssen zeigen, dass $u_1 = u_1'$ und $u_2 = u_2'$ gilt. Hierzu schreibe $u_1 = \lambda_1 e_1$, $u_2 = \lambda_2 e_2$, sowie $u_1' = \lambda_1' e_1$ und $u_2' = \lambda_2' e_2$. Es folgt

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ 0 \end{pmatrix} = \lambda_1 e_1 + \lambda_2 e_2 = u = \lambda_1' e_1 + \lambda_2' e_2 = \begin{pmatrix} \lambda_1' \\ \lambda_2' \\ 0 \end{pmatrix}.$$

Also folgt $\lambda_1 = \lambda_1'$ sowie $\lambda_2 = \lambda_2'$, also $u_1 = u_1'$ sowie $u_2 = u_2'$. □

Proposition 2.2.7. *Seien $U, W, U_1, \dots, U_m$ Unterräume eines Vektorraums V.*

(a) *Die Summe $U + W$ ist genau dann direkt, wenn $U \cap W = 0$.*

(b) *Die Summe $U_1 + \dots + U_m$ ist genau dann direkt, wenn für alle $u_1 \in U_1, \dots, u_m \in U_m$ gilt*

$$u_1 + \dots + u_m = 0 \quad \Rightarrow \quad u_1 = \dots = u_m = 0.$$

Beweis. (a) Sei die Summe direkt und sei $v \in U \cap W$. Dann sind

$$0 = \underbrace{0}_{\in U} + \underbrace{0}_{\in W} = \underbrace{v}_{\in U} + \underbrace{(-v)}_{\in W}$$

zwei Schreibweisen der Null. Diese müssen nach Definition der direkten Summe gleich sein, also $v = 0$.

Sei umgekehrt $U \cap W = 0$ und sei $u + w = u' + w'$ mit $u, u' \in U$ und $w, w' \in W$. Dann ist $u - u' = w' - w \in U \cap W$ also gleich Null und damit $u = u'$ und $w = w'$.

(b) "$\Rightarrow$" folgt direkt aus der Definition. Wir zeigen "$\Leftarrow$": Sei $u_1 + \dots + u_m = u_1' + \dots + u_m'$ mit $u_j, u_j' \in U_j$. Dann ist

$$u_1 - u_1' + \dots + u_m - u_m' = 0,$$

also $u_1 = u_1', \dots, u_m = u_m'$. □

2.3 Linearkombinationen und Lineare Unabhängigkeit

Seien $v_1, \dots, v_m$ Elemente des Vektorraums V. Eine **Linearkombination** der Vektoren $v_1, \dots, v_m$ ist jeder Vektor der Gestalt

$$\lambda_1 v_1 + \cdots + \lambda_m v_m$$

wobei $\lambda_1, \dots, \lambda_m \in K$. Die Menge aller Linearkombinationen ist der **Spann** von $v_1, \dots, v_m$, also

$$\mathrm{Spann}(v_1, \dots, v_m) = \left\{\lambda_1 v_1 + \cdots + \lambda_m v_m : \lambda_1, \dots, \lambda_m \in K\right\}.$$

Beispiele 2.3.1. (a) Seien $K = \mathbb{R}$, $V = \mathbb{R}^3$. Seien weiter $v_1, v_2 \in V$, wobei keiner ein Vielfaches des anderen ist. Dann ist $\mathrm{Spann}(v_1, v_2)$ die eindeutig bestimmte Ebene, die die Punkte $0, v_1, v_2$ enthält.

(b) Sei $K = \mathbb{Q}$. der Vektor $\begin{pmatrix} 7 \\ 2 \\ 9 \end{pmatrix}$ liegt im Spann der Vektoren $\begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, denn es gilt

$$\begin{pmatrix} 7 \\ 2 \\ 9 \end{pmatrix} = 2\begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} + 3\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Proposition 2.3.2. *Sei V ein Vektorraum und $u_1, \dots, u_m \in V$. Dann ist* $\mathrm{Spann}(u_1, \dots, u_m)$ *ein Untervektorraum von V.*

Beweis. Sei $U = \mathrm{Spann}(u_1, \dots, u_m)$. Wir müssen zeigen:

$$u, v \in U,\ \lambda, \mu \in K \Rightarrow \lambda u + \mu v \in U.$$

Seien also $u = \lambda_1 u_1 + \cdots + \lambda_m u_m$ und $v = \mu_1 u_1 + \cdots + \mu_m u_m$ und $\lambda \in K$. Dann ist

$$\lambda u + \mu v = (\lambda\lambda_1 + \mu\mu_1)u_1 + \cdots + (\lambda\lambda_m + \mu\mu_m)u_m \in U,$$ □

Definition 2.3.3. Gilt $V = \mathrm{Spann}(v_1, \dots, v_m)$, so sagen wir, dass die Vektoren $v_1, \dots, v_m$ den Vektorraum V **aufspannen**. Ist dies der Fall, sagen wir, dass die Menge $\left\{v_1, \dots, v_m\right\}$ ein **Erzeugersystem** von V ist. Man nennt m dann auch die **Länge** des Erzeugersystems.

Beispiel 2.3.4. Sei $V = K^n$. Die Vektoren $e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ spannen der Raum V

auf, denn es gilt für $v \in V$,

$$v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 e_1 + \cdots + x_n e_n.$$

Definition 2.3.5. Ein Vektorraum V heißt **endlich-dimensional**, falls es ein endliches Erzeugersystem gibt. Ansonsten heißt er unendlich-dimensional.

Beispiele 2.3.6. (a) Der Raum K^n ist endlich-dimensional.

(b) Der Vektorraum $K[x]$ der Polynome ist nicht endlich-dimensional.

Beweis. Wir definieren den **Grad** eines Polynoms $f(x) = a_0 + \cdots + a_n x^n$ als

$$\mathrm{grad}(f) = \begin{cases} \max\left\{k \in \mathbb{N}_0 : a_k \neq 0\right\} & f \neq 0, \\ -\infty & f = 0. \end{cases}$$

So ist zum Beispiel der Grad von $1 + 2x + x^2$ gleich 2. Wir zeigen, dass es zu je $f_1, \dots, f_m \in V$ ein f gibt, das nicht im Spann der $f_1, \dots, f_m$ liegt. Sei hierzu N das Maximum der Grade $\mathrm{grad}(f_j)$, $j = 1, \dots, m$. Dann liegt

$$f(x) = x^{N+1}$$

nicht im Spann der f_j. □

(c) Sei S eine nichtleere Menge. Ist S endlich, dann ist $V = \mathrm{Abb}(S, K)$ endlich-dimensional.

Beweis. Sei $S = \left\{s_1, \dots, s_n\right\}$ Für $1 \le j \le n$ sei $\delta_j : S \to K$ definiert durch

$$\delta_j(t) = \begin{cases} 1 & t = s_j, \\ 0 & t \neq s_j. \end{cases}$$

Wir behaupten, dass $\delta_1, \dots, \delta_n$ ein Erzeugersystem ist. Sei hierzu $f \in V$ und sei $\lambda_j = f(s_j) \in K$ für $j = 1, \dots, n$. Wir zeigen

$$f = \lambda_1 \delta_1 + \cdots + \lambda_n \delta_n.$$

Sei $s \in S$. Dann gibt es genau ein $1 \le i \le n$ mit $s = s_i$. Es ist dann

$$f(s) = f(s_i) = \lambda_i = \lambda_1 \delta_1(s) + \cdots + \lambda_n \delta_n(s).$$

Damit folgt die Behauptung. □

* * *

Definition 2.3.7. Die Vektoren $v_1, \dots, v_m \in V$ heißen **linear unabhängig** oder nur **unabhängig**, falls es keine Linearkombination der Null gibt, außer der trivialen, also wenn gilt

$$\lambda_1 v_1 + \cdots + \lambda_m v_m = 0 \quad \Rightarrow \quad \lambda_1 = \cdots = \lambda_m = 0.$$

Mit anderen Worten, die Vektoren sind unabhängig, wenn es nur eine Linearkombination der Null gibt. Man sagt in diesem Fall auch, dass die Menge der Vektoren $\{v_1, \dots, v_n\}$ unabhängig ist.

Andernfalls heissen sie **abhängig**.

Beispiele 2.3.8. (a) Die Vektoren $e_1, \dots, e_n \in K^n$ sind unabhängig, denn ist $\lambda_1 e_1 + \cdots + \lambda_n e_n = 0$, so heißt das

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \lambda_1 e_1 + \cdots + \lambda_n e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

also $\lambda_1 = \cdots = \lambda_n = 0$.

(b) Sei $K = \mathbb{Q}$. Die Vektoren $\begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 7 \\ 3 \\ 8 \end{pmatrix} \in \mathbb{Q}^3$ sind abhängig, denn

$$2\begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} + 3\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} + (-1)\begin{pmatrix} 7 \\ 3 \\ 8 \end{pmatrix} = 0.$$

Proposition 2.3.9. *Seien $v_1, \dots, v_n$ Vektoren in V. Dann sind äquivalent:*

(a) *Die Vektoren $v_1, \dots, v_n$ sind abhängig.*

(b) *Einer der Vektoren ist eine Linearkombination der anderen.*

(c) *Es gibt ein $v \in \mathrm{Spann}(v_1, \dots, v_n)$ mit zwei verschiedenen Darstellungen*

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n = \mu_1 v_1 + \cdots + \mu_n v_n.$$

Beweis. (a)⇒(b): Es gebe eine nichttriviale Darstellung der Null

$$0 = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \lambda_n v_n.$$

Nach Umnummerierung koennen wir $\alpha_1 \neq 0$ annehmen. Dann folgt

$$v_1 = \frac{-\alpha_2}{\alpha_1} v_2 + \cdots + \frac{-\alpha_n}{\alpha_1} v_n.$$

(b)⇒(c): Nach Umnummerierung sei $v_1 = \beta_2 v_2 + \cdots + \beta_n v_n$, so hat also v_1 zwei verschiedene Darstellungen.

(c)⇒(a): Es gelte

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n = \mu_1 v_1 + \cdots + \mu_n v_n.$$

Dann folgt

$$0 = v - v = (\lambda_1 - \mu_1) v_1 + \cdots + (\lambda_n - \mu_n) v_n$$

und es gibt ein j mit $\lambda_j - \mu_j \neq 0$. □

Lemma 2.3.10. *Sind $v_1, \dots, v_m$ linear unabhängig und ist $w \in V$, dann sind äquivalent:*

(a) $w \in \mathrm{Spann}(v_1, \dots, v_m)$,

(b) $w, v_1, \dots, w_m$ *sind abhängig.*

Beweis. (a)⇒(b) folgt aus Proposition 2.3.9.

(b)⇒(a): Sei $\lambda w + \lambda_1 v_1 + \cdots + \lambda_m v_m = 0$ eine nichttriviale Linearkombination der Null. Dann ist $\lambda \neq 0$, denn sonst wäre schon $\lambda_1 v_1 + \cdots + \lambda_m v_m = 0$ nichttrivial, was der Unabhängigkeit der v_j wiederspricht. Also folgt

$$w = \frac{-\lambda_1}{\lambda} v_1 + \cdots + \frac{-\lambda_n}{\lambda} v_m.$$ □

Lemma 2.3.11. *Sind $v_1, \dots, v_n$ unabhängig und ist $w_1, \dots, w_m$ ein Erzeugersystem, dann gilt*

$$n \leq m.$$

Proof. Seien $v_1, \dots, v_n$ unabhängig und sei $w_1, \dots, w_m$ ein Erzeugersystem. Wir müssen zeigen, dass $n \leq m$ gilt. Der Vektor v_1 hat eine Darstellung $v_1 = \lambda_1 w_1 + \cdots + \lambda_m w_m$ und eines der λ_j muss $\neq 0$ sein. Nach Umnummerierung sei $\lambda_m \neq 0$. Dann folgt

$$w_m = \frac{1}{\lambda_m} v_1 + \frac{-\lambda_1}{\lambda_n} w_1 + \cdots + \frac{-\lambda_{m-1}}{\lambda_m} w_{m-1}.$$

Also ist $v_1, w_1, \dots, w_{n-1}$ immer noch ein Erzeugersystem. Wir wiederholen den Schluss wie folgt: Gegeben, dass $v_1, \dots, v_j, w_1, \dots, w_{m-j}$ ein Erzeugersystem ist und $j < n$. Dann ist v_{j+1} eine Linearkombination dieser Vektoren, also ist nach Proposition 2.3.9 die Menge der $v_1, \dots, v_j, v_{j+1}, w_1, \dots, w_{m-j}$ abhängig. Sei $k \geq 0$ maximal mit der Eigenschaft, dass $v_1, \dots, v_{j+1}, w_1, \dots, w_k$ unabhängig sind. Es gibt dann eine Linearkombination

$\mu_1 v_1 + \cdots + \mu_{j+1} v_{j+1} + \lambda_1 w_1 + \cdots + \lambda_k w_k + \lambda_{k+1} w_{k+1} = 0$. Der Koeffizient λ_{k+1} muss $\neq 0$ sein, da $v_1, \ldots, v_{j+1}, w_1, \ldots, w_k$ unabhängig sind. Also ist $w_{k+1} = \frac{-\mu_1}{\lambda_{k+1}} v_1 + \cdots + \frac{-\mu_{j+1}}{\lambda_{k+1}} v_{j+1} + \frac{-\lambda_1}{\lambda_{k+1}} w_1 + \cdots + \frac{-\lambda_k}{\lambda_{k+1}} w_k$ und wir koennen w_{k+1} streichen. Nach Umnummerierung erhalten wir ein Erzeugersystem $v_1, \ldots, v_{j+1}, w_1, \ldots, w_{n-j-1}$. Der Prozess stoppt, wenn $j + 1 = n$ ist, also folgt $n \leq m$. □

Definition 2.3.12. Sei V ein Vektorraum. Ein unabhängiges Erzeugersystem nennt man auch eine **Basis** von V.

Definition 2.3.13. Ein Erzeugersystem $v_1, \ldots, v_n$ heisst **minimal**, falls keine echte Teilmenge ein Erzeugersystem ist. Man kann dann also keinen Vektor weglassen.

Satz 2.3.14. *Für ein Erzeugersystem $\mathcal{E}$ sind äquivalent:*

(a) *$\mathcal{E}$ ist eine Basis.*

(b) *Die Länge von $\mathcal{E}$ ist minimal.*

(c) *$\mathcal{E}$ ist minimal.*

Jedes Erzeugersystem enthält eine Basis.

Proof. (a)⇒(b) folgt aus Lemma 2.3.11.

(b)⇒(c) ist klar.

(c)⇒(a): $\mathcal{E}$ sei minimal. Ist $\mathcal{E}$ abhängig, dann ist nach Proposition 2.3.9 einer der Vektoren von $\mathcal{E}$ eine Linearkombination der anderen, kann also entfernt werden, was der Minimalität von $\mathcal{E}$ widerspricht.

Der Zusatz folgt, da man aus einem gegebenen Erzeugersystem Vektoren herausstreichen kann, bis man ein minimales Erzeugersystem erhalten hat. □

Definition 2.3.15. Eine unabhängige Menge $\{v_1, \ldots, v_k\} \subset V$ heisst **maximal**, falls jede groessere Menge in V abhängig ist.

Satz 2.3.16. *Sei $\mathcal{D} = \{v_1, \ldots, v_k\}$ linear unabhängig in V. Dann sind äquivalent:*

(a) *$\mathcal{D}$ ist eine Basis.*

(b) *$\mathcal{D}$ hat maximale Länge unter allen unabhängigen Mengen.*

(c) *$\mathcal{D}$ ist maximal.*

Jede unabhängige Menge lässt sich zu einer Basis erweitern.

Proof. (a)⇒(b): Da $\mathcal{D}$ ein Erzeugersystem ist, ist nach Lemma 2.3.11 die Länge von $\mathcal{D}$ $\geq$ der Länge jeder unabhängigen Menge.

(b)⇒(c) ist klar.

(c)⇒(a): Es ist nur zu zeigen, dass $\mathcal{D}$ erzeugend ist. Falls aber ein Vektor $w \notin \mathrm{Spann}(v_1, \dots, v_k)$ existiert, dann ist nach Lemma 2.3.10 die Menge $\{w, v_1, \dots, w_k\}$ immer noch unabhängig, also ist $\mathcal{D}$ nicht maximal. □

Definition 2.3.17. Sowohl aus Satz 2.3.14 als auch aus Satz 2.3.16 folgt, dass alle Basen von V dieselbe Länge haben. Diese wird die **Dimension** von V,

$$\dim(V)$$

genannt.

Bemerkung 2.3.18. (a) Eine Basis ist ein minimales Erzeugersystem oder ein maximales unabhängiges System.

(b) Ist $n = \dim V$, dann ist jedes Erzeugersystem der Länge n eine Basis.

(c) Ist $n = \dim V$, dann ist jedes unabhängige System der Länge n eine Basis.

* * *

2.4 Dimensionsformel für Unterräume

Proposition 2.4.1. *Jeder Unterraum eines endlich-dimensionalen Raums ist endlich-dimensional.*

Beweis. Sei $V = \mathrm{Spann}(v_1, \dots, v_n)$ und U ein Unterraum. Jede unabhängige Menge $\{u_1, \dots, u_k\}$ in U lässt sich zu einer Basis von V erweitern, also folgt $k \leq n$. Sei k die maximale Länge einer unabhängigen Menge $\{u_1, \dots, u_k\}$ in U. Wir behaupten, dass $u_1, \dots, u_k$ bereits ein Erzeugersystem von U ist. Wäre dem nicht so, dann gäbe es einen Vektor $u_0 \in U$ mit $u_0 \notin \mathrm{Spann}(u_1, \dots, u_k)$. Dann ist nach Lemma 2.3.10 aber $u_0, u_1, \dots, u_k$ unabhängig, was der Maximalität von k widerspricht. □

Satz 2.4.2. *Sei $U \subset V$ ein Unterraum des endlich-dimensionalen Vektorraums V. Dann existiert ein Unterraum $W \subset V$ so dass*

$$V = U \oplus W.$$

Der Raum W wird ein **Komplementärraum** *zu U genannt.*

Beweis. Nach Proposition 2.4.1 ist U endlich-dimensional, also hat U eine Basis $v_1, \dots, v_k$. Nach Satz 2.3.16 lässt sich diese zu einer Basis $v_1, \dots, v_n$ von V verlängern. Setze $W = \mathrm{Spann}(v_{k+1}, \dots, v_n)$. Dann ist $V = U + W$, da sich jedes $v \in V$ als Linearkombination der $v_1, \dots, v_n$ schreiben lässt. Ferner ist $U \cap W = 0$, denn ist $v \in U \cap W$, dann ist einerseits $v = \lambda_1 v_1 + \dots + \lambda_k v_k$ für geeignete Koeffizienten $\lambda_1, \dots, \lambda_k \in K$ und andererseits $v = \lambda_{k+1} v_{k+1} + \dots + \lambda_n v_n$ für Koeffizienten $\lambda_{k+1}, \dots, \lambda_n \in K$. Es folgt

$$0 = v - v = \lambda_1 v_1 + \dots + \lambda_k v_k + (-\lambda_{k+1}) v_{k+1} + \dots + (-\lambda_n) v_n,$$

woraus wegen der Unabhängigkeit $\lambda_1 = \dots = \lambda_n = 0$ und damit $v = 0$ folgt. □

Beispiel 2.4.3. Sei $K = \mathbb{Q}$ und $V = \mathbb{Q}^2$, sowie $U = \mathbb{Q}e_1$. Dann ist $\mathbb{Q}e_2$ ein Komplementärraum. Allerdings ist auch $\mathbb{Q}(\lambda e_1 + e_2)$ ein Komplementärraum für jedes $\lambda \in \mathbb{Q}$. Damit gibt es unendlich viele verschieden Komplementärräume.

Beispiel 2.4.4. Es ist $\dim(K^n) = n$, da $e_1, \dots, e_n$ eine Basis ist.

Proposition 2.4.5. (a) *Ist $U \subset V$ ein Unterraum des endlich-dimensionalen Vektorraums V, so gilt* $\dim U \le \dim V$. *Gleichheit gilt nur für $U = V$.*

(b) *Ist* $\dim V = n$*, so ist jedes Erzeugersystem der Länge n eine Basis. Ebenso ist jede unabhängige Menge der Länge n eine Basis.*

(c) *Für jeden Vektorraum V gilt*

$$V = 0 \quad \Leftrightarrow \quad \dim V = 0.$$

Beweis. (a) Jede Basis von U kann zu einer Basis von V verlängert werden, woraus die Ungleichgung folgt. Zur zweiten beachte, dass eine Basis von U der Länge $\dim V$ schon eine Basis von V ist, da sie nicht mehr verlängert werden kann.

(b) Sei $v_1, \dots, v_n$ ein Erzeugersystem der Länge n. Nach Satz 2.3.14 enthält es eine Basis. Diese muss aber die Länge $n = \dim V$ haben, also ist $v_1, \dots, v_n$ selbst eine Basis.

Sei $\{v_1, \dots, v_n\}$ unabhängig. Dann kann sie nach Satz 2.3.16 zu einer Basis verlängert werden. Diese muss die Länge n haben, also ist sie gleich $v_1, \dots, v_n$.

(c) Ist $V = 0$, so ist die leere Menge eine Basis und umgekehrt. □

Satz 2.4.6 (Dimensionsformel für Unterräume)**.** *Sind U und W Unterräume des endlich-dimensionalen Raums V, dann gilt*

$$\dim(U+W) = \dim U + \dim W - \dim(U \cap W).$$

Beweis. Sei $v_1, \dots, v_k$ eine Basis von $U \cap W$. Wir setzen sie fort zu einer Basis $v_1, \dots, v_k, v_{k+1}, \dots, v_m$ von U. Andererseits können wir sie aber auch zu einer Basis $v_1, \dots, v_k, v_{m+1}, \dots, v_n$ von W fortsetzen. Wir behaupten, dass $v_1, \dots, v_k, v_{k+1}, \dots, v_m, v_{m+1}, \dots, v_n$ eine Basis von $U + W$ ist. Da

$$U, W \subset \mathrm{Spann}(v_1, \dots, v_n) \subset U + W,$$

ist $v_1, \dots, v_n$ ein Erzeugersystem. Wir zeigen Unabhängigkeit. Sei hierzu $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$ eine Linearkombination der Null. Seien

$$\begin{aligned} v &= \lambda_1 v_1 + \dots + \lambda_k v_k \in U \cap W, \\ u &= \lambda_{k+1} v_{k+1} + \dots + \lambda_m v_m \in U, \\ w &= \lambda_{m+1} v_{m+1} + \dots + \lambda_n v_n \in W. \end{aligned}$$

Dann ist $v + u + w = 0$. Also ist $w = -v - u \in U$ und ebenso ist $u = -v - w \in W$ und damit $v, u, w \in U \cap W$. Es existieren also $\alpha_1, \dots, \alpha_k \in K$ mit

$$u = \alpha_1 v_1 + \dots + \alpha_k v_k = \lambda_{k+1} v_{k+1} + \dots + \lambda_m v_m.$$

Wegen der Unabhängigkeit von $v_1, \dots, v_m$ folgt $u = 0$ und damit $\lambda_{k+1} = \dots = \lambda_m = 0$. Ebenso folgt $w = 0$ und $\lambda_{m+1} = \dots = \lambda_n = 0$ und wegen $v + u + w = 0$ schliesslich $v = 0$ und $\lambda_1 = \dots = \lambda_k = 0$.

Damit ist

$$\dim(U+W) = n = m + (n - m + k) - k = \dim U + \dim W - \dim(U \cap W). \qquad \square$$

Beispiel 2.4.7. Sei $V = K^3$, sowie

$$U = \left\{ \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} : a, b \in K \right\}$$

$$W = \left\{ \begin{pmatrix} 0 \\ b \\ c \end{pmatrix} : b, c \in K \right\}$$

Dann ist
$$U \cap W = \left\{ \begin{pmatrix} 0 \\ b \\ 0 \end{pmatrix} : b \in K \right\},$$
man hat $\dim U = \dim W = 2$ und $\dim U \cap W = 1$. Schliesslich ist $V = U + W$ und
$$\dim V = 3 = 2 + 2 - 1 = \dim U + \dim W - \dim(U \cap W).$$

Korollar 2.4.8. *Gilt* $V = U + W$ *und* $\dim V = \dim U + \dim W$, *dann ist* $U \cap W = 0$, *die Summe ist dann also direkt.*

Proof. Aus der Dimensionsformel folgt dann $\dim(U \cap W) = 0$, also $U \cap W = 0$. □

2.5 Lineare Abbildungen

Definition 2.5.1. Seien V, W Vektorräume über dem Körper K. Eine **lineare Abbildung** ist eine Abbildung $T : V \to W$ mit

- $T(v + v') = T(v) + T(v')$
- $T(\lambda v) = \lambda T(v)$

für alle $v, v' \in V$ und jedes $\lambda \in K$.

Beispiele 2.5.2. (a) $T : V \to W$ die **Nullabbildung** $T(v) = 0$ ist linear.

(b) Die identische Abbildung $\mathrm{Id} : V \to V; v \mapsto v$ ist linear.

(c) Sein $V = K^3$ und $W = K^2$. Dann ist die Abbildung
$$T\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a + b \\ c \end{pmatrix}$$
linear, wie man leicht nachrechnet.

(d) Ist $K = \mathbb{R}$ und V der Raum aller stetigen Abbildungen $f : [0,1] \to \mathbb{R}$, dann ist $T : V \to \mathbb{R}$,
$$T(f) = \int_0^1 f(x)\, dx$$
linear (Analysis).

(e) Sei $V = K[x]$ und sei $T : V \to V$ gegeben durch $T(f) = f'$ mit
$$(a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n)' = a_1 + 2a_2 x + \cdots + n a_n x^{n-1}$$

die **formale Ableitung**. Dann ist T linear.

Lemma 2.5.3. *Eine lineare Abbildung schickt die Null auf die Null.*

Beweis. Sei $T : V \to W$ linear. Dann gilt $T(0) = T(0-0) = T(0) - T(0) = 0$. $\square$

Proposition 2.5.4. *Sei $v_1, \dots, v_n$ eine Basis von V. Für jede Menge $\left\{w_1, \dots, w_n\right\}$ in W existiert genau eine lineare Abbildung $T : V \to W$ mit*

$$\begin{aligned} T(v_1) &= w_1 \\ &\vdots \\ T(v_n) &= w_n \end{aligned}$$

Beweis. Seien $w_1, \dots, w_n$ gegeben. Definiere $T : V \to W$ durch

$$T(\lambda_1 v_1 + \cdots + \lambda_n v_n) = \lambda_1 w_1 + \cdots + \lambda_n w_n,$$

für beliebige $\lambda_1, \dots, \lambda_n \in K$. Diese Vorschrift definiert eine Abbildung $T : V \to W$. Wir zeigen, dass diese linear ist. Seien $v, v' \in V$. Schreibe diese in der Basis als

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n, \qquad v' = \lambda_1' v_1 + \cdots + \lambda_n' v_n.$$

Für $\lambda, \mu \in K$ ist dann

$$\begin{aligned} T(\lambda v + \mu v') &= T\Big((\lambda\lambda_1 v_1 + \cdots + \lambda\lambda_n v_n) + (\mu\lambda_1' v_1 + \dots \mu\lambda_n' v_n\Big) \\ &= T\Big((\lambda\lambda_1 + \mu\lambda_1')v_1 + \cdots + (\lambda\lambda_n + \mu\lambda_n')v_n\Big) \\ &= (\lambda\lambda_1 + \mu\lambda_1')w_1 + \cdots + (\lambda\lambda_n + \mu\lambda_n')w_n \\ &= \lambda\Big(\lambda_1 w_1 + \cdots + \lambda_n w_n\Big) + \mu\Big(\lambda_1' w_1 + \cdots + \lambda_n' w_n\Big) \\ &= \lambda T(v) + \mu T(v'). \end{aligned}$$

Für die Eindeutigkeit von T sei S eine weitere lineare Abbildung mit $S(v_j) = w_j$. Für ein beliebiges $v = \lambda_1 v_1 + \cdots + \lambda_n v_n \in V$ rechnen wir mit Hilfe der Linearität von S und der Definition von T:

$$\begin{aligned} S(v) &= S(\lambda_1 v_1 + \cdots + \lambda_n v_n) \\ &= \lambda_1 S(v_1) + \cdots + \lambda_n S(v_n) \\ &= \lambda_1 w_1 + \cdots + \lambda_n w_n \\ &= T(\lambda_1 v_1 + \cdots + \lambda_n v_n) = T(v) \end{aligned}$$ $\square$

Proposition 2.5.5 (Komposition und Umkehrabbildung)**.** (a) *Sind $T : V \to W$ und $S : W \to U$ lineare Abbildungen, so ist die Komposition $S \circ T : V \to U$ linear.*

(b) *Ist $T : V \to W$ linear und bijektiv, so ist die Umkehrabbildung $T^{-1} : W \to V$ ebenfalls linear. In diesem Fall nennen wir T einen* **linearen Isomorphismus***.*

Beweis. (a) Seien $v, v' \in V$. Dann gilt

$$\begin{aligned} S \circ T(v + v') &= S(T(v + v') \\ &= S(T(v) + T(v')) \\ &= S(T(v)) + S(T(v')) \\ &= S \circ T(v) + S \circ T(v'). \end{aligned}$$

Ist außerdem $\lambda \in K$, so gilt

$$S \circ T(\lambda v) = S(T(\lambda v)) = S(\lambda T(v)) = \lambda S(T(v)) = \lambda S \circ T(v).$$

(b) Seien $w, w' \in W$ dann gilt

$$\begin{aligned} T^{-1}(w + w') &= T^{-1}\Big(T(T^{-1}(w)) + T(T^{-1}(w'))\Big) \\ &= T^{-1}\Big(T\Big(T^{-1}(w)) + T^{-1}(w')\Big)\Big) \\ &= T^{-1} \circ T\Big(T^{-1}(w)) + T^{-1}(w')\Big) \\ &= T^{-1}(w) + T^{-1}(w'). \end{aligned}$$

Ist ferner $\lambda \in K$, so gilt

$$\begin{aligned} T^{-1}(\lambda w) &= T^{-1}\Big(\lambda T(T^{-1}(w))\Big) \\ &= T^{-1}\Big(T(\lambda T^{-1}(w))\Big) = T^{-1} \circ T(\lambda T^{-1}(w)) = \lambda T^{-1}(w). \end{aligned}$$

□

Beispiel 2.5.6. Sei V ein endlich-dimensionaler Vektorraum. Die Wahl einer Basis $\mathcal{B} = \big\{v_1, \dots, v_n\big\}$ induziert einen Isomorphismus

$$\begin{aligned} \Phi_{\mathcal{B}} : K^n &\xrightarrow{\cong} V, \\ \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} &\mapsto \lambda_1 v_1 + \dots + \lambda_n v_n. \end{aligned}$$

Umgekehrt liefert jeder Isomorphismus $\phi : K^n \to V$ eine Basis $\phi(e_1), \dots, \phi(e_n)$.

Definition 2.5.7. Sind $S, T : V \to W$ zwei lineare Abbildungen, so definieren wir ihre

Summe durch:

$$S + T : V \to W, \qquad v \mapsto S(v) + T(v).$$

Diese Summe wird auch als die **punktweise Summe** bezeichnet. Für $\lambda \in K$ definieren wir

$$\lambda T : V \to W, \qquad v \mapsto \lambda T(v).$$

Lemma 2.5.8. *Die Summe $S + T$ und das Produkt λT sind wieder lineare Abbildungen.*

Beweis. Für $v, v' \in V$ ist

$$\begin{aligned}(S+T)(v+v') &= S(v+v') + T(v+v') \\ &= S(v) + S(v') + T(v) + T(v') \\ &= S(v) + T(v) + S(v') + T(v') \\ &= (S+T)(v) + (S+T)(v').\end{aligned}$$

Die Beweise der anderen Eigenschaften bestehen ebenso in einer einfachen Auflösung der Definitionen. □

Definition 2.5.9. Sei $T : V \to W$ linear. Der **Kern** von T ist definiert als die Menge

$$\ker T = \{v \in V : T(v) = 0\}.$$

Beispiele 2.5.10. (a) Sei $T : K^2 \to K$ gegeben durch $T\begin{pmatrix} a \\ b \end{pmatrix} = a + b$, dann ist

$$\ker T = \left\{\begin{pmatrix} a \\ b \end{pmatrix} : a + b = 0\right\} = K\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

(b) Sei $T : \mathbb{Q}[x] \to \mathbb{Q}[x]$ die formale Ableitung: $T(f) = f'$. Dann ist $\ker T$ die Menge aller konstanten Polynome.

Satz 2.5.11. (a) *Sei $T : V \to W$ linear. Dann ist der Kern von T ein Untervektorraum von V.*

(b) *Eine lineare Abbildung $T : V \to W$ ist genau dann injektiv, wenn* $\ker T = 0$ *ist.*

Beweis. (a) Seien $v, v' \in \ker T$. Dann gilt

$$T(v + v') = T(v) + T(v') = 0 + 0 = 0,$$

also ist $v + v' \in \ker T$. Ist $\lambda \in K$, so gilt

$$T(\lambda v) = \lambda T(v) = \lambda 0 = 0,$$

also ist auch $\lambda v \in \ker T$.

(b) Sei T injektiv und sei $T(v) = 0 = T(0)$. Aus der Injektivität folgt dann $v = 0$, also ist $\ker T = 0$. Sei umgekehrt $\ker T = 0$ und seien $v, v' \in V$ mit $T(v) = T(v')$. Dann ist $T(v - v') = T(v) - T(v') = 0$, also $v - v' \in \ker T = 0$ und damit $v - v' = 0$, also $v = v'$, so dass T injektiv ist. □

Beispiele 2.5.12. (a) Die lineare Abbildung $T : K^2 \to K^2; \begin{pmatrix} a \\ b \end{pmatrix} \mapsto \begin{pmatrix} b \\ a \end{pmatrix}$ ist injektiv.

(b) Die lineare Abbildung $T : K^2 \to K, \begin{pmatrix} a \\ b \end{pmatrix} \mapsto a + b$ ist nicht injektiv.

Definition 2.5.13. Sei $T : V \to W$ linear. Das **Bild** von T ist die Menge

$$\operatorname{Bild} T = \big\{ T(v) : v \in V \big\}.$$

Beispiele 2.5.14. (a) Sei $T : K^2 \to K^3$ gegeben durch $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \\ x \end{pmatrix}$, dann ist das Bild die Menge aller Vektoren $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ mit $z = x$.

Proposition 2.5.15. *Sei $T : V \to W$ linear. Dann ist das Bild von T ein Untervektorraum von W.*

Beweis. Seien $\lambda, \mu \in K$ und $w, w' \in \operatorname{Bild}(T)$. Dann gibt es $v, v' \in V$ mit $w = T(v)$ und $w' = T(v')$. Es folgt

$$T\big(\lambda v + \mu v'\big) = \lambda T(v) + \mu T(v') = \lambda w + \mu w',$$

also $\lambda w + \mu w' \in \operatorname{Bild}(T)$. □

Satz 2.5.16 (Dimensionsformel für lineare Abbildungen)**.** *Sei $T : V \to W$ linear.*

(a) *Das Bild von T ein Untervektorraum von W.*

(b) *Ist V endlich-dimensional, dann gilt*

$$\dim V = \dim\big(\ker T\big) + \dim\big(\operatorname{Bild} T\big).$$

Insbesondere ist das Bild endlich-dimensional.

Beweis. (a) Das Bild enthält den Nullvektor, also ist es nicht leer. Seien w, w' im Bild, also etwa $w = T(v)$ und $w' = T(v')$, dann ist $w + w' = T(v) + T(v') = T(v + v')$ wieder im Bild. Ist schliesslich $\lambda \in K$, so gilt $\lambda w = \lambda T(v) = T(\lambda v)$, also $\lambda w \in \operatorname{Bild} T$.

(b) Sei $v_1, \dots, v_k$ eine Basis von $\ker T$. Erweitere diese zu einer Basis $v_1, \dots, v_n$ von V. Wir behaupten, dass $T(v_{k+1}), \dots, T(v_n)$ eine Basis von $\operatorname{Bild} T$ ist. Zunächst zeigen wir Unabhängigkeit. Dazu sei $\lambda_{k+1}T(v_{k+1}) + \dots + \lambda_n T(v_n) = 0$. Dann ist $\lambda_{k+1}v_{k+1} + \dots + \lambda_n v_n$ im Kern von T, also ausdrückbar in der Form $\lambda_1 v_1 + \dots + \lambda_k v_k$. Es folgt $\lambda_1 v_1 + \dots \lambda_k v_k + (-\lambda_{k+1})v_{k+1} + \dots + (-\lambda_n)v_n = 0$ und daher wegen der Unabhängigkeit $\lambda_{k+1} = \dots = \lambda_n = 0$, also sind $T(v_{k+1}). \dots, T(v_n)$ unabhängig. Es bleibt zu zeigen, dass sie das Bild erzeugen. Sei also $w \in \operatorname{Bild} T$, also etwa $w = T(v)$ und $v = \mu_1 v_1 + \dots \mu_n v_n$ für geeignete $\mu_j \in K$. Dann ist

$$\begin{aligned} w &= T(\mu_1 v_1 + \dots + \mu_n v_n) \\ &= \underbrace{\mu_1 T(v_1) + \dots + \mu_k T(v_k)}_{=0} + \mu_{k+1}T(v_{k+1}) + \dots + \mu_n T(v_n) \\ &= \mu_{k+1}T(v_{k+1}) + \dots + \mu_n T(v_n). \end{aligned}$$

Also ist $T(v_{k+1}), \dots, T(v_n)$ auch ein Erzeugersystem von $\operatorname{Bild} T$. □

Korollar 2.5.17. *Seien V, W endlich-dimensionale K-Vektorräume.*

(a) *Gilt* $\dim V > \dim W$, *so gibt es keine injektive lineare Abbildung $V \to W$.*

(b) *Gilt* $\dim V < \dim W$, *so gibt es keine surjektive lineare Abbildung $V \to W$.*

Beweis. (a) Sei $T : V \to W$ linear. Dann gilt

$$\begin{aligned} \dim \ker T &= \dim V - \dim \operatorname{Bild} T \\ &\geq \dim V - \dim W > 0. \end{aligned}$$

(b) Mit derselben Notation gilt

$$\dim \operatorname{Bild} T = \dim V - \dim \ker T \leq \dim V < \dim W.$$ □

Satz 2.5.18. *Ist V endlich-dimensional und $T : V \to V$, so sind die folgenden äquivalent:*

(a) *T ist injektiv,*

(b) *T ist surjektiv,*

(c) *T ist bijektiv.*

In diesem Fall gilt insbesondere für jeden Untervektorraum $U \subset V$, dass

$$\dim T(U) = \dim U.$$

Beweis. Wir haben

$$\dim V = \dim\Big(\ker T\Big) + \dim\Big(\operatorname{Bild} T\Big).$$

daher ist

$$\begin{aligned} T \text{ injektiv} &\Leftrightarrow \dim\ker T = 0 \\ &\Leftrightarrow \dim V = \dim \operatorname{Bild} T \\ &\Leftrightarrow T \text{ surjektiv.} \end{aligned}$$

Für den Zusatz beachte, dass $\ker(T) = 0$ und daher ist nach der Dimensionsformel $\dim U = \dim T(U) + \dim\ker(T|_U) = \dim T(U)$. □

* * *

2.6 Matrizen

Definition 2.6.1. Seien V, W endlich-dimensionale Vektorräume mit Basen $\mathcal{A} = (v_1, \dots, v_n)$ und $\mathcal{B} = (w_1, \dots, w_m)$. Sei $T : V \to W$ linear. Dann existieren eindeutig bestimmte Zahlen $a_{i,j} \in K$ mit

$$T(v_j) = a_{1,j}w_1 + \cdots + a_{m,j}w_m$$

für $j = 1, \dots, n$. Wir schreiben diese Körperelemente in ein rechteckiges Schema, eine **Matrix**

$$\begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix}.$$

Wir schreiben $\mathrm{M}_{m\times n}(K)$ für die Menge aller Matrizen über K mit m Zeilen und n Spalten. Ist $m = n$, so schreiben wir auch $\mathrm{M}_n(K)$ für $M_{n\times n}(K)$. Wir erhalten also zu jeder linearen Abbildung $T : V \to W$ eine Matrix $\mathbf{M} = \mathbf{M}(T)$. Diese hängt allerdings von der

Wahl der Basen ab, also schreiben wir besser

$$\mathbf{M}_{\mathcal{A},\mathcal{B}}(T).$$

Definition 2.6.2. Auf der Menge der Matrizen $\mathrm{M}_{m\times n}(K)$ definieren wir eine Addition

$$\begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix} + \begin{pmatrix} b_{1,1} & \dots & b_{1,n} \\ \vdots & & \vdots \\ b_{m,1} & \dots & b_{m,n} \end{pmatrix} = \begin{pmatrix} a_{1,1}+b_{1,1} & \dots & a_{1,n}+b_{1,n} \\ \vdots & & \vdots \\ a_{m,1}+b_{m,1} & \dots & a_{m,n}+b_{m,n} \end{pmatrix}$$

und eine skalare Multiplikation

$$\lambda \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix} = \begin{pmatrix} \lambda a_{1,1} & \dots & \lambda a_{1,n} \\ \vdots & & \vdots \\ \lambda a_{m,1} & \dots & \lambda a_{m,n} \end{pmatrix}$$

Mit diesen Operationen wird $\mathrm{M}_{m\times n}(K)$ ein K-Vektorraum. Jede Durchnummerierung der Einträge liefert einen linearen Isomorphismus $\mathrm{M}_{m\times n}(K) \xrightarrow{\cong} K^{mn}$. Wir folgern also

$$\dim \mathrm{M}_{m\times n}(K) = mn.$$

Proposition 2.6.3. *Seien Basen $(v_1,\dots,v_n)$ und $(w_1,\dots,w_m)$ der Vektorräume V und W gegeben. Dann ist die Abbildung*

$$\mathbf{M} : \mathrm{Lin}(V,W) \to \mathrm{M}_{m\times n}(K),$$

die jeder linearen Abbildung ihre Matrix zuordnet, ein linearer Isomorphismus.

Beweis. Wir zeigen zunächst, dass M linear ist. Seien $\mathbf{M}(T) = (a_{i,j})$ und $\mathbf{M}(S) = (b_{i,j})$. Wir wollen zeigen, dass $\mathbf{M}(S+T) = \mathbf{M}(S) + \mathbf{M}(T)$. Sei hierzu $1 \le j \le n$. Dann gilt

$$\begin{aligned} (T+S)(v_j) &= T(v_j) + S(v_j) \\ &= a_{1,j}w_1 + \dots + a_{m,j}w_m \\ &\quad + b_{1,j}w_1 + \dots + b_{m,j}w_m \\ &= (a_{1,j}+b_{1,j})w_1 + \dots + (a_{m,j}+b_{m,j})w_m. \end{aligned}$$

Hieraus folgt $\mathbf{M}(S+T) = \mathbf{M}(S) + \mathbf{M}(T)$. Die Gleichung $\mathbf{M}(\lambda T) = \lambda\mathbf{M}(T)$ zeigt man ebenso. Die Abbildung M ist injektiv, denn ist $\mathbf{M}(T) = 0$, so folgt

$$T(v_j) = a_{1,j}w_1 + \dots + a_{m,j}w_m = 0$$

und damit $T = 0$. Schliesslich ist M surjektiv, denn sei $A = (a_{i,j})$ irgendeine Matrix, dann definiert nach Proposition 2.5.4 die Vorschrift

$$T(v_j) = a_{1,j}w_1 + \cdots + a_{m,j}w_m$$

eine lineare Abbildung T. Für diese gilt $\mathbf{M}(T) = A$. □

* * *

2.7 Matrixmultiplikation

Proposition 2.7.1. *Seien U, V, W endlich-dimensionale Vektorräume. Fixiere Basen von U, V und W. Seien $S : U \to V$ und $T : V \to W$. Schreibe $A = \mathbf{M}(T)$ und $B = \mathbf{M}(S)$. Für die Matrix $C = \mathbf{M}(T \circ S)$ gilt dann*

$$\boxed{C_{i,k} = \sum_{j=1}^{n} A_{i,j}B_{j,k},}$$

wobei $n = \dim V$.

Beweis. Seien die Basen von U, V und W mit $u_1, \ldots, u_p$, sowie $v_1, \ldots, v_n$ und $w_1, \ldots, w_m$ bezeichnet. Dann gilt

$$\begin{aligned} T \circ S(u_k) &= T(S(u_k)) = T\Big(\sum_{j=1}^{n} B_{j,k}v_j\Big) \\ &= \sum_{j=1}^{n} B_{j,k}T(v_j) = \sum_{j=1}^{n} B_{j,k}\sum_{i=1}^{m} A_{i,j}w_i = \sum_{i=1}^{m} \underbrace{\sum_{j=1}^{n} A_{i,j}B_{j,k}}_{=C_{i,k}} w_i. \end{aligned}$$ □

Definition 2.7.2. Wir definieren die **Matrixmultiplikation**

$$\mathrm{M}_{m\times n}(K) \times \mathrm{M}_{n\times p}(K) \to \mathrm{M}_{m\times p}(K)$$

durch die Vorschrift

$$(A, B) \mapsto C$$

mit

$$\boxed{C_{i,j} = \sum_{k=1}^{n} A_{i,k}B_{k,j}.}$$

Die Matrixmultiplikation wurde also so gemacht, dass für komponierbare Abbildungen S und T gilt

$$\mathbf{M}(T \circ S) = \mathbf{M}(T)\mathbf{M}(S).$$

Beispiele 2.7.3. (a)

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} ax+bz & ay+bw \\ cx+dz & cy+dw \end{pmatrix}.$$

Insbesondere

$$\begin{pmatrix} a & \\ & d \end{pmatrix}\begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} ax & ay \\ dz & dw \end{pmatrix}$$

und

$$\begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix}\begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} x+z & y+w \\ z & w \end{pmatrix}.$$

(b)

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & j \end{pmatrix}\begin{pmatrix} x & y & z \\ u & v & w \\ r & s & t \end{pmatrix} = \begin{pmatrix} ax+bu+cr & ay+bv+cs & az+bw+ct \\ dx+eu+fr & dy+ev+fs & dz+ew+ft \\ gx+hu+jr & gy+hv+js & gz+hw+jt \end{pmatrix}.$$

(c) Eine Diagonalmatrix multipliziert die Zeilen mit den Diagonaleinträgen, also

$$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}\begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{pmatrix}\begin{pmatrix} \lambda_1 a_{1,1} & \dots & \lambda_1 a_{1,n} \\ \vdots & & \vdots \\ \lambda_n a_{n,1} & \dots & \lambda_n a_{n,n} \end{pmatrix}$$

(d) Sei $A \in \mathrm{M}_{m\times n}(K)$. Die Abbildung $K^n \to K^m$ definiert durch $x \mapsto Ax$, also

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{1,1}x_1 + \cdots + a_{1,n}x_n \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n \end{pmatrix}$$

ist eine lineare Abbildung. Der Kern dieser Abbildung ist die Lösungsmenge des linearen Gleichungssystems

$$\begin{aligned} a_{1,1}x_1 + \cdots + a_{1,n}x_n &= 0 \\ &\vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n &= 0. \end{aligned}$$

Proposition 2.7.4. *Die Matrixmultiplikation ist assoziativ und distributiv, d.h. es gilt*

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

und

$$\begin{aligned} A(B+B') &= AB + AB' \\ (A+A')B &= AB + A'B \end{aligned}$$

für alle Matrizen A, A', A, B', C mit den passenden Formaten, so dass die Produkte existieren.

Beweis. Mit Hilfe von Proposition 2.6.3 folgt dies sofort aus den entsprechenden Eigenschaften linearer Abbildungen und der Tatsache, dass $\mathbf{M}(T \circ S) = \mathbf{M}(T)\mathbf{M}(S)$. □

Proposition 2.7.5. *Jede lineare Abbildung $K^n \to K^n$ ist von der Form $x \mapsto Ax$ für eine eindeutig bestimmte Matrix $A \in \mathrm{M}_n(K) = \mathrm{M}_{n \times n}(K)$.*

Beweis. Klar nach Proposition 2.6.3. □

Definition 2.7.6. Wir nennen eine Matrix $A \in \mathrm{M}_n(K)$ **invertierbar**, falls die induzierte lineare Abbildung $K^n \to K^n$ bijektiv ist. Die Umkehrabbildung ist dann ebenfalls durch eine Matrix gegeben, die wir mit A^{-1} bezeichnen. Wir schreiben $\mathrm{GL}_n(K)$ für die Menge aller invertierbaren Matrizen in $\mathrm{M}_n(K)$.

Proposition 2.7.7. *Eine Matrix $A \in \mathrm{M}_n(K)$ ist genau dann invertierbar, wenn es eine Matrix $A^{-1} \in \mathrm{M}_n(K)$ gibt mit*

$$AA^{-1} = I \quad \text{oder} \quad A^{-1}A = I,$$

wobei

$$I = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$$

die **Einheitsmatrix** *ist.*

Beweis. A^{-1} ist gerade die Matrix der Umkehrabbildung. □

Proposition 2.7.8. *Die Menge $\mathrm{GL}_n(K)$ aller invertierbaren Matrizen wird mit der Matrixmultiplikation eine* **Gruppe**.

Proof. Die Matrixmultiplikation is assoziativ, es gibt ein neutrales Element $I = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$ und Invertierbarkeit besagt, dass es inverse Elemente gibt. □

Definition 2.7.9. Für eine Matrix $A \in \mathrm{M}_{m \times n}(K)$ sei die **transponierte Matrix** $A^t \in \mathrm{M}_{n \times n}(K)$ definiert durch

$$A^t = \begin{pmatrix} A_{1,1} & \dots & A_{m,1} \\ \vdots & & \vdots \\ A_{1,n} & \dots & A_{m,n} \end{pmatrix},$$

die Matrix wird also an der Diagonalen gespiegelt, oder, was dasselbe ist, die Indizes i, j werden vertauscht, also

$$A^t_{i,j} = A_{j,i}.$$

Beispiele 2.7.10. (a) $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^t = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$,

(b) $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^t = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$.

(c) $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}^t = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$.

Proposition 2.7.11. *Für* $A, A' \in \mathrm{M}_{m\times n}(K)$ *und* $B \in \mathrm{M}_{n,p}(K)$ *gilt*

(a) $(A + A')^t = A^t + (A')^t, \qquad (A^t)^t = A,$

(b) $(AB)^t = B^t A^t,$

(c) *Ist* $m = n$ *so ist* A *genau dann invertierbar, wenn* A^t *dies ist und dann ist*

$$(A^t)^{-1} = (A^{-1})^t.$$

Beweis. (a) ist klar.

(b) Es gilt

$$(AB)^t_{i,j} = (AB)_{j,i} = \sum_k A_{j,k} B_{k,i} = \sum_j A^t_{k,j} B^t_{i,k} = \sum_j B^t_{i,k} A^t_{k,j} = (B^t A^t)_{i,j}.$$

(c) Ist A invertierbar, so rechnen wir

$$I = I^t = \left(A^{-1}A\right)^t = A^t (A^{-1})^t,$$

also ist auch A^t invertierbar und ie behauptete Formel gilt. Im anderen Fall vertauschen wir die Rollen von A und A^t. □

Beispiel 2.7.12. Es gilt

$$\begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ & 1 \end{pmatrix} = I$$

und

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

Beispiel 2.7.13. Wir wollen an einem Beispiel zeigen, wie man die Matrix zu einer linearen Abbildung bestimmt. Hierzu sei V der Raum aller Polynome vom Grad ≤ 3 und $T : V \to V$ sei die formale Ableitung, also

$$T(a_0 + a_1 x + a_2 x^2 + a_3 x^3) = a_1 + 2a_2 x + 3a_3 x^2.$$

Wir bestimmen die Matrix $A = \mathrm{M}_{\mathcal{A}}(T)$ bezüglich der Basis $\mathcal{A} = (1, x, x^2, x^3)$. Der erste Basisvektor 1 wird auf die Null geworfen, deshalb ist die erste Spalte von A gleich Null. Der zweite Basisvektor wird auf den ersten geworfen, deshalb ist die zweite Spalte $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$. Der dritte Basisvektor wird auf das 2-fache des zweiten geworfen, also ist die dritte Spalte $\begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}$. Der vierte Basisbvektor x^3 wird auf das dreifache des dritten geworfe, also ist die vierte Spalte gleich $\begin{pmatrix} 0 \\ 0 \\ 3 \\ 0 \end{pmatrix}$ Damit ist die Matrix gleich

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

* * *

2.8 Basiswechsel

Definition 2.8.1. Seien also $\mathcal{B} = (v_1, \dots, v_n)$ und $\mathcal{D} = (v_1', \dots, v_n')$ Basen von V. Wir definieren die **Basiswechselmatrix** $S = S_{\mathcal{B},\mathcal{D}}$ als die Matrix $S_{\mathcal{B},\mathcal{D}} = (s_{j,k})_{1 \le j,k \le n}$, die durch

$$v_k' = \sum_{j=1}^n s_{k,j} v_j.$$

gegeben ist. Man drückt also jeden Vektor der einen Basis durch die andere aus und packt die Koeffizienten in eine Matrix.Man kann dies auch in der Form

$$S \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} v_1' \\ \vdots \\ v_n' \end{pmatrix}$$

ausdrücken, wenn man akzeptiert, dass man in einen Spaltenvektor auch Vektoren hineinschreiben darf. Mit dieser Konvention kann man also $S\mathcal{B} = \mathcal{D}$ schreiben.

Lemma 2.8.2. *Jede Basiswechselmatrix S ist invertierbar. Gilt $S\mathcal{B} = \mathcal{D}$, dann ist $\mathcal{B} = S^{-1}\mathcal{D}$.*

Proof. Es sei $S = (s_{i,j})$ gegeben durch

$$v'_k = \sum_{j=1}^{n} s_{k,j} v_j.$$

In der umgekehrten Richtung sei die Matrix $R = (r_{i,j})$ gegeben durch

$$v_k = \sum_{j=1}^{n} r_{k,j} v'_j.$$

Dann gilt für jedes $1 \leq k \leq n$

$$v_k = \sum_{j=1}^{n} r_{k,j} v'_j = \sum_{j=1}^{n} r_{k,j} \sum_{i=1}^{n} s_{j,i} v_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} r_{k,j} s_{j,i} \right) v_i.$$

Da die Linearkombination von v_k eindeutig ist, folgt

$$\left(\sum_{j=1}^{n} r_{k,j} s_{j,i} \right) = \begin{cases} 1 & k = i, \\ 0 & k \neq i. \end{cases}$$

Dies ist gleichbedeutend mit

$$RS = I.$$ □

Bemerkung 2.8.3. Nach Beispiel 2.5.6 induziert jede Basis $\mathcal{B}$ einen Isomorphismus $\Phi_{\mathcal{B}} : K^n \to V$. Ein Basiswechsel induziert dann ein kommutatives Diagramm

$$\begin{tikzcd} K^n \arrow[r, "\Phi_{\mathcal{B}}"] \arrow[d, "S"'] & V \\ K^n \arrow[ur, "\Phi_{\mathcal{D}}"'] & \end{tikzcd}$$

wobei $S = S_{\mathcal{B},\mathcal{D}}$ die Basiswechselmatrix ist.

Definition 2.8.4. Zwei Matrizen $A, B \in \mathrm{M}_n(K)$ heissen **konjugiert**, falls es eine invertierbare Matrix $S \in \mathrm{GL}_n(K)$ gibt, so dass

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

Satz 2.8.5 (Basiswechselsatz). *Sei $T : V \to V$ und seien $\mathcal{B}, \mathcal{D}$ Basen von V. Seien $B = M_{\mathcal{B}}(T)$ und $D = M_{\mathcal{D}}(T)$ die Matrizen, die T in den beiden Basen darstellen. Sei $S = S_{\mathcal{B},\mathcal{D}}$ die Basiswechselmatrix. Dann gilt*

$$D = SBS^{-1}.$$

Das heisst, Matrizen von T zu verschiedenen Basen sind konjugiert.

Proof. Seien $\mathcal{B} = (v_1, \dots, v_n)$ und $\mathcal{D} = (w_1, \dots, w_n)$. Dann gilt

$$w_k = \sum_{j=1}^{n} s_{k,j} v_j, \qquad v_l = \sum_{i=1}^{n} r_{l,i} w_i,$$

wobei $S^{-1} = (r_{i,j})$.

Sei $B = (b_{i,j})$ und $D = (d_{i,j})$. Dann haben wir $Tv_j = \sum_{l=1}^{n} b_{j,l} v_l$. Wir erhalten einerseits

$$Tw_k = \sum_{j=1}^{n} d_{k,j} w_j = \sum_{j=1}^{n} d_{k,j} \sum_{l=1}^{n} s_{j,l} v_l = \sum_{l=1}^{n} (DS)_{k,l} v_l$$

und andererseits

$$Tw_k = T\left(\sum_{j=1}^{n} s_{k,j} v_j\right) = \sum_{j=1}^{n} s_{k,j} T v_j = \sum_{j=1}^{n} s_{k,j} \sum_{l=1}^{n} b_{j,l} v_l = \sum_{l=1}^{n} (SB)_{k,l} v_l.$$

Da die v_l unabhängig sind, folgt $DS = SB$ wie verlangt. □

* * *

2.9 Gauß-Verfahren

Ab jetzt rechnen wir nur noch mit quadratischen Matrizen. Für unsere Zwecke reicht dies, denn man kann jede Matrix durch Nullen zu einer quadratischen Matrix auffüllen. Wir schreiben

$$\mathrm{M}_n(K)$$

für die Menge der $n \times n$ Matrizen über K.

Definition 2.9.1. Sei $B \in \mathrm{M}_n(K)$. Eine **Zeilentransformation** ist eine der folgenden Operationen

1. Für $\lambda \in K$ addiere das λ-fache der j-ten Zeile zur i-ten. $(j \neq i)$
2. Multipliziere die j-te Zeile mit $\mu \in K^{\times}$.
3. Vertausche zwei Zeilen.

Beispiele 2.9.2. (a) $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 \\ 5 & 8 \end{pmatrix}$, (das doppelte der ersten zur zweiten addiert),

(b) $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & 4 \\ 3 & 4 \end{pmatrix}$, (die erste mit 2 multipliziert),

(c) $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix}$, (die Zeilen vertauscht).

Die erste Operation ist gegeben durch $B \rightsquigarrow A_{i,j}(\lambda)B$, wobei

$$A_{i,j}(\lambda) = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & \lambda & \ddots & \\ & & & 1 \end{pmatrix},$$

mit einem λ in der i-ten Zeile und j-ten Spalte und Einsen auf der Diagonalen.

Operation 2. ist gegeben durch $B \rightsquigarrow M_i(\mu)B$, wobei

$$M_i(\mu) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \mu & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix},$$

mit μ in der i-ten Zeile und Spalte.

Die dritte Operation ist gegeben durch $B \rightsquigarrow S_{i,j}B$, wobei

$$S_{i,j} = \begin{pmatrix} 1 & & & & \\ & 0 & & 1 & \\ & & 1 & & \\ & 1 & & 0 & \\ & & & & 1 \end{pmatrix},$$

mit den nichtdiagonalen Einsen in der i-ten Zeile und j-ten Spalte und umgekehrt.

Die Matrizen $A_{i,j}(\lambda), M_i(\mu), S_{i,j}$ werden **Elementarmatrizen** genannt. Es gilt also:

Zeilentransformationen = Linksmultiplikationen mit Elementarmatrizen.

Lemma 2.9.3. *Die Elementarmatrizen sind alle invertierbar. Wenn also eine Matrix B aus einer Matrix A durch wiederholte Zeilentransformationen hervorgeht, existiert eine invertierbare Matrix T mit $B = TA$.*

Beweis. Man rechnet nach:

$$\begin{aligned} A_{i,j}(\lambda)A_{i,j}(-\lambda) &= I, \\ M_i(\mu)M_i(\mu^{-1}) &= I, \\ S_{i,j}S_{i,j} &= I. \end{aligned}$$

□

Definition 2.9.4. Eine Matrix $A \in \mathrm{M}_n(K)$ heißt **obere Dreiecksmatrix**, wenn A unterhalb der Diagonale nur Nullen hat, wenn also gilt $A_{i,j} = 0$ für $i > j$, d.h. wenn A von der Form

$$A = \begin{pmatrix} A_{1,1} & & * \\ & \ddots & \\ 0 & & A_{n,n} \end{pmatrix}$$

ist.

Satz 2.9.5. *Jede Matrix kann durch wiederholte Zeilentransformationen auf obere Dreiecksform gebracht werden, wobei man außerdem erreichen kann, dass auf der Diagonalen nur Nullen und Einsen stehen. Man sagt dazu, dass man A in* **Zeilenstufenform** *bringen kann.*

Beweis. Ist die erste Spalte von A gleich Null, so kann man induktiv mit der Untermatrix A' weitermachen, für die

$$A = \begin{pmatrix} 0 & * \\ 0 & A' \end{pmatrix}$$

gilt. Ist die erste Spalte ungleich Null, so kann man durch Zeilenvertauschung erreichen, dass $a_{1,1} \neq 0$. Dann multipliziert man die erste Zeile mit $a_{1,1}^{-1}$ und erreicht $a_{1,1} = 1$. Subtrahiere dann das $a_{j,1}$-fache der ersten Zeile von der j-ten für $j = 2, \dots, n$

und erreiche

$$A = \begin{pmatrix} 1 & * & \dots & * \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}.$$

Mach nun Induktiv mit A_1 weiter. □

Beispiel 2.9.6.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 1 & 2 \\ 3 & 4 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & -7 & -10 \\ 3 & 4 & 1 \end{pmatrix}$$
$$\rightsquigarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 10/7 \\ 3 & 4 & 1 \end{pmatrix}$$
$$\rightsquigarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 10/7 \\ 0 & -2 & -8 \end{pmatrix}$$
$$\rightsquigarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 10/7 \\ 0 & 0 & (20/7) - 8 \end{pmatrix}$$
$$\rightsquigarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 10/7 \\ 0 & 0 & 1 \end{pmatrix}$$

Proposition 2.9.7. *Für jede Matrix $A \in \mathrm{M}_n(K)$ gibt es invertierbare Matrizen S, T, die Produkte von Elemenentarmatrizen sind, so dass*

$$A = S\begin{pmatrix} I & \\ & 0 \end{pmatrix}T,$$

wobei I die $k \times k$ Einheitsmatrix ist für ein $0 \leq k \leq n$.

Proof. Für $n = 1$ ist die Behauptung klar. Ist E eine Elementarmatrix, dann ist $(E^tA^t)^t = AE$ und daher entspricht mit E von rechts einer **Spaltentransformation**, die man analog zu den Zeilentransformationen definiert. Ist $A = 0$, ist nichts zu zeigen. Ist $A \neq 0$, dann kann man durch Spaltenvertauschung erreichen, dass die erste Spalte $\neq 0$ ist. Danach kann man durch Zeilentausch $a_{1,1} \neq 0$ erreichen. Durch weitere Zeilentrafos erreicht man, dass die erste Spalte gleich e_1 ist und dann reduziert man durch Spaltentrafos auf den Fall, dass $A = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & A' \end{array}\right)$. Induktiv gilt die Behauptung für A' und damit folgt die Proposition. □

Definition 2.9.8. Verfahren zum Finden aller Lösungen eines Gleichungssystems $Ax = b$ mit gegebenen $A \in M_n(K)$ und $b \in K^n$:

1. Bringe A in Zeilenstufenform A'. und führe alle Transformationen gleichzeitig an dem Vektor b aus. Notiere die ausgeführten Transformationen in der Weise, dass das entsprechende Produkt von Elementarmatrizen S notiert wird.

$$A \rightsquigarrow A' = SA,$$
$$b \rightsquigarrow b' = Sb.$$

2. Löse das System $A'x = b'$. Das ist vergleichsweise einfach. Dann ist jede Lösung des modifizierten Systems auch eine des ursprünglichen Systems, denn

$$A'x = b' \Leftrightarrow SAx = SB \Leftrightarrow Ax = b,$$

da S invertierbar ist. Dieses Verfahren wird das **Gauß-Verfahren** genannt.

Beispiel 2.9.9. Löse $Ax = b$ mit $K = \mathbb{Q}$ und $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, sowie $b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. Dies machen wir wie folgt:

$$\left(\begin{array}{cc|c} 1 & 2 & 1 \\ 3 & 4 & 2 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|c} 1 & 2 & 1 \\ 0 & -2 & -1 \end{array}\right)$$
$$\rightsquigarrow \left(\begin{array}{cc|c} 1 & 2 & 1 \\ 0 & 1 & 1/2 \end{array}\right).$$

Das modifizierte Gleichungssystem ist

$$x_1 + 2x_2 = 1$$
$$x_2 = \frac{1}{2}$$

Der einzige Lösungsvektor ist $x = \begin{pmatrix} 0 \\ 1/2 \end{pmatrix}$.

Definition 2.9.10. Ein **affiner Unterraum** eines Vektorraums V ist eine Teilmenge der Gestalt

$$S = v_0 + U$$

für einen linearen Unterraum U.

Beispiel 2.9.11. Jede Gerade in der Ebene $\mathbb{R}^2$ oder im Raum $\mathbb{R}^3$ ist ein affiner Unterraum. Ein affiner Unterraum ist genau dann ein Untervektorraum, wenn er die Null enthält.

Jede Matrix $A \in \mathrm{M}_n(K)$ liefert eine lineare Abbildung $K^n \to K^n$. Wir identifizieren die

Matrix mit dieser Abbildung und schreiben

$$\begin{aligned}\ker A &= \left\{x \in K^n : Ax = 0\right\}\\ \operatorname{Bild} A &= \left\{Ax : x \in K^n\right\}.\end{aligned}$$

Satz 2.9.12. *Ein lineares Gleichungssystem $Ax = b$ ist genau dann lösbar, wenn $b \in \operatorname{Bild} A$. Ist dies der Fall, so ist die Lösungsmenge $L = \left\{x \in K^n : Ax = b\right\}$ ein affiner Unterraum*

$$L = v_0 + U,$$

wobei $U = \ker A$ und v_0 ein beliebiges Element aus L ist.

Beweis. Die erste Aussage ist offensichtlich. Sei $b \in \operatorname{Bild} A$. Dann existiert ein $v_0 \in K^n$ mit $Av_0 = b$. Sei $U = \ker A$. Wir zeigen $L = v_0 + U$.

"$\subset$" Sei $v \in L$, also $Av = b$. Sei $u = v - v_0$, dann folgt $Au = Av - Av_0 = b - b = 0$, also ist $u \in U$ und damit ist $v = v_0 + u \in v_0 + U$.
"$\supset$" Sei $u \in U$, dann gilt $A(v_0 + u) = Av_0 + Au = b + 0 = b$, also ist $v_0 + u \in L$. □

Beispiel 2.9.13. Sei $K = \mathbb{Q}$ und $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 2 & 1 \end{pmatrix}$, sowie $b = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$. Wir lösen das System mit Zeilentransformationen:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 2 & 4 & 6 & 2 \\ 3 & 2 & 1 & -1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -4 & -8 & -4 \end{array}\right)$$
$$\rightsquigarrow \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Basislösung $v_0 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$. Ferner ist

$$U = \ker A = \mathbb{Q}\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix},$$

also

$$L = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + \mathbb{Q}\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

Lemma 2.9.14. *Eine quadratische Matrix der Form* $S = \left(\begin{array}{c|c} A & B \\ \hline 0 & D \end{array}\right)$ *ist genau dann invertierbar, wenn A und D es sind. In diesem Fall gilt*

$$\left(\begin{array}{c|c} A & B \\ \hline 0 & D \end{array}\right)^{-1} = \left(\begin{array}{c|c} A^{-1} & B' \\ \hline 0 & D^{-1} \end{array}\right)$$

für eine Matrix B'.

Beweis. Sei S invertierbar. Dann muss A invertierbar sein, denn ist $v \in \ker A$, dann ist $\begin{pmatrix} v \\ 0 \end{pmatrix}$ im Kern von S, also ist $v = 0$. Ist dann $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ die inverse Matrix, dann folgt $I = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}\left(\begin{array}{c|c} A & B \\ \hline 0 & D \end{array}\right) = \left(\begin{array}{c|c} * & * \\ \hline \gamma A & * \end{array}\right)$. Damit ist $\gamma = 0$ und $\alpha A = I$, sowie $\delta D = I$ wie verlangt. Seien umgekehrt A, D invertierbar, dann gilt

$$\left(\begin{array}{c|c} A & B \\ \hline 0 & D \end{array}\right)\left(\begin{array}{c|c} A^{-1} & X \\ \hline 0 & D^{-1} \end{array}\right) = \left(\begin{array}{c|c} I & AX + BD^{-1} \\ \hline 0 & I \end{array}\right).$$

Setzt man also $X = -A^{-1}BD^{-1}$, so ist $\left(\begin{array}{c|c} A^{-1} & X \\ \hline 0 & D^{-1} \end{array}\right)$ eine Inverse zu S. □

Korollar 2.9.15. *Eine obere Dreiecksmatrix ist genau dann invertierbar, wenn alle Diagonalelemente* $\neq 0$ *sind und es gilt dann*

$$\begin{pmatrix} \lambda_1 & & * \\ 0 & \ddots & \\ & & \lambda_n \end{pmatrix}^{-1} = \begin{pmatrix} \lambda_1^{-1} & & * \\ 0 & \ddots & \\ & & \lambda_n^{-1} \end{pmatrix}$$

Proof. □

Satz 2.9.16. *Ist die Matrix* $A \in \mathrm{M}_n(K)$ *invertierbar, dann kann die erweiterte Matrix* $(A, I) \in \mathrm{M}_{n,2n}(K)$ *durch Zeilentransformationen in die Form* (I, B) *gebracht werden. Dann ist* $B = A^{-1}$ *die Inverse zu A.*

Beweis. Sei S ein Produkt der Elementarmatrizen, die A auf eine Zeilenstufenform D bringen. Da $D = SA$ und A, S invertierbar, ist auch D invertierbar. Da D auf Zeilenstufenform ist, sind alle Diagonaleinträge gleich 1. Nun kann man durch weitere Zeilentransformationen die Matrix zur Einheitsmatrix machen, hat also $SA = I$, damit ist $S = A^{-1}$. □

Folgerung. Jede invertierbare Matrix ist ein Produkt von Elementarmatrizen.

Beispiel 2.9.17. Bestimme die Inverse zu $A = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$ über $K = \mathbb{Q}$.

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -1 & -1 & 1 \end{array}\right)$$
$$\rightsquigarrow \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & -1 \end{array}\right)$$
$$\rightsquigarrow \left(\begin{array}{cc|cc} 1 & 0 & -1 & 2 \\ 0 & 1 & 1 & -1 \end{array}\right).$$

In der Tat rechnent man nach, dass

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

ist und damit $A^{-1} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$.

Beispiel 2.9.18. Bestimme die Inverse zu $S = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$. Hierzu rechnen wir

$$\left(\begin{array}{ccc|ccc} 0 & 1 & 1 & 1 & & \\ 1 & 0 & 1 & & 1 & \\ 1 & 1 & 0 & & & 1 \end{array}\right)$$
$$\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & & 1 & \\ 0 & 1 & 1 & 1 & & \\ 1 & 1 & 0 & & & 1 \end{array}\right)$$
$$\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & & 1 & \\ 0 & 1 & 1 & 1 & & \\ 0 & 1 & -1 & & -1 & 1 \end{array}\right)$$
$$\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & & 1 & \\ 0 & 1 & 1 & 1 & & \\ 0 & 0 & -2 & -1 & -1 & 1 \end{array}\right)$$
$$\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & & 1 & \\ 0 & 1 & 1 & 1 & & \\ 0 & 0 & 1 & 1/2 & 1/2 & -1/2 \end{array}\right)$$
$$\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & & 1 & \\ 0 & 1 & 0 & 1/2 & -1/2 & 1/2 \\ 0 & 0 & 1 & 1/2 & 1/2 & -1/2 \end{array}\right)$$
$$\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1/2 & 1/2 & 1/2 \\ 0 & 1 & 0 & 1/2 & -1/2 & 1/2 \\ 0 & 0 & 1 & 1/2 & 1/2 & -1/2 \end{array}\right)$$

Damit ist $S^{-1} = \frac{1}{2}\begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$, was man leicht überprüft.

* * *

2.10 Rang einer Matrix

Definition 2.10.1. Ist $A \in \mathrm{M}_{m\times n}(K)$ eine Matrix, so definieren wir den **Spaltenrang** von A als

$$\mathrm{SRang}(A) = \dim \mathrm{Spann}(s_1,\dots,s_n) = \dim \mathrm{Bild}(A).$$

Analog definiere den **Zeilenrang**

$$\mathrm{ZRang}(A) = \dim \mathrm{Spann}(z_1,\dots,z_m),$$

wobei die z_j die Zeilen von A sind.

Es gilt dann

$$\mathrm{ZRang}(A) = \mathrm{SRang}(A^t) = \dim \mathrm{Bild}(A^t).$$

Lemma 2.10.2. (a) *Sei $A \in \mathrm{M}_{m\times n}(K)$. Wir schreiben die lineare Abbildung $K^n \to K^m$, $x \to Ax$ ebenfalls als A. Dann gilt*

$$\mathrm{Bild}(A) = \mathrm{Spann}(s_1,\dots,s_n),$$

wobei die s_j die Spalten der Matrix A sind.

(b) *Sei A eine Matrix. Sei $B = (A, 0)$ eine Matrix, die durch hinzufügen von Nullspalten, also der Form $\begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix}$ entsteht. Dann gilt*

$$\mathrm{SRang}(A) = \mathrm{SRang}(B) \quad \textit{und} \quad \mathrm{ZRang}(A) = \mathrm{ZRang}(B).$$

Dasselbe gilt bei Hinzufügen von Nullzeilen.

Proof. (a) Die Spalten sind die Bilder der Standard Basisvektoren $e_1,\dots,e_n$, damit spannen sie das Bild auf.

(b) Der Spaltenrang ändert sich nicht durch Hinzufügen von Nullen. Für den Zeilenrang, beachte dies: ist $V \subset K^n$ der Spann der Zeilen von A, dann wirft die Abbildung $K^n \to K^{n+K}$, $x \mapsto (x, 0)$ den Raum V isomorpoh auf den Zeilenspann von B). Dieselbe Aussage für Nullzeilen erhält man durch Transposition. □

Satz 2.10.3. *Für jede Matrix ist Spaltenrang gleich Zeilenrang, also*

$$\mathrm{SRang}(A) = \mathrm{ZRang}(A), \qquad A \in \mathrm{M}_n(K).$$

Beweis. Indem wir durch Nullen auffüllen, koennen wir $m = n$ annehmen. Nach Proposition 2.9.7 schreiben wir $A = SDT$ mit $D = \left(\begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array}\right)$, wobei I die $k \times k$ Einheitsmatrix ist und S, T invertierbar sind. Mit Satz 2.5.18 folgt dann

$$\mathrm{SRang}(A) = \dim \mathrm{Bild}(SDT) = \dim SDT(K^n) = \dim SD(K^n) = \dim D(K^n) = k$$

und

$$\mathrm{ZRang}(A) = \dim \mathrm{Bild}(A^t) = \dim \mathrm{Bild}(T^t DS^t) = k. \qquad \square$$

Definition 2.10.4. Ist $T : V \to W$ linear, so definieren wir den **Rang** von T als

$$\mathrm{Rang}(T) = \dim \mathrm{Bild}(T).$$

Fassen wir eine Matrix als lineare Abbildung auf, ist der Rang gleich dem Spaltenrang, also auch gleich dem Zeilenrang.

Proposition 2.10.5. *Für eine Matrix $A \in \mathrm{M}_n(K)$ sind äquivalent*

(a) *A ist invertierbar.*

(b) $\mathrm{Rang}(A) = n$.

Proof. (a)⇒(b): Ist A invertierbar, dann ist A surjektiv, also
$\mathrm{Rang}(A) = \dim \mathrm{Bild}(A) = \dim K^n = n$.
(b⇒(a): Ist $\mathrm{Rang}(A) = n$, dann ist A surjektiv, also bijektiv und damit invertierbar. $\square$

* * *

3 Determinanten, Eigenwerte und Jordan-Normalform

3.1 Determinanten

Sei $\mathrm{Per}(n)$ die Menge aller **Permutationen** in n Elementen, d.h., die Gruppe aller bijektiven Abbildungen

$$\sigma : \{1,\ldots,n\} \to \{1,\ldots,n\}.$$

Eine **Transposition** ist eine Permutation τ, die zwei Zahlen vertauscht und den Rest gleich lässt. Für $1 \le i < j \le n$ sei $\tau_{i,j}$ die Transposition, die i und j vertauscht. Es gilt $\tau_{i,j}^2 = \mathrm{Id}$.

Satz 3.1.1. *Jede Permutation $\sigma \in \mathrm{Per}(n)$ lässt sich als Produkt von Transpositionen schreiben:*

$$\sigma = \tau_1 \cdots \tau_k$$

Die Transpositionen $\tau_1,\ldots,\tau_k$ sind nicht eindeutig bestimmt, aber die Parität von k ist eindeutig bestimmt. Das bedeutet, dass für jede andere Darstellung

$$\sigma = \delta_1 \cdots \delta_m$$

mit Transpositionen δ_j gilt

$$(-1)^k = (-1)^m.$$

Wir nennen diese Zahl das **Signum** *von σ, also*

$$\mathrm{sign}(\sigma) = (-1)^k.$$

Das Signum ist eine Abbildung $\mathrm{sign} : \mathrm{Per}(n) \to \{\pm 1\}$ *mit*

$$\mathrm{sign}(\alpha\beta) = \mathrm{sign}(\alpha)\,\mathrm{sign}(\beta).$$

Beispiel 3.1.2. Sei $\sigma \in \mathrm{Per}(3)$ definiert durch

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

Dann ist $\mathrm{sign}(\sigma) = 1$, denn $\sigma = \tau_1\tau_2$ mit

$$\tau_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

Beweis des Satzes. Sei $\sigma \in \mathrm{Per}(n)$ mit $\sigma \neq \mathrm{Id}$. Sei k die kleinste Zahl mit $\sigma(k) \neq k$. Sei $j = \sigma(k)$ und betrachte die Permutation $\sigma_1 = \tau_{k,j}\sigma$. Dann folgt $\sigma_1(i) = i$ für alle $i \le k$. Wir wiederholen diesen Schritt bis wir am Ende

$$\mathrm{Id} = \tau_k \cdots \tau_1 \sigma$$

mit Transpositionen τ_j erhalten. Hieraus folgt $\sigma = \tau_1 \cdots \tau_k$ wie behauptet. Sei nun $\sigma = \delta_1 \cdots \delta_m$ eine zweite Darstellung. Für eine beliebige Permutation γ sei

$$F(\gamma) = \left\{ (i,j) : \begin{array}{c} 1 \le i < j \le n \\ \gamma(i) > \gamma(j) \end{array} \right\}$$

die Menge der **Fehlstände** von γ.

Wir beweisen nun: Ist τ eine Transposition, so gilt

$$|F(\sigma\tau)| = |F(\sigma)| + \text{ eine ungerade Zahl.} \tag{$*$}$$

Um dies zu zeigen sei $\tau = \tau_{i,j}$. Sind k, k' beide von i und j verschieden, so folgt

$$(k,k') \in F(\sigma) \quad \Leftrightarrow \quad (k,k') \in F(\sigma\tau).$$

Der Uebergang von σ zu $\sigma\tau$ ändert also an der Anzahl dieser Fehlstände nichts. Wir betrachten daher die Fälle, wenn $k' = i$ oder $k' = j$ ist. Ist $k < i < j$ so gilt

$$(k,i) \in F(\sigma) \quad \Leftrightarrow \quad (k,j) \in F(\sigma\tau)$$

und ebenso mit i und j vertauscht. Dasselbe passiert im Fall $i < j < k$. Die Anzahl sonder Fehlstände ändert sich also beim Uebergang von σ zu $\sigma\tau$ ebenfalls nicht.

Der interessante Fall ist $i < k < j$. Dann gilt

$$\begin{aligned} (i,k) \in F(\sigma) \quad &\Leftrightarrow \quad (k,j) \notin F(\sigma\tau), \\ (i,k) \notin F(\sigma) \quad &\Leftrightarrow \quad (k,j) \in F(\sigma\tau). \end{aligned}$$

Das heisst also, dass diese Stellen die Anzahl der Fehlstände, in denen k vorkommt, beim Uebergang zu $\sigma\tau$ entweder gar nicht, oder um ± 2 geandert wird.

Schliesslich gilt

$$(i,j) \in F(\sigma) \quad \Leftrightarrow \quad (i,j) \notin F(\sigma\tau).$$

An dieser Stelle wird also $|F(\sigma)|$ um ± 1 geändert und damit ist $(*)$ bewiesen.

Ist also $\sigma = \tau_1 \cdots \tau_k$, so folgt

$$(-1)^{F(\sigma)} = (-1)^{F(\tau_1 \cdots \tau_{k-1})+1} = \cdots = (-1)^k$$

und damit ist das Signum wohldefiniert. Zur Multiplikativität seien $\alpha = \tau_1 \cdots \tau_r$ und

$\beta = \eta_1 \cdots \eta_s$ mit Trasnpositionen τ_j und η_j. Dann gilt

$$\operatorname{sign}(\alpha\beta) = \operatorname{sign}(\tau_1 \cdots \tau_r \eta_1 \cdots \eta_s) = (-1)^{r+s} = (-1)^r (-1)^s = \operatorname{sign}(\alpha)\operatorname{sign}(\beta).$$

□

Definition 3.1.3. Sei $A \in \mathrm{M}_n(K)$. Definiere die **Determinante** von A durch

$$\det(A) = \sum_{\sigma \in \operatorname{Per}(n)} \operatorname{sign}(\sigma)\, A_{1,\sigma(1)} A_{2,\sigma(2)} \cdots A_{n,\sigma(n)}.$$

Beispiele 3.1.4. (a) $n = 1$, $\det(A) = A_{1,1}$.

(b) $n = 2$, $\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$.

(c) $n = 3$, $\det\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & j \end{pmatrix} = aej + bfg + cdh - ceg - bdj - afh$. Berechnung in diesem Fall:

Addiere die Produkte der gleichfarbigen Einträge

$$\left(\begin{array}{ccc|cc} a & b & c & a & b \\ d & e & f & d & e \\ g & h & j & g & h \end{array}\right)$$

dann subtrahiere die Produkte in der umgedrehten Diagonalen

$$\left(\begin{array}{ccc|cc} a & b & c & a & b \\ d & e & f & d & e \\ g & h & j & g & h \end{array}\right).$$

Das Ganze kompakter

$$+\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & j \end{pmatrix} \qquad\qquad -\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & j \end{pmatrix}$$

Proposition 3.1.5. *Für jede quadratische Matrix $A \in \mathrm{M}_n(K)$ gilt*

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

also

$$\det A = \sum_{\sigma \in \operatorname{Per}(n)} \operatorname{sign}(\sigma) A_{\sigma(1),1} \cdots A_{\sigma(n),n}.$$

Beweis. Es gilt $A^t_{i,j} = A_{i,j}$. Für $\sigma \in \mathrm{Per}(n)$ ordne das Produkt $A_{1,\sigma(1)} \cdots A_{n,\sigma(n)}$ nach dem zweiten Index und erhalte $A_{\sigma^{-1}(1),1} \cdots A_{\sigma^{-1}(n),n}$, wobei σ^{-1} die zu σ inverse Permutation ist. Ist $\sigma = \tau_1 \cdots \tau_k$ als Produkt von Transpositionen, so ist $\sigma^{-1} = \tau_k \cdots \tau_1$, also folgt $\mathrm{sign}(\sigma^{-1}) = \mathrm{sign}(\sigma)$. Damit ist

$$\begin{aligned}
\det A &= \sum_\sigma \mathrm{sign}(\sigma) A_{1,\sigma(1)} \cdots A_{n,\sigma(n)} \\
&= \sum_\sigma \mathrm{sign}(\sigma) A_{\sigma^{-1}(1),1} \cdots A_{\sigma^{-1}(n),n} \\
&= \sum_\sigma \mathrm{sign}(\sigma) A_{\sigma(1),1} \cdots A_{\sigma(n),n} = \det A^t.
\end{aligned}$$

□

Proposition 3.1.6. *Sei* $A = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix} \in \mathrm{M}_n(K)$ *mit* $X \in \mathrm{M}_k(K)$ *für ein* $k < n$. *Dann gilt*

$$\det(A) = \det \begin{pmatrix} X & Y \\ & Z \end{pmatrix} = \det(X)\ \det(Z).$$

Beweis. Sei $\sigma \in \mathrm{Per}(n)$ mit $\prod_{i=1}^n A_{i,\sigma(i)} = A_{1,\sigma(1)} \cdots A_{n,\sigma(n)} \neq 0$. Dann kann keine der Stellen $(i, \sigma(i))$ in der Nullmatrix links unten liegen. Es gilt also

$$i > k \Rightarrow \sigma(i) > k.$$

Das heisst, $\sigma(\{k+1, \dots, n\}) = \{k+1, \dots, n\}$ und da σ bijektiv ist, folgt auch $\sigma(\{1, \dots, k\}) = \{1, \dots, k\}$. Die Permutation σ ist also ein Produkt $\sigma = \eta\tau$ mit $\eta \in \mathrm{Per}(\{1, \dots, k\})$ und $\tau \in \mathrm{Per}(\{k+1, \dots, n\})$ und umgekehrt tritt jedes solche Produkt $\eta\tau$ auf. Also folgt

$$\begin{aligned}
\det \begin{pmatrix} X & Y \\ & W \end{pmatrix} &= \sum_\sigma \mathrm{sign}(\sigma) A_{1,\sigma(1)} \cdots A_{n,\sigma(n)} \\
&= \sum_\eta \sum_\tau \mathrm{sign}(\eta\tau) A_{1,\eta(1)} \cdots A_{k,\eta(k)} A_{k+1,\tau(k+1)} \cdots, A_{n,\tau(n)} \\
&= \sum_\eta \mathrm{sign}(\eta) A_{1,\eta(1)} \cdots A_{k,\eta(k)}\ \sum_\tau \mathrm{sign}(\tau) A_{k+1,\tau(k+1)} \cdots, A_{n,\tau(n)} \\
&= \sum_\eta \mathrm{sign}(\eta) X_{1,\eta(1)} \cdots X_{k,\eta(k)}\ \sum_\tau \mathrm{sign}(\tau) Z_{1,\tau(1)} \cdots, Z_{n,\tau(n-k)} \\
&= \det(X)\ \det(Z).
\end{aligned}$$

□

Korollar 3.1.7. *Es gilt*

$$\det \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \lambda_1 \cdots \lambda_n.$$

Proof. Wiederholte Anwendung der Proposition. □

Lemma 3.1.8. *Hat die Matrix A zwei gleiche Zeilen oder zwei gleiche Spalten, so ist*

$\det A = 0$.

Beweis. Wegen $\det(A) = \det(A^t)$ reicht es, anzunehmen, dass A zwei gleiche Zeilen hat. Es gebe also $i \neq j$ mit $a_{i,k} = a_{j,k}$ für jedes k. Dann gilt

$$\begin{aligned}\det A &= \sum_{\sigma} \operatorname{sign}(\sigma) \prod_{\nu=1}^{n} a_{\nu,\sigma(\nu)} \\ &= \sum_{\sigma:\sigma(i)<\sigma(j)} \operatorname{sign}(\sigma) \prod_{\nu=1}^{n} a_{\nu,\sigma(\nu)} + \sum_{\sigma:\sigma(i)>\sigma(j)} \operatorname{sign}(\sigma) \prod_{\nu=1}^{n} a_{\nu,\sigma(\nu)}.\end{aligned}$$

Sei $\tau = \tau_{i,j}$ die Transposition, die i und j vertauscht. In der zweiten Summe ersetzen wir σ durch $\sigma\tau$. Dann folgt

$$\det A = \sum_{\sigma:\sigma(i)<\sigma(j)} \operatorname{sign}(\sigma) \prod_{\nu=1}^{n} a_{\nu,\sigma(\nu)} + \sum_{\sigma:\sigma(i)<\sigma(j)} \operatorname{sign}(\sigma\tau) \prod_{\nu=1}^{n} a_{\nu,\sigma\tau(\nu)}$$

Für jedes k gilt $a_{\tau(\nu),k} = a_{\nu,k}$. Daher sehen wir, indem wir das Produkt nach $\tau(\nu)$ ordnen:

$$\prod_{\nu=1}^{n} a_{\nu,\sigma\tau(\nu)} = \prod_{\nu=1}^{n} a_{\tau(\nu),\sigma(\nu)} = \prod_{\nu=1}^{n} a_{\nu,\sigma(\nu)}.$$

Da $\operatorname{sign}(\sigma\tau) = -\operatorname{sign}(\sigma)$, ergibt sich $\det A = 0$. □

Beispiel 3.1.9. Es ist $\det\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = 0$.

Lemma 3.1.10. *Die Determinanten-Abbildung* $\det : \mathrm{M}_n(K) \to K$ *ist linear in jeder Zeile oder Spalte.*

Genauer: sind $a_1, \dots, a_n, a_j'$ *Spaltenvektoren, so gilt*

$$\det(a_1, \dots, a_j + a_j', \dots, a_n) = \det(a_1, \dots, a_j, \dots, a_n) + \det(a_1, \dots, a_j', \dots, a_n)$$

und für $\lambda \in K$ *gilt*

$$\det(a_1, \dots, \lambda a_j, \dots, a_n) = \lambda \det(a_1, \dots, a_j, \dots, a_n)$$

und analog für Zeilen.

Beweis. Es gilt

$$\det\begin{pmatrix} a_{1,1} & \dots & a_{1,j}+a'_{1,j} & \dots & a_{1,n} \\ \vdots & & \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,j}+a'_{n,j} & \dots & a_{n,n} \end{pmatrix} = \sum_\sigma \operatorname{sign}(\sigma) a_{\sigma(1),1} \cdots (a_{\sigma(j),j} + a'_{\sigma(j),j}) \cdots a_{\sigma(n),n}$$
$$= \sum_\sigma \operatorname{sign}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(j),j} \cdots a_{\sigma(n),n}$$
$$+ \sum_\sigma \operatorname{sign}(\sigma) a_{\sigma(1),1} \cdots a'_{\sigma(j),j} \cdots a_{\sigma(n),n}.$$

Die zweite Gleichung folgt ähnlich. Die Aussage über Zeilen folgt aus der für Spalten und $\det A = \det A^t$. □

* * *

3.2 Determinante und Zeilentransformationen

Satz 3.2.1. *Die Determinante hat folgende Eigenschaften.*

(a) *Falls die Matrix B aus A durch Addition des λ-fachen einer Zeile zu einer anderen hervorgeht (Zeilentrafo 1), so gilt*

$$\det B = \det A.$$

(b) *Falls B aus A durch Multiplikation einer Zeile mit $\mu \in K$ hervorgeht (Zeilentrafo 2), so gilt*

$$\det B = \mu \det A.$$

(c) *Falls die Matrix B aus A durch Vertauschung zweier Zeilen hervorgeht (Zeilentrafo 3), so gilt*

$$\det B = -\det A.$$

Beweis. (a) Seien $z_1, \dots, z_n$ Zeilenvektoren. Dann gilt

$$\det\begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_j + \lambda z_i \\ \vdots \\ z_n \end{pmatrix} = \det\begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_j \\ \vdots \\ z_n \end{pmatrix} + \lambda \underbrace{\det\begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_i \\ \vdots \\ z_n \end{pmatrix}}_{=0 \text{ (zwei gleiche Spalten)}} = \det\begin{pmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_j \\ \vdots \\ z_n \end{pmatrix}.$$

Teil (b) folgt ebenso. Für (c) rechne

$$\det\begin{pmatrix} z_1\\ \vdots\\ z_i\\ \vdots\\ z_j\\ \vdots\\ z_n\end{pmatrix} = \begin{pmatrix} z_1\\ \vdots\\ z_i+z_j\\ \vdots\\ z_j\\ \vdots\\ z_n\end{pmatrix} = \begin{pmatrix} z_1\\ \vdots\\ z_i+z_j\\ \vdots\\ z_j-(z_i+z_j)\\ \vdots\\ z_n\end{pmatrix} = \begin{pmatrix} z_1\\ \vdots\\ z_i+z_j\\ \vdots\\ -z_i\\ \vdots\\ z_n\end{pmatrix} = \begin{pmatrix} z_1\\ \vdots\\ z_j\\ \vdots\\ -z_i\\ \vdots\\ z_n\end{pmatrix} = -\begin{pmatrix} z_1\\ \vdots\\ z_j\\ \vdots\\ z_i\\ \vdots\\ z_n\end{pmatrix}.$$ □

Wir erinnern an die Elementarmatrizen:

$$A_{i,j}(\lambda) = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & \lambda & \ddots & \\ & & & 1\end{pmatrix}, \qquad M_i(\mu) = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \mu & & \\ & & & & 1 & \\ & & & & & \ddots\end{pmatrix},$$

$$S_{i,j} = \begin{pmatrix} 1 & & & & \\ & 0 & & 1 & \\ & & 1 & & \\ & 1 & & 0 & \\ & & & & 1\end{pmatrix},$$

Wobei wir jetzt auch $\mu = 0$ zulassen wollen. Dann ist allerdings $M_i(\mu)$ nicht mehr invertierbar.

Lemma 3.2.2. *Jede Matrix $A \in \mathrm{M}_n(K)$ ist ein Produkt von Elementarmatrizen.*

Beweis. Sei $A \in \mathrm{M}_n(K)$. Nach Proposition 2.9.7 gilt $A = S\begin{pmatrix} I & \\ & 0\end{pmatrix}T$ mit Produkten von Elementarmatrizen S, T. Der mittlere Faktor ist $M_{k+1}(0)\cdots M_n(0)$, wobei k die Länge der Eiheitsmatrix ist. □

Beispiel 3.2.3. Wir haben

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

Satz 3.2.4. *Für Matrizen $A, B \in \mathrm{M}_n(K)$ gilt*

$$\det(AB) = \det(A)\det(B).$$

Ferner gilt

$$A \text{ ist invertierbar} \quad \Leftrightarrow \quad \det A \neq 0.$$

Beweis. Für die Elementarmatrizen sieht man leicht, dass

$$\det A_{i,j}(\lambda) = 1, \quad \det M_i(\mu) = \mu, \quad \det S_{i,j} = -1.$$

Daher folgt aus Satz 3.2.1, dass $\det(SA) = \det S \det A$, falls S eine Elementarmatrix ist. Seien A, B beliebig. Nach Lemma 3.2.2 können wir schreiben:

$$A = S_1 \cdots S_k, \qquad B = T_1 \cdots T_m,$$

wobei $S_1, \dots, S_k, T_1, \dots T_m$ Elementarmatrizen sind. Es folgt

$$\begin{aligned} \det(AB) &= \det(S_1 \cdots S_k T_1 \cdots T_m) \\ &= \underbrace{\det(S_1) \cdots \det(S_k)}_{=\det(S_1 \cdots S_k)} \underbrace{\det(T_1) \cdots \det(T_m)}_{=\det(T_1 \cdots T_m)} \\ &= \det(A)\det(B). \end{aligned}$$

Sei nun A invertierbar. Dann ist $1 = \det(I) = \det(AA^{-1}) = \det(A)\det(A^{-1})$, also ist $\det(A) \neq 0$. Ist umgekehrt $\det A \neq 0$, so schreibe $A = S_1 \cdots S_k$ als Produkt von Elementarmatrizen. Dann folgt $0 \neq \det(A) = \det(S_1) \cdots \det(S_k)$, also ist $\det(S_j) \neq 0$ für jedes S_j. Daher ist keine der S_j gleich einer Matrix $M_i(0)$ und daher sind alle S_j invertierbar, also ist A invertierbar. □

Definition 3.2.5. Sei $A \in \mathrm{M}_n(K)$, $n \geq 2$ und für $1 \leq i, j \leq n$ sei $C(i, j) \in \mathrm{M}_{n-1}(K)$ die Matrix, die aus A durch Wegstreichen der i-ten Zeile und j-ten Spalte entsteht.

Beispiel 3.2.6. Ist $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$, dann ist

$$C(2,2) = \begin{pmatrix} 1 & \cancel{2} & 3 \\ \cancel{4} & \cancel{5} & \cancel{6} \\ 7 & \cancel{8} & 9 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 7 & 9 \end{pmatrix}$$

Es ist leicht zu sehen, dass

$$\det(C(i,j)) := \det \begin{pmatrix} a_{1,1} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,j-1} & a_{i-1,j+1} & \dots & a_{i-1,n} \\ a_{i+1,1} & \dots & a_{i+1,j-1} & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{n,n} \end{pmatrix}$$

$$= (-1)^{i+j} \det \begin{pmatrix} a_{1,1} & \dots & a_{1,j-1} & 0 & a_{1,j+1} & \dots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \dots & a_{i-1,n} \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ a_{i+1,1} & \dots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,j-1} & 0 & a_{n,j+1} & \dots & a_{n,n} \end{pmatrix}$$

gilt.

Satz 3.2.7 (Laplace Entwicklungssatz)**.** *Sei* $1 \le j_0 \le n$. *Dann gilt*

$$\det A = \sum_{i=1}^{n} (-1)^{i+j_0} a_{i,j_0} \det(C(i,j_0)).$$

Für $1 \le i_0 \le n$ *gilt*

$$\det A = \sum_{j=1}^{n} (-1)^{i_0+j} a_{i_0,j} \det(C(i_0,j)).$$

Beweis.

$$(-1)^{i+j_0}a_{i,j_0}\det(C(i,j_0)) = a_{i,j_0}\det\begin{pmatrix} a_{1,1} & \dots & a_{1,j_0-1} & 0 & a_{1,j_0+1} & \dots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,j_0-1} & 0 & a_{i-1,j_0+1} & \dots & a_{i-1,n} \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ a_{i+1,1} & \dots & a_{i+1,j_0-1} & 0 & a_{i+1,j_0+1} & \dots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,j_0-1} & 0 & a_{n,j_0+1} & \dots & a_{n,n} \end{pmatrix}$$

$$= \det\begin{pmatrix} a_{1,1} & \dots & a_{1,j_0-1} & 0 & a_{1,j_0+1} & \dots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,j_0-1} & 0 & a_{i-1,j_0+1} & \dots & a_{i-1,n} \\ 0 & \dots & 0 & a_{i,j_0} & 0 & \dots & 0 \\ a_{i+1,1} & \dots & a_{i+1,j_0-1} & 0 & a_{i+1,j_0+1} & \dots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,j_0-1} & 0 & a_{n,j_0+1} & \dots & a_{n,n} \end{pmatrix}$$

$$= \sum_{\substack{\sigma\in\mathrm{Per}(n)\\ \sigma(i)=j_0}} \mathrm{sign}(\sigma)a_{1,\sigma(1)}\cdots a_{n,\sigma(n)}.$$

Also ist

$$\sum_{i=1}^{n}(-1)^{i+j_0}a_{i,j_0}\det(C(i,j_0)) = \sum_{i=1}^{n}\sum_{\substack{\sigma\in\mathrm{Per}(n)\\ \sigma(i)=j_0}} \mathrm{sign}(\sigma)a_{1,\sigma(1)}\cdots a_{n,\sigma(n)} = \det A.$$

Die andere Aussage beweist man analog. □

Beispiel 3.2.8.

$$\det\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 0 & 3 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 5 & 6 & 0 & 1 & 2 \\ 7 & 8 & 0 & 3 & 2 \end{pmatrix} = -54.$$

* * *

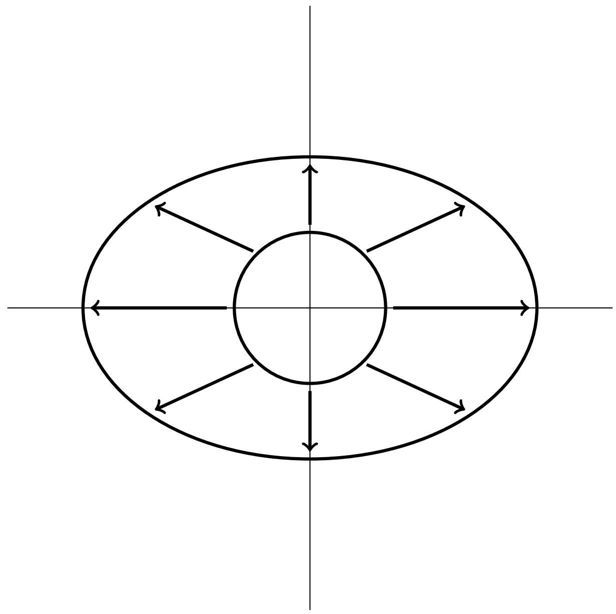
3.3 Eigenwerte

Definition 3.3.1. Sei $T : V \to V$ linear. Ein Vektor $v \in V$ heißt **Eigenvektor** von T zum **Eigenwert** $\lambda \in K$, falls

- $v \neq 0$ und

- $T(v) = \lambda v$.

Beispiel 3.3.2. Betrachte die Matrix $A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \in M_n(\mathbb{R})$. Die Abbildung $x \mapsto Ax$ von $\mathbb{R}^2$ nach $\mathbb{R}^2$ streckt den Vektor $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ um den Faktor 2 und den Vektor $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ um den Faktor 3. Sie bildet also den Kreis mit Radius 1 auf eine Ellipse ab:



Wir betrachten nun die Basis $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ und entsprechend die Basiswechselmatrix $S = \begin{pmatrix} 1 & 1 \\ 2 & \end{pmatrix}$, so gilt $Se_1 = v_1$ und $Se_2 = v_2$. Wir berechnen die Inverse zu $S^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$. Aus $Se_j = v_j$ folgt dann $S^{-1}v_j = e_j$. Sei dann $B = SAS^{-1}$, so gilt

$$Bv_1 = SAS^{-1}v_1 = SAe_1 = S(2e_1) = 2v_1$$

und ebenso $Bv_2 = 3v_2$. Die Matrix $B = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$ beschreibt also dieselbe Abbildung nur in der Basis v_1, v_2.

Beispiel 3.3.3. Sei $K = \mathbb{Q}$ und $T = \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}$. Dann ist der Vektor $v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ein Eigenvektor zum Eigenwert 2. Ferner ist

$$T\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

also ist $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ein Eigenvektor zu Eigenwert 3.

Satz 3.3.4. *Sei $T : V \to V$ linear und seien $v_1, \dots, v_k$ Eigenvektoren zu paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_k$. Dann sind die $v_1, \dots, v_k$ unabhängig.*

Beweis. Induktion nach k. Für $k = 1$ ist nichts zu zeigen, da ein Eigenvektor stets $\neq 0$ ist.

$k \to k+1$: Seien

$$\mu_1 v_1 + \cdots + \mu_{k+1} v_{k+1} = 0 \tag{I}$$

eine Linearkombination der Null. Wir wenden T hierauf an und erhalten

$$\mu_1 \lambda_1 v_1 + \cdots + \mu_{k+1} \lambda_{k+1} v_{k+1} = 0. \tag{II}$$

Wir multiplizieren (I) mit λ_{k+1} und erhalten

$$\mu_1 \lambda_{k+1} v_1 + \cdots + \mu_{k+1} \lambda_{k+1} v_{k+1} = 0. \tag{III}$$

Nun ziehen wir (III) von (II) ab und erhalten

$$\mu_1(\lambda_1 - \lambda_{k+1}) v_1 + \cdots + \mu_k(\lambda_k - \lambda_{k+1}) v_k = 0. \tag{II)-(III}$$

Nach Induktionsvoraussetzung ist $\mu_j(\lambda_j - \lambda_{k+1}) = 0$ für jedes $j = 1, \dots, k$. Da $\lambda_j \neq \lambda_{k+1}$, folgt $\mu_1 = \cdots = \mu_k = 0$. Damit folgt aus (I), dass $\mu_{k+1} v_{k+1} = 0$ und wegen $v_{k+1} \neq 0$, ist auch $\mu_{k+1} = 0$. □

Definition 3.3.5. Eine Matrix, die zu einer Diagonalmatrix $\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ konjugiert ist, heißt **diagonalisierbar**.

Proposition 3.3.6. *Hat die Matrix $A \in \mathrm{M}_n(K)$ paarweise verschiedene Eigenwerte $\lambda_1, \dots, \lambda_n$, also n-Stück, dann ist A diagonalisierbar.*

Beweis. Seien $v_1, \dots, v_n$ Eigenvektoren zu den Eigenwerten. Dann sind $v_1, \dots, v_n$ unabhängig nach Satz 3.3.4. Daher ist die Matrix S mit den Spalten $v_1, \dots, v_n$ invertierbar und es gilt

$$S^{-1} A S e_k = S^{-1} A v_k = \lambda_k S^{-1} v_k = \lambda_k e_k,$$

mit anderen Worten:

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

□

Beispiel 3.3.7. Sei $A = \begin{pmatrix} -7 & 2 \\ -24 & 7 \end{pmatrix}$. Finde S mit $S^{-1}AS = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$.

* * *

3.4 Polynome

Sei K ein Körper. Der **Polynomring** $K[x]$ ist die Menge der formalen Ausdrücke der Form

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

mit $n \in \mathbb{N}$ und $a_0, \ldots, a_n \in K$. Man hat eine Addition:

$$(a_0 + a_1x + a_2x^2 + \cdots + a_nx^n) + (b_0 + b_1x + b_2x^2 + \cdots + b_nx^n)$$

$$= (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \cdots + (a_n + b_n)x^n,$$

wobei man bei verschiedenem Grad ein Polynom durch Nullen auffüllt. Ferner hat man eine Multiplikation

$$\begin{aligned} &(a_0 + a_1x + a_2x^2 + \cdots + a_mx^m)(b_0 + b_1x + b_2x^2 + \cdots + b_nx^n) \\ &= a_0b_0 + (a_1b_0 + a_0b_1)x + \cdots + \Big(\sum_{k+l=p} a_kb_l \Big)x^p + \ldots \end{aligned}$$

Präziser definieren wir den Polynomring als Menge aller Folgen $(a_0, a_1, a_2, \ldots)$ in K, die nur endlich viele Glieder $\neq 0$ haben. Die Addition ist dann

$$(a_0, a_1, a_2, \ldots) + (b_0, b_1, b_2, \ldots) = (a_0 + b_0, a_1 + b_1, \ldots)$$

und die Multiplikation ist

$$(a_0, a_1, a_2, \ldots)(b_0, b_1, b_2, \ldots) = (c_0, c_1, c_2, \ldots)$$

mit $c_p = \sum_{i+j=p} a_jb_i$. Praktischer ist es allerdings, Polynome mit einer Unbekannten x zu schreiben.

Definition 3.4.1. Sei $p(x) = a_0 + a_1x + \cdots + a_nx^n$ ein Polynom über dem Körper K. Dann definiert p eine Abbildung

$$\begin{aligned} \tilde{p} : K &\to K, \\ \lambda &\mapsto a_0 + a_1\lambda + \cdots + a_n\lambda^n, \end{aligned}$$

wobei jetzt die Potenzen, Produkte und Summen in K genommen werden. Man nennte eine solche Abbildung $K \to K$ eine **polynomiale Abbildung**.

Beispiel 3.4.2. Sei $K = \mathbb{F}_2$, dann gilt für jedes $t \in K$, dass

$$1 + t = 1 + t^2,$$

denn es gibt ja nur die Elemente 0 und 1. Aber die Polynome $1 + x$ und $1 + x^2$ sind verschiedene Elemente in $\mathbb{F}_2[x]$. Man kann also $K[x]$ **nicht** mit der Menge der polynomialen Abbildungen $K \to K$ identifizieren! Im Allgemeinen gilt also

Polynome $\neq$ polynomiale Abbildungen!

Wir werden allerdings feststellen, dass diese Identifikation doch klappt, falls der Körper K unendlich viele Elemente hat.

Definition 3.4.3. Für ein Polynom $f(x) = a_0 + a_1 x + \cdots + a_n x^n$, das nicht Null ist, sei der **Grad** von f das größte $k \in \mathbb{N}_0$ mit $a_k \neq 0$. Man erweitert diese Definition durch

$$\operatorname{grad}(0) = -\infty.$$

Dann gilt für beliebige Polynome $f(x), g(x)$, dass

$$\begin{aligned} \operatorname{grad}(fg) &= \operatorname{grad}(f) + \operatorname{grad}(g), \\ \operatorname{grad}(f + g) &\leq \big(\operatorname{grad}(f), \operatorname{grad}(g)\big), \end{aligned}$$

wenn man formal $-\infty + n = n + (-\infty) = -\infty$ rechnet.

Satz 3.4.4 (Division mit Rest)**.** *Sind $f, g \in K[x]$ und ist $g \neq 0$, so gibt es eindeutig bestimmte Polynome $q, r \in K[x]$ so dass*

$$f = qg + r \qquad \textit{und} \qquad \operatorname{grad}(r) < \operatorname{grad}(g).$$

Beweis. Existenz: Wir geben ein Verfahren zur Berechnung an. Schreibe

$$\begin{aligned} f(x) &= a_n x^n + \cdots + a_0 \\ g(x) &= b_m x^m + \cdots + b_0 \end{aligned}$$

mit $a_n \neq 0$ und $b_m \neq 0$. Ist $n < m$, so setze $r = f$ und $q = 0$, denn dann ist ja

$$f = qg + r \qquad \text{und} \qquad \operatorname{grad}(r) < \operatorname{grad}(g).$$

Ist hingegen $n \geq m$, so setze $q_1 = \frac{a_n}{b_m} x^{n-m}$ und $f_1 = f - q_1 g$. Dann folgt $\text{grad}(f_1) < \text{grad}(f)$. Ersetze nun f durch f_1 und wiederhole diesen Vorgang bis der Grad kleiner wird als $\text{grad}(g)$.

Nun zur *Eindeutigkeit:* Seien $q', r' \in K[x]$ zwei weitere Polynome mit $f = q'g + r'$ und $\text{grad}(r) < \text{grad}(g)$. Dann gilt

$$0 = f - f = (q - q')g + (r - r').$$

Also gilt $(q - q')g = (r' - r)$ und damit

$$\begin{aligned} \text{grad}(q - q') + \text{grad}(g) &= \text{grad}(r' - r) \\ &\leq \max(\text{grad}(r'), \text{grad}(r)) < \text{grad}(g). \end{aligned}$$

Daraus folgt $\text{grad}(q - q') < 0$, also $\text{grad}(q - q') = -\infty$ und damit $q = q'$, woraus auch $r = r'$ folgt. □

Beispiel 3.4.5. Sei $K = \mathbb{Q}$ und $f(x) = 3x^3 + 2x + 1$, sowie $g(x) = x^2 + 4x$. Dann ist $q(x) = 3x - 12$ und $r(x) = 50x + 1$.

Sei $f \in K[x]$, etwa

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n.$$

Ein $\lambda \in K$ mit $f(\lambda) = 0$ heißt **Nullstelle** des Polynoms f.

Proposition 3.4.6. *Ist $\lambda \in K$ eine Nullstelle von $f \in K[x]$, $f \neq 0$, dann existiert ein $g(x) \in K[x]$ und ein $p \in \mathbb{N}$ mit*

$$f(x) = (x - \lambda)^p g(x)$$

und $g(\lambda) \neq 0$. Es folgt dann $\text{grad}(g) = \text{grad}(f) - p$. *Die Zahl p heißt die* **Vielfachheit** *der Nullstelle λ von f.*

Beweis. Nach Division mit Rest existieren $g_1(x)$ und $r(x)$ mit

$$f(x) = (x - \lambda)g_1(x) + r(x)$$

und $\text{grad}(r) < \text{grad}(x - \lambda) = 1$. Daher ist $r(x)$ konstant. Aber wegen

$$0 = f(\lambda) = (\lambda - \lambda)g(\lambda) + r = r$$

ist $r = 0$ und wir erhalten $f(x) = (x - \lambda)g_1(x)$. Ist nun $g_1(\lambda) \neq 0$, setzen wir $p = 1$ und $g = g_1$. Andernfalls wiederholen wir den Schritt mit g_1 an der Stelle von f, so lange bis wir auf ein Polynom g mit $g(\lambda) \neq 0$ stoßen. □

Korollar 3.4.7. *Ist $f \in K[x]$, $f \neq 0$ und ist k die Anzahl der Nullstellen von f. Dann gilt*

$$k \leq \mathrm{grad}(f).$$

Beweis. Sind $\lambda_1, \dots, \lambda_k$ die verschiedenen Nullstellen von f, dann gilt

$$f(x) = (x - \lambda_1)g_1(x)$$

für ein $g_1 \in K[x]$, $g_1 \neq 0$. Dann ist

$$0 = f(\lambda_2) = \underbrace{(\lambda_2 - \lambda_1)}_{\neq 0} g_1(\lambda_2)$$

und also $g_1(\lambda_2) = 0$. Daher gibt es ein $g_2(x) \neq 0$ mit $g_1(x) = (x - \lambda_2)g_2(x)$ und daher

$$f(x) = (x - \lambda_1)(x - \lambda_2)g_2(x).$$

Wir wiederholen dies Argument bis zu

$$f(x) = (x - \lambda_1)(x - \lambda_2)\cdots(x - \lambda_k)g_k(x),$$

mit $g_k \neq 0$, was soviel heisst wie $\mathrm{grad}(g_k) \geq 0$. Daher ist $\mathrm{grad}(f) = k + \mathrm{grad}(g_k) \geq k$. □

Definition 3.4.8. Sei $f \in K[x]$, $f \neq 0$ und $\lambda \in K$. Dann heißt

$$\mu(f, \lambda) = \max\left\{n \in \mathbb{N} : \exists_{g \in K[x]} f(x) = (x - \lambda)^n g(x)\right\}$$

die **Vielfachheit** der Nullstelle λ. Es ist $\mu(f, \lambda) = 0$ falls λ gar keine Nullstelle ist.

Sind $\lambda_1, \dots, \lambda_k$ die verschiedenen Nullstellen von f und $m_j = \mu(f, \lambda_j)$ die Vielfachheiten, dann gilt

$$f(x) = (a - \lambda_1)^{m_1} \cdots (x - \lambda_k)^{m_k} g(x)$$

für ein Polynom g, das keine Nullstelle hat.

Beispiel 3.4.9. Das Polynom $g(x) = 1 + x^2$ hat in $K = \mathbb{R}$ keine Nullstelle.

Satz 3.4.10. *Hat der Körper K unendlich viele Elemente, dann ist jedes Polynom p durch seine polynomiale Abbildung $\tilde{p} : K \to K$ eindeutig festgelegt. In diesem Fall gilt also*

Polynome = polynomiale Abbildungen!

Beweis. Sei K unendlich und es gelte $\tilde{f} = \tilde{g}$ für zwei Polynome f, g. Wir müssen zeigen, dass $f = g$ gilt.

Für $\alpha \in K$ gilt $(f-g)^{\sim}(\alpha) = f(\alpha) - g(\alpha) = \tilde{f}(\alpha) - \tilde{g}(\alpha) = 0$. Das heißt, das Polynom $h = f - g$ hat unendlich viele Nullstellen. Nach Korollar 3.4.7 muss es das Nullpolynom sein, also $f - g = 0$ oder $f = g$. □

Definition 3.4.11. Ein Körper K heisst **algebraisch abgeschlossen**, falls jedes nichtkonstante Polynom $f(x) \in K[x]$ eine Nullstelle hat.

Beispiele 3.4.12. (a) $K = \mathbb{Q}$ ist nicht algebraisch abgeschlossen, weil $f(x) = x^2 - 2$ keine Nullstelle hat.

(b) $K = \mathbb{R}$ ist nicht algebraisch abgeschlossen, weil $x^2 + 1$ keine Nullstelle hat.

(c) Der **Fundamentalsatz der Algebra** besagt, dass der Körper $\mathbb{C}$ algebraisch abgeschlossen ist.

FAKT: Zu jedem Koerper K existiert ein Oberkoerper $\widehat{K} \supset K$, der algebraisch abgeschlossen ist. (Beweis in der Algebra-Vorlesung.)

* * *

3.5 Das charakteristische Polynom

Definition 3.5.1. Sei $A \in \mathrm{M}_n(K)$ mit Einträgen $(a_{i,j})$. Dann ist die Determinante

$$\det(A) = \sum_{\sigma} \mathrm{sign}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$$

ein "Polynom" in den Einträgen $a_{i,j}$. Daraus folgt: Ist $F(x)$ eine Matrix

$$F(x) = \begin{pmatrix} F_{1,1}(x) & \dots & F_{1,n}(x) \\ \vdots & & \vdots \\ F_{n,1}(x) & \dots & F_{n,n}(x) \end{pmatrix},$$

wobei die Einträge Polynome $F_{i,j}(x) \in K[x]$ sind, dann ist $\det(F(x))$ ein Polynom in $K[x]$. Wir wenden dies an auf

$$F(x) = xI - A = \begin{pmatrix} x - a_{1,1} & \dots & -a_{1,n} \\ \vdots & \ddots & \vdots \\ -a_{n,1} & \dots & x - a_{n,n} \end{pmatrix}.$$

Das Polynom

$$\chi_A(x) = \det(xI - A)$$

nennt man das **charakteristische Polynom** von A.

Beispiel 3.5.2. Sei $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, dann ist

$$\chi_A(x) = \det\begin{pmatrix} x-1 & -2 \\ -3 & x-4 \end{pmatrix} = (x-1)(x-4) - 6 = x^2 - 5x - 2.$$

Satz 3.5.3. *Die Eigenwerte einer Matrix $A \in \mathrm{M}_n(K)$ sind genau die Nullstellen des charakteristischen Polynoms $\chi_A(x)$.*

Beweis. Sei λ ein Eigenwert von A. Dann existiert ein Vektor $v \neq 0$ in K^n mit $Av = \lambda v$, also $(\lambda I - A)v = 0$. Daher ist $\ker(\lambda I - A) \neq 0$, also ist die Matrix $(\lambda I - A)$ nicht invertierbar, also folgt

$$\chi_A(\lambda) = \det(\lambda I - A) = 0$$

nach Satz 3.2.4. Ist umgekehrt λ eine Nullstelle von $\chi_A(x)$, dann ist die Matrix $(\lambda I - A)$ nicht invertierbar, hat also einen nichttrivialen Kern, damit gibt es einen Vektor $v \in K^n$ mit $v \neq 0$ und $Av = \lambda v$. □

Beispiel 3.5.4. Sei $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Dann ist

$$\chi_A(x) = \det\begin{pmatrix} x-1 & -1 \\ -1 & x-1 \end{pmatrix} = (x-1)^2 - 1 = x^2 - 2x = x(x-2).$$

Dieses Polynom hat genau die Nullstellen $x = 0$ und $x = 2$. Hierzu gehören die Eigenvektoren $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Korollar 3.5.5. *Jede Matrix in $\mathrm{M}_n(\mathbb{C})$ hat einen Eigenwert.*

Korollar 3.5.6. *Ist K algebraisch abgeschlossen und ist $f(x) \in \mathrm{M}_n(K)$ ein nichtkonstantes Polynom, dann zerfällt f in Linearfaktoren, d.h.,*

$$f(x) = c(x - \lambda_1)\cdots(x - \lambda_n),$$

wobei $c, \lambda_1, \dots, \lambda_n \in K$ sind.

Beweis. Ist λ eine Nullstelle von f, dann ist $f(x) = (x - \lambda)g(x)$ für ein Polynom g. Wir wiederholen dies für g, bis wir bei einer Konstanten c ankommen. □

Definition 3.5.7. Zwei Matrizen $A, B \in \mathrm{M}_n(K)$ heißen **konjugiert**, falls es eine invertierbare Matrix $S \in \mathrm{GL}_n(K)$ gibt mit

$$B = S^{-1}AS.$$

Satz 3.5.8. *Sind die Matrizen $A, B \in \mathrm{M}_n(K)$ konjugiert, dann haben sie dasselbe charakteristische Polynom.*

Beweis. Sei $B = S^{-1}AS$, dann gilt

$$\begin{aligned}\chi_B(x) &= \det(xI - B) = \det(xS^{-1}S - S^{-1}AS)\\ &= \det(S^{-1}(xI - A)S)\\ &= \det(S^{-1})\det(xI - A)\det(S) = \det(S^{-1}S)\det(xI - A)\\ &= \det(xI - A) = \chi_A(x).\end{aligned}$$

□

Beispiel 3.5.9. Sei $A = \begin{pmatrix} 2 & \\ & 1 \end{pmatrix}$ und $S = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$. Man berechnet die inverse Matrix als $S^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$. Dann ist

$$B = S^{-1}AS = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 2 & \\ & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ -2 & 0 \end{pmatrix}.$$

Wir haben $\chi_A(x) = (x-2)(x-1)$ und

$$\chi_B(x)\det\begin{pmatrix} x-3 & -1 \\ 2 & x \end{pmatrix} = (x-3)x + 2 = x^2 - 3x + 2 = (x-2)(x-1).$$

Definition 3.5.10. Sei $T : V \to V$ linear, wobei V endlich-dimensional ist. Wähle dann eine Basis α von V und betrachte die Matrix $A = \mathrm{M}_\alpha(T)$. Definiere dann das **charakteristische Polynom** von T als

$$\chi_T(x) = \chi_A(T) = \det(xI - A).$$

Nach Satz 3.5.8 und dem Basiswechselsatz 2.8.5 hängt χ_T nicht von der Wahl der Basis ab.

Definition 3.5.11. Eine Matrix $A \in \mathrm{M}_n(K)$ heißt **trigonalisierbar**, wenn sie zu einer oberen Dreiecksmatrix konjugiert ist.

Satz 3.5.12. *Eine quadratische Matrix A ist genau dann trigonalisierbar, wenn ihr charakteristisches Polynom in ein Produkt von Linearfaktoren zerfällt. In diesem Fall ist A konjugiert zu einer Matrix der Form*

$$D = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

wobei $\lambda_1, \dots, \lambda_n$ die (nicht notwendig verschiedenen) Nullstellen von χ_A sind.
Ist K algebraisch abgeschlossen, dann ist jede Matrix trigonalisierbar.

Beweis. Sei A trigonalisierbar, also konjugiert zu einer Matrix

$$D = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Dann ist $\chi_A(x) = \chi_D(x) = \det \begin{pmatrix} x-\lambda_1 & & * \\ & \ddots & \\ & & x-\lambda_n \end{pmatrix} = (x-\lambda_1)\cdots(x-\lambda_n)$.

Umgekehrt sei $\chi_A(x) = (x-\lambda_1)\cdots(x-\lambda_n)$. Dann hat A einen Eigenvektor v_1 zum Eigenwert λ_1. Ergänze v_1 zu einer Basis $v_1, \dots, v_n$ von K^n. Sei dann S die Matrix mit den Spalten $v_1, \dots, v_n$. Dann ist der Rang von S gleich n, also ist S invertierbar. Setze $A_1 = S^{-1}AS$. Dann folgt $A_1 e_1 = S^{-1}ASe_1 = S^{-1}Av_1 = \lambda_1 S^{-1}v_1 = \lambda_1 e_1$. Also ist A_1 von der Form $\begin{pmatrix} \lambda_1 & * \\ 0 & C \end{pmatrix}$ für eine Matrix C mit $\chi_C(x) = (x-\lambda_2)\cdots(x-\lambda_n)$. Eine Induktion nach n liefert die Existenz einer invertierbaren $(n-1)\times(n-1)$ Matrix T mit
$T^{-1}CT = \begin{pmatrix} \lambda_2 & & * \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$. Mit der Matrix $U = \begin{pmatrix} 1 & \\ & T \end{pmatrix}$ folgt dann
$U^{-1}A_1U = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$. □

Satz 3.5.13. *Für $A, B \in \mathrm{M}_n(K)$ gilt*

$$\chi_{AB} = \chi_{BA}.$$

Beweis. Ist A invertierbar, so gilt $\chi_{AB} = \chi_{A^{-1}(AB)A} = \chi_{BA}$.

Betrachte nun den Fall, dass A nicht invertierbar ist. Wir können annehmen, dass K unendlich viele Elemente hat, da wir sonst zu einem Erweiterungskörper übergehen.

Dann ist die Matrix $A - \lambda I$ für unendlich viele $\lambda \in K$ invertierbar, da das Polynom $\lambda \mapsto \det(A - \lambda I)$ nur endlich viele Nullstellen hat. Für jedes solche λ gilt

$$\chi_{(A-\lambda I)B}(x) = \chi_{B(A-\lambda I)}(x)$$

für alle $x \in K$. Fixiere ein $x \in K$. Dann hat die Polynomabbildung

$$\lambda \mapsto \chi_{(A-\lambda I)B}(x) - \chi_{B(A-\lambda I)}(x)$$

unendlich viele Nullstellen, ist also identisch Null. Man kann also $\lambda = 0$ einsetzen und findet

$$\chi_{AB}(x) - \chi_{BA}(x) = 0$$

für jedes $x \in K$. Da K unendlich viele Elemente hat, ist das Polynom auf der linken Seite Null, also $\chi_{AB} = \chi_{BA}$. □

Definition 3.5.14. Sei $A = (a_{i,j})$. Dann gilt

$$\begin{aligned}\chi_A(x) &= (x - a_{1,1}) \dots (x - a_{n,n}) + \text{ Terme vom Grad } \leq n-2\\ &= x^n - (a_{1,1} + \cdots + a_{n,n})x^{n-1} + \text{ Terme vom Grad } \leq n-2.\end{aligned}$$

Wir definieren daher

$$\operatorname{tr}(A) = a_{1,1} + \cdots + a_{n,n}$$

und nennen diesen Ausdruck die **Spur** der Matrix A.

Satz 3.5.15. *Die Abbildung* $\operatorname{tr} : \mathrm{M}_n(K)$ *ist linear und erfüllt*

$$\operatorname{tr}(AB) = \operatorname{tr}(BA)$$

für $A, B \in \mathrm{M}_n)K)$ *und damit insbesondere auch* $\operatorname{tr}(SAS^{-1}) = \operatorname{tr}(A)$, *falls* $S \in \mathrm{GL}_n(L)$.

Beweis. Dies folgt aus der Tatsache, dass $-\operatorname{tr}(A)$ der $(n-1)$-te Koeffizient von $\chi_A(x)$ ist. □

Definition 3.5.16. Ist $T : V \to V$ eine lineare Abbildung auf dem endlich-dimensionalen K-Vektorraum V, dann definieren wir das **charakteristische Polynom** von T durch $\chi_T(x) = \chi_A(x)$, wobei A eine Matrix ist, die T in einer Basis von V darstellt. Der letzte Satz stellt sicher, dass $\chi_T(x)$ nicht von der Wahl der Basis abhängt.

Definition 3.5.17. Sei $\lambda \in K$ und $A \in \mathrm{M}_n(K)$. Wir definieren den **Eigenraum** zu λ durch

$$\mathrm{Eig}_A(\lambda) = \left\{v \in K^n : Av = \lambda v\right\}.$$

Es gilt dann

$$\lambda \text{ ist Eigenwert} \quad \Leftrightarrow \quad \mathrm{Eig}_A(\lambda) \neq 0.$$

Ist λ eine Eigenwert von A, so nennen ist die **algebraische Vielfachheit** $m(A,\lambda)$ von λ gleich der Vielfachheit der Nullstelle λ von χ_A.

Wir definieren die **geometrische Vielfachheit** von λ als die Dimension des Eigenraums $\mathrm{Eig}_A(\lambda)$.

Proposition 3.5.18. *Sind $A, S \in \mathrm{M}_n(K)$ und ist S invertierbar, so gilt für jedes $\lambda \in K$*

$$S\Big(\mathrm{Eig}_A(\lambda)\Big) = \mathrm{Eig}_{SAS^{-1}}(\lambda).$$

Proof. Es gilt

$$\begin{aligned} v \in \mathrm{Eig}_A(\lambda) &\Leftrightarrow Av = \lambda v \\ &\Leftrightarrow SAv = \lambda Sv \\ &\Leftrightarrow SAS^{-1}Sv = \lambda Sv \\ &\Leftrightarrow Sv \in \mathrm{Eig}_{SAS^{-1}}(\lambda). \end{aligned}$$

□

Beispiele 3.5.19. (a) Sei $A = \begin{pmatrix} 3 & \\ & 3 \end{pmatrix}$. Dann ist $\chi_A(x) = (x-3)^2$, also ist die algebraische Vielfachheit des Eigenwertes 3 gleich 2. In diesem Fall ist dies auch gleich der geometrischen Vielfachheit, denn

$$\mathrm{Eig}(A,3) = K^2.$$

(b) Sei $B = \begin{pmatrix} 3 & 1 \\ & 3 \end{pmatrix}$. Dann ist $\chi_B(x) = (x-3)^2$, also ist auch hier die algebraische Vielfachheit gleich 2, aber die geometrische ist gleich 1, denn

$$\mathrm{Eig}(B,3) = Ke_1.$$

Korollar 3.5.20. *Zerfällt $\chi_A(x)$ in Linearfaktoren, also etwa*

$$\chi_A(x) = \prod_{j=1}^{k}(x-\lambda_j)^{m_j},$$

wobei $\lambda_1, \dots, \lambda_k$ die verschiedenen Eigenwerte mit algebraischen Vielfachheiten m_i sind, dann ist

$$\mathrm{tr}(A) = \sum_{j=1}^{k} m_j \lambda_j.$$

Beweis. Nach Satz 3.5.12 ist A trigonalisierbar und nach Satz 3.5.15 können wir also annehmen, dass A eine obere Dreiecksmatrix ist. Die Diagonaleinträge sind dann $\lambda_1, \ldots, \lambda_k$, wobei λ_j gemäß der Vielfachheit m_j wiederholt wird. □

Proposition 3.5.21. *Die geometrische Vielfachheit eines Eigenwertes ist stets kleiner oder gleich der algebraischen, also*

$$\dim \operatorname{Eig}(A, \lambda) \leq m(A, \lambda).$$

Beweis. Sei $v_1, \ldots, v_k$ eine Basis von $\operatorname{Eig}(A, \lambda)$. Erweitere diese zu einer Basis $v_1, \ldots, v_n$ von K^n. Sei S die Matrix mit den Spaltenvektoren $v_1, \ldots, v_n$. Für
$1 \leq j \leq k = \dim \operatorname{Eig}(A, \lambda)$ gilt

$$S^{-1}ASe_j = S^{-1}Av_j = \lambda_j S^{-1}v_j = \lambda_j e_j.$$

Daher ist

$$S^{-1}AS = \begin{pmatrix} \lambda I_k & * \\ 0 & C \end{pmatrix}$$

und es folgt $\chi_A(x) = (x - \lambda)^k \chi_C(x)$. □

Satz 3.5.22. *Für eine Matrix $A \in \mathrm{M}_n(K)$ sind äquivalent:*

(a) *A ist diagonalisierbar,*

(b) *$\chi_A(x)$ zerfällt in Linearfaktoren und $\dim \operatorname{Eig}(T, \lambda) = m(T, \lambda)$ für jedes $\lambda \in K$,*

(c) *$K^n = \bigoplus_{\lambda \in K} \operatorname{Eig}(T, \lambda)$. Die Summe ist in Wirklichkeit endlich, da nur endlich viele λ einen nichtverschwindenden Beitrag leisten.*

(d) *Es gibt eine Basis von K^n, die aus Eigenvektoren besteht.*

Beweis. (a)⇒(b) Ist A diagonalisierbar, so gibt es eine Basis α bzgl der T durch eine Diagonalmatrix mit Diagonale $\lambda_1, \ldots, \lambda_1, \lambda_2, \ldots, \lambda_k$, wobei jeder Eigenwert λ_j gemäß seiner algebraischen Vielfachheit m_j wiederholt wird. Dann ist $\chi_T(x) = \prod_{j=1}^k (x - \lambda_j)^{m_j}$ und $m_j = \dim \operatorname{Eig}(T, \lambda_j)$.

(b)⇒(c) Die Summe $\operatorname{Eig}(T, \lambda_1) \oplus \cdots \oplus \operatorname{Eig}(T, \lambda_k)$ ist direkt nach Satz 3.3.4. Ferner ist $\dim V = \sum_j \dim \operatorname{Eig}(T, \lambda_j)$ und daher folgt die Behauptung.

(c)⇒(d): Man wählt Basen der Eigenräume und setzt sie zu einer Basis des ganzen Raums zusammen.

(d)⇒(a) ist klar. □

Definition 3.5.23. Sei $T : V \to V$ linear mit $n = \dim V < \infty$. Wir sagen: T ist **diagonalisierbar**, falls die Matrix von T in einer (und damit jeder) Basis diagonalisierbar ist. Für $\lambda \in K$ definieren wir den **Eigenraum** als

$$\mathrm{Eig}_T(\lambda) = \big\{v \in V : Tv = \lambda v\big\}.$$

Proposition 3.5.24. *Für eine lineare Abbildung $T : V \to V$ mit endlich-dimensionalem V sind äquivalent:*

(a) *T ist diagonalisierbar,*

(b) $V = \bigoplus_{\lambda \in K} \mathrm{Eig}_T(\lambda)$.

(c) *V hat eine Basis aus Eigenvektoren.*

Proof. Wähle eine Basis, also einen Isomorphismus $\phi : V \xrightarrow{\cong} K^n$. Sei $A \in \mathrm{M}_n(K)$ die Matrix, die T darstellt, die also $T = \phi^{-1}A\phi$ erfüllt. Dann ist $\mathrm{Eig}_A(\lambda) = \phi\Big(\mathrm{Eig}_T(\lambda)\Big)$ und die Aequivalenz folgt aus Satz 3.5.22. □

Satz 3.5.25. (a) *Ist $T : V \to V$ diagonalisierbar und ist $U \subset V$ ein Unterraum mit $T(U) \subset U$, dann ist $T|_U : U \to U$ diagonalisierbar.*

(b) *Sind $S, T : V \to V$ linear, beide diagonalisierbar und gilt $ST = TS$, dann sind die beiden lineare Abbildungen simultan diagonalisierbar, d.h. es gibt eine Basis $v_1, \dots, v_n$ bestehend aus simultanen Eigenvektoren.*

Beweis. (a) Seien $\lambda_, \dots, \lambda_k$ die verschiedenen Eigenwerte von T und $V_j = \mathrm{Eig}(T, \lambda_j)$. Dann gilt $V = V_1 \oplus \dots \oplus V_k$. Sei nun $u \in U$, dann gilt $u = u_1 + \dots + u_k$ mit eindeutig bestimmten $u_j \in V_j$. Wir müssen zeigen, dass jedes u_j wieder in U liegt.

Anwendung von T einerseits und Multiplikation mit λ_k andererseits liefert

$$\begin{aligned} Tu &= \lambda_1 u_1 + \dots + \lambda_k u_k, \\ \lambda_k u &= \lambda_k u_1 + \dots + \lambda_k u_k. \end{aligned}$$

Wir nehmen die Differenz und sehen

$$(\lambda_1 - \lambda_k)u_1 + \dots + (\lambda_{k-1} - \lambda_k)u_{k-1} \in U.$$

Wir wenden denselben Schluss nochmals an:

$$(\lambda_1 - \lambda_k)(\lambda_1 - \lambda_{k-1})u_1 + \dots + (\lambda_{k-2} - \lambda_k)(\lambda_{k-2} - \lambda_{k-1})u_{k-2} \in U.$$

Wir wiederholen dies bis zu

$$(\lambda_1 - \lambda_k)\cdots(\lambda_1 - \lambda_2)u_1 \in U.$$

Da die Eigenwerte verschieden sind, folgt $u_1 \in U$. Aus Symmetriegründen folgt $u_j \in U$ für jedes j.

(b) Sei λ ein Eigenwert von S und v ein Eigenvektor. Dann gilt

$$S(Tv) = T(Sv) = T(\lambda v) = \lambda Tv,$$

also liegt Tv wieder in $\mathrm{Eig}(S,\lambda)$. Anwendung von (a) auf den Unterraum $U = \mathrm{Eig}(S,\lambda)$ zeigt, dass T auf diesem Raum diagonalisierbar ist, d.h. $\mathrm{Eig}(S,\lambda)$ hat eine Basis aus simultanen Eigenvektoren. Da dies für jeden Eigenraum gilt, folgt die Behauptung. □

* * *

3.6 Der Satz von Cayley-Hamilton

Definition 3.6.1. Sei $T : V \to V$ linear. Für jedes $k \in \mathbb{N}$ definiere die k-te Potenz von T durch

$$T^k = \underbrace{T \circ T \circ \cdots \circ T}_{k\text{-mal}}$$

Ferner sei $T^0 = \mathrm{Id}_V$. Ist dann

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n$$

ein Polynom, dann setze

$$f(T) = a_0 \mathrm{Id}_V + a_1 T + \cdots + a_n T^n.$$

Dann ist $f(T)$ eine lineare Abbildung auf V.

Beispiel 3.6.2. Sei $V = K^2$ und $T = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. Sei $f(x) = 1 + x + x^2$. Es ist $T^2 = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}$. Also ist $f(T) = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} + \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} = \begin{pmatrix} 4 & 4 \\ 4 & 8 \end{pmatrix}$.
Ist hingegen $g(x) = 1 - 3x + x^2$, so gilt

$$g(T) = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} - 3\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} + \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0.$$

Satz 3.6.3 (Cayley-Hamilton). *Ist $A \in \mathrm{M}_n(K)$ und ist $f(x) = \chi_A(x) = \det(x - A)$ das charakteristische Polynom, dann gilt*

$$f(A) = 0.$$

Beweis. Für jedes Polynom $g(x)$ gilt $g(S^{-1}AS) = S^{-1}g(A)S$. Der Körper K besitzt eine Erweiterung $\widehat{K} \supset K$, so dass $\widehat{K}$ algebraisch abgeschlossen ist. Für die Aussage des Satzes spielt der Koerper keine Rolle, also koennen wir K durch $\widehat{K}$ ersetzen und annehmen, dass das charakteristische Polynom χ_A in Linearfaktoren zerfällt. Dann ist die Matrix A trigonalisierbar. Wir können also annehmen, dass $A = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ gilt.

Es ist $f(x) = (x - \lambda_1) \cdots (x - \lambda_n)$.

Wir zeigen durch Induktion, dass

$$(A - \lambda_1) \cdots (A - \lambda_j) e_j = 0.$$

Für $j = 1$ ist dies klar.

$j \to j + 1$: Da A eine obere Dreiecksmatrix ist, liegt $(A - \lambda_{j+1}) e_{j+1}$ in $\mathrm{Spann}(e_1, \dots, e_j)$. Nach Induktionsvoraussetzung gilt

$$\begin{aligned} (A - \lambda_1) \cdots (A - \lambda_j) e_1 &= 0 \\ &\vdots \\ (A - \lambda_1) \cdots (A - \lambda_j) e_j &= 0 \end{aligned}$$

Damit folgt

$$(A - \lambda_1) \cdots (A - \lambda_j) \underbrace{(A - \lambda_{j+1}) e_{j+1}}_{\in \mathrm{Spann}(e_1, \dots, e_j)} = 0. \qquad \square$$

Beispiel 3.6.4. Sei $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Dann ist

$$\chi_A(x) = \det \begin{pmatrix} x - 1 & -2 \\ -3 & x - 4 \end{pmatrix} = (x - 1)(1 - 4) - 6 = x^2 - 5x - 2.$$

Ferner ist $A^2 = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}$. Daher

$$\chi_A(A) = A^2 - 5A - 2I_2 = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix} - 5 \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Definition 3.6.5. Ein Polynom f heißt **normiert**, falls der Leitkoeffizient gleich 1 ist, wenn also f von der Form

$$a_0 + a_1 x + \cdots + a_{n-1} x^{n-1} + x^n$$

ist.

Satz 3.6.6. *Für eine gegebene Matrix $A \in \mathrm{M}_n(K)$ existiert ein eindeutig bestimmtes normiertes Polynom $m_A(x)$, das* **Minimalpolynom**, *so dass*

$$m_A(A) = 0$$

und der Grad von m_A ist minimal unter allen Polynomen g mit $g(A) = 0$.
Ist g ein Polynom mit $g(A) = 0$, so existiert ein Polynom q, so dass

$$g(x) = q(x) m_A(x).$$

Sind A und B konjugiert, so ist $m_A = m_B$.

Beweis. Nach dem Satz von Cayley-Hamilton gibt es ein Polynom $f \neq 0$ mit $f(A) = 0$. Sei ν der minimale Grad eines Polynoms g mit $g(A) = 0$. Sei $m(x)$ ein Polynom vom Grad ν mit $m(A) = 0$. Indem wir durch den Leitkoeffizienten teilen, können wir m als normiert annehmen. Damit ist die Existenz bewiesen. Wir brauchen die Eindeutigkeit. Sei also $n(x)$ ein weiteres normiertes Polynom vom Grad ν mit $n(A) = 0$. Dann ist der Grad von $g(x) = m(x) - n(x)$ echt kleiner als ν und es ist $g(A) = m(A) - n(A) = 0$. da der Grad ν minimal war, ist $g(x) = 0$, also $m = n$. Damit ist der erste Teil des Satzes bewiesen.

Für den zweiten sei $g \neq 0$ ein Polynom mit $g(A) = 0$. Dann ist der Grad von g größer oder gleich ν. Nach Polynomdivision existieren Polynome $q(x), r(x)$ mit $\mathrm{grad}(r) < \nu$ und

$$g(x) = q(x)m(x) + r(x).$$

Es folgt $r(A) = g(A) - q(A)m(A) = 0 - 0 = 0$. Da $\mathrm{grad}(r) < \nu$, folgt $r = 0$, also $g(x) = q(x)m(x)$.

Schliesslich die letzte Aussage des Satzes. Sei $B = S^{-1}AS$, dann gilt $m_A(B) = m_A(S^{-1}AS) = S^{-1}m_A(A)S = 0$. Damit existiert ein Polynom q mit $m_A(x) = q(x)m_B(x)$. Der Schluss funktioniert mit umgekehrten Rollen ebensogut, also existiert auch ein Polynom $p(x)$ mit $m_B(x) = p(x)m_A(x)$. Daher $m_A(x) = p(x)q(x)m_A(x)$, woraus folgt, dass p und q konstant sind. Da m_A und m_B normiert sind, sind $p = q = 1$, also $m_A = m_B$. □

Beispiel 3.6.7. Sei $A = \begin{pmatrix} 1 & & \\ & 2 & \\ & & 2 \end{pmatrix}$. Dann ist $\chi_A(x) = (x-1)(x-2)^2$. Das Minimalpolynom m_A muss ein Teiler von χ_A sein, also bleibt $(x-1), (1-2), (x-2)^2, (x-1)(x-2), (x-1)(x-2)^2$. Nur die letzten beiden annullieren A, also ist $m_A(x) = (x-1)(x-2)$.

Proposition 3.6.8. *Das charakteristische Polynom von A zerfalle in Linearfaktoren (etwa wenn $K = \bar{K}$). Seien $\lambda_1, \dots, \lambda_k$ die verschiedenen Eigenwerte von A. Sei*

$$\chi_A(x) = \prod_{j=1}^{k} (x - \lambda_j)^{n_j}$$

das charakteristische Polynom. Dann ist das Minimalpolynom m_A von der Form

$$m_A(x) = \prod_{j=1}^{k} (x - \lambda_j)^{m_j},$$

wobei gilt

$$1 \le m_j \le n_j$$

für $j = 1, \dots, k$. Insbesondere hat das Minimalpolynom dieselben Nullstellen wie das charakteristische Polynom.

Beweis. Da m_A das charakteristische Polynom teilt, muss es von der angegebenen Form sein für Exponenten $0 \le m_j \le n$. Es bleibt zu zeigen, dass jeder Eigenwert eine Nullstelle von m_A ist. Sei λ_i ein Eigenwert und sei v ein Eigenvektor zu diesem Eigenwert, also $Av = \lambda_i v$. Es gilt

$$\begin{aligned} 0 = m_A(A)v &= \prod_j (A - \lambda_j)^{m_j} v \\ &= \prod_j (\lambda_i - \lambda_j)^{m_j} v \end{aligned}$$

Der Faktor $\prod_j (\lambda_i - \lambda_j)^{m_j}$ kann aber nur Null sein, wenn $m_i \ge 1$ ist. □

Beispiele 3.6.9. (a) Ist $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$, dann ist $\chi_A(x) = m_A(x) = (x-2)^2$.

(b) Ist $B = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, dann ist $\chi_B(B) = (x-2)^2$ und $m_B(x) = x - 2$.

Satz 3.6.10. *Ist $A \in \mathrm{M}_n(K)$ diagonalisierbar, dann hat das Minimalpolynom nur einfache Nullstellen. Das heißt, sind $\lambda_1, \dots, \lambda_k$ die verschiedenen Eigenwerte, dann gilt*

$$m_A(x) = \prod_{j=1}^{k}(x - \lambda_j).$$

Beweis. Indem wir A durch $S^{-1}AS$ ersetzen für ein $S \in \mathrm{GL}_n(K)$, können wir annehmen, dass A von Diagonalgestalt ist. Dann ist $\prod_{j=1}^{k}(A - \lambda_j \mathrm{Id})$ ein Produkt von Diagonalmatrizen. Für jede Position $\nu = 1, \dots, n$ hat mindestens einer der Faktoren $(A - \lambda_j)$ eine Null in der ν-ten Position. Daher ist das Produkt gleich Null. □

3.7 Nilpotente Endomorphismen und Hauptraumzerlegung

Lemma 3.7.1. *Sei $T : V \to V$ linear und sei $W \subset V$ ein Unterraum mit*

$$T(W) = T(V) \subset W.$$

Dann hat T eine Matrixdarstellung der Form $\begin{pmatrix} T|_W & 0 \\ 0 & 0 \end{pmatrix}$. Genauer laesst sich jede Basis von W zu einer Basis von V verlaengern, in der T eine solche Matrixdarstellung hat.

Proof. Sei $k = \dim W$. Ist $v_1, \dots, v_n$ eine Basis, in der T eine solche Darstellung hat, dann ist $v_1, \dots, v_k$ eine Basis von W und diese laesst sich durch jede andere Basis von W ersetzen. Damit folgt der Zusatz aus der Hauptaussage. Diese ist aequivalent dazu, dass es einen Unterraum $U \subset \ker(T)$ gibt, so dass $V = W \oplus U$. Für gegebenes $v \in V$ existiert ein $w \in W$ mit $T(v) = T(w)$, also $v = (v - w) + w \in \ker(T) + W$ und daher folgt $V = \ker(T) + W$. Wählt man in $\ker(T)$ einen Komplementärraum U zu $W \cap \ker(T)$, so folgt $V = W \oplus U$ wie verlangt. □

Definition 3.7.2. Ein Endomorphismus $T : V \to v$ heißt **nilpotent**, falls es ein $m \in \mathbb{N}$ gibt, so dass

$$T^m = 0$$

gilt. Ist T nilpotent, so gibt es zu jedem $v \in V$ eine Zahl $m(v) \in \mathbb{N}_0$ so dass $T^{m(v)}v \neq 0 = T^{m(v)+1}v$.

Satz 3.7.3. *Sei $T : V \to V$ nilpotent und $\dim V < \infty$. Dann existiert eine Basis, genannt* **Jordan-Basis**, *bezüglich der T dargestellt wird durch eine Matrix der Form*

$$J = \begin{pmatrix} J_{d_1} & & \\ & \ddots & \\ & & J_{d_k} \end{pmatrix},$$

wobei J_d die $d \times d$-Matrix

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

bezeichnet. Man nennt die Matrizen J_d auch **Jordan-Bloecke** *und die Gesamtmatrix J eine* **Jordan-Matrix**. *Die Jordan-Bloecke sind in bis auf die Reihenfolge eindeutig bestimmt.*

Beweis. Induktion nach $n = \dim(V)$. Für $n = 0$ ist nichts zu zeigen. Sei also $n > 0$. Wir nehmen $T \neq 0$ an. Da T nicht invertierbar ist, ist $U := \mathrm{Bild}(T) \neq V$, also $m = \dim U < n$. Da $T(U) \subset U$ gibt es nach Induktionsvoraussetzung eine Jordan-Basis von U. Sei dann $v_1, \dots, v_r$ Teil dieser Basis, das ein Jordan-Kästchen aufspannt. Es ist dann $T(v_{j+1}) = v_j$ und v_r liegt nicht in $T\Big(\mathrm{Bild}(T)\Big)$. Ferner liegt v_1 im Kern von T und $\dim\Big(\mathrm{Bild}(T) \cap \ker(T)\Big)$ ist genau die Anzahl der Jordan Bloecke von $T|_{\mathrm{Bild}(T)} : \mathrm{Bild}(T) \to \mathrm{Bild}(T)$.

Wähle ein $v_{r+1} \in V$ mit $T(v_{r+1}) = v_r$, so ist auf dem von $v_1, \dots, v_{r+1}$ aufgespannten Raum das Jordan-Kästchen um eins verlängert. Wir tun dies für jeden Jordan Block und erhalten durch Hinzunahme der neuen Vektoren v_{r+1} einen T-stabilen Unterraum W mit $U \subset W \subset V$ und $T(W) = T(V) = \mathrm{Bild}(T)$. Nach Lemma 3.7.1 folgt die Behauptung.

Nun zur Eindeutigkeit. Aus der Definition einer Jordan-Matrix folgt, dass

$$\dim\ker(T^k) - \dim\ker(T^{k-1})$$

genau die Anzahl der Jordan-Bloecke der Länge k ist. Je zwei Jordan-Matrizen zu T haben also dieselben Bloecke bis auf Permutation. □

Definition 3.7.4. Sei $T : V \to V$ linear. Ein Unterraum $U \subset V$ heisst T**-stabil**, falls

$$T(U) \subset U.$$

Eine Zerlegung $V = U \oplus W$ heisst T**-stabile Zerlegung**, falls U und W beide T-stabil sind. Wählt man Basen von U und W, dann stellt sich T durch eine Matrix der Form $\begin{pmatrix} T|_U & 0 \\ 0 & T|_W \end{pmatrix}$ dar.

Lemma 3.7.5. *Sei* $\dim V = n$ *und* $T : V \to V$ *linear. Dann gilt für jedes* $k \in \mathbb{N}$*, dass*

$$\ker T^{n+k} = \ker T^n \quad \textit{und} \quad \mathrm{Bild}(T^{n+k}) = \mathrm{Bild}(T^n).$$

Ferner haben wir eine T*-stabile Zerlegung*

$$V = \ker(T^n) \oplus \mathrm{Bild}(T^n).$$

Beweis. Die T-Stabilität der Räume ist klar. Seien $U_j = \ker T^j$ und $W_j = \mathrm{Bild}(T_j)$, dann gilt $V_0 = 0 \subset V_1 \subset \cdots$ und $V = W_0 \supset W_1 \supset W_2 \supset \cdots$. Die Dimensionsformel besagt, dass $n = \dim U_j + \dim W_j$ für jedes j gilt. Wir zeigen, dass es ein $k \in \mathbb{N}_0$ gibt, so dass $U_0 \ne U_1 \ne \cdots \ne U_k = U_{k+1} = \dots$ und $V = W_0 \supsetneq W_1 \supsetneq \cdots \supsetneq W_k = W_{k+1} = \cdots$. Hierzu reicht es zu zeigen, dass aus $U_k = U_{k+1}$ schon $U_{k+1} = U_{k+2}$ folgt. Die Inklusion "$\subset$" gilt immer. Sei also $v \in \ker(T^{k+2})$, dann ist $Tv \in \ker(T^{k+1}) = \ker(T^k)$, so dass $v \in \ker(T^{k+1})$ ist.

Da nun die Dimension bei jedem Schritt $U_j \ne U_{j+1}$ um mindestens 1 wächst, ist spätestens bei U_n Schluss. Die duale Aussage über die Bilder folgt aus der Dimensionsformel.

Für die letzte Aussage reicht es $\ker(T^n) \cap \mathrm{Bild}(T^n) = 0$ zu zeigen, da dann die Behauptung aus der Dimensionsformel folgt. Sei $v \in \ker(T^n) \cap \mathrm{Bild}(T^n)$. Dann folgt $v = T^n u$ für ein $u \in V$ und wegen $T^n v = 0$ folgt $T^{2n} u = 0$, dann ist nach dem ersten Teil schon $T^n u = 0$, also $v = 0$. □

Definition 3.7.6. Für $\lambda \in K$ sei

$$\ker\Big((T-\lambda)^n\Big)$$

der **Hauptraum** zu λ.

Satz 3.7.7 (Hauptraum-Zerlegung)**.** *Sei* $\dim V = n$ *und* $T : V \to V$ *linear. Nimm an, dass das charakteristische Polynom in Linearfaktoren zerfällt, also*

$$\chi_T(x) = \prod_{j=1}^k (x-\lambda_j)^{n_j}.$$

Dann gilt

$$V = \bigoplus_{j=1}^k \ker(T-\lambda_j)^n.$$

Beweis. Wir zeigen nun den Satz mit Induktion nach n. Für $n = 1$ ist nichts zu zeigen. Sei also $n \ge 2$ und die Behauptung für Räume kleinerer Dimension bereits gezeigt. Nach Lemma 3.7.5 ist

$$V = \ker(T-\lambda_1)^n \oplus \mathrm{Bild}(T-\lambda_1)^n.$$

Da $\ker(T-\lambda_1)^n \neq 0$ hat der Raum $W = \mathrm{Bild}(T-\lambda_1)^n$ eine Dimension $< n$. Ferner ist $T(W) \subset W$, nach der Induktionsvorausseztung zerfällt W in eine direkte Summe von Haupträumen. Auf W ist λ_1 kein Eigenwert, denn der λ_1-Eigenraum liegt ganz in $\ker(T-\lambda_1)^n$. Damit folgt

$$V = \bigoplus_{j=1}^{k} \ker(T-\lambda_j)^n$$

wie gewünscht. □

3.8 Jordan-Matrizen

Definition 3.8.1. Ein **Jordan-Block** ist eine Matrix der Form

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix},$$

wobei $\lambda \in K$ und $J_k(\lambda) \in \mathrm{M}_k(K)$ ist.

Eine **Jordan-Matrix** ist eine Matrix der Form

$$\begin{pmatrix} J_{k_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{k_s}(\lambda_s) \end{pmatrix},$$

d.h., es sind Jordan-Blöcke auf der Diagonale, sonst Nullen.

Beispiele 3.8.2. (a) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ ist ein Jordan-Block.

(b) $\begin{pmatrix} 1 & 1 & \\ & 1 & \\ & & 2 \end{pmatrix}$ ist eine Jordan-Matrix.

(c) $\begin{pmatrix} 1 & 1 & \\ & 2 & \\ & & 1 \end{pmatrix}$ ist keine Jordan-Matrix.

Satz 3.8.3 (Jordan-Normalform-Satz)**.** *Sei $A \in \mathrm{M}_n(K)$. Das charakteristische Polynom χ_A zerfalle in Linearfaktoren. Dann ist A konjugiert zu einer Jordan-Matrix*

$$\begin{pmatrix} J_{k_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{k_s}(\lambda_s) \end{pmatrix}.$$

Hierbei sind $\lambda_1, \dots, \lambda_s$ die (nicht notwendig verschiedenen) Eigenwerte von A. Die Jordan-Matrix ist eindeutig bestimmt bis auf die Reihenfolge der Blöcke. Man nennt diese Jordan-Matrix die **Jordan-Normalform** *von A.*

Beweis. Sei $K^n = \ker(A - \lambda_1)^n \oplus \cdots \oplus \ker(A - \lambda_k)^n$ die Hauptraumzerlegung. Da die Haupträume invariant sind, reicht es zu zeigen, dass A auf jedem dieser Räume durch eine Jordan-Matrix dargestellt wird. Wir können also annehmen, dass A nur einen einzigen Eigenwert λ hat. Dann ist $(A - \lambda)$ nilpotent, es gibt also eine Basis bezgl der $(A - \lambda)$ in Jordan-Form ist, dann ist in derselben Basis auch A in Jordan-Form. Die Eindeutigkeit der Jordan-Bloecke folgt aus der Eindeutigkeit bei nilpotenten Matrizen. □

Beispiele 3.8.4. (a) Die Jordan-Normalform von $\begin{pmatrix} 1 & 1 \\ & 2 \end{pmatrix}$ ist $\begin{pmatrix} 1 & \\ & 2 \end{pmatrix}$.

(b) Die Jordan-Normalform von $\begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$ ist $\begin{pmatrix} 2 & 1 \\ & 2 \end{pmatrix}$.

Proposition 3.8.5. *Sei $A \in \mathrm{M}_n(K)$ und seien $\lambda_1, \dots, \lambda_k$ die verschiedenen Eigenwerte. Sei*

$$m_A(x) = \prod_{j=1}^{k} (x - \lambda_j)^{m_j}$$

das Minimalpolynom. Dann ist m_j die Länge des längsten Jordan-Blocks mit dem Eigenwert λ_j.

Beweis. Wir nehmen an, dass A in Jordan-Normalform ist und die Blöcke nach Eigenwerten geordnet sind. Wir können dann A durch die Untermatrix zu einem festen Eigenwert ersetzen und also annehmen, dass A nur einen einzigen Eigenwert λ hat. Dann besteht $A - \lambda$ ebenfalls aus Jordan-Blöcken, aber zum Eigenwert Null. Man stellt nun fest, dass $J_k(0)^{k-1} \neq 0$, aber $J_k(0)^k = 0$ ist. Damit folgt die Behauptung. □

Beispiele 3.8.6. (a) Ist $A = \begin{pmatrix} 2 & 1 & \\ & 2 & \\ & & 2 \end{pmatrix}$, dann ist $m_A(x) = (x-2)^2$.

(b) Ist $A = \begin{pmatrix} 2 & 1 & \\ & 2 & 1 \\ & & 2 \end{pmatrix}$, dann ist $m_A(x) = (x-2)^3$.

Proposition 3.8.7. *Sei K algebraisch abgeschlossen. Für jede Matrix A in $\mathrm{M}_2(K)$ oder $\mathrm{M}_3(K)$ ist die JNF durch das charakteristische Polynom und das Minimalpolynom eindeutig festgelegt.*

Beweis. Es reicht, den Fall $A \in \mathrm{M}_3(K)$ zu betrachten. In diesem Fall hat A maximal drei verschiedene Eigenwerte und für das charakteristische Polynom gibt es folgende Fälle:

1. Fall: Es gibt drei verschiedene Eigenwerte, $\chi_A(x) = (x-\lambda)(x-\mu)(x-\nu)$. Dann ist A diagonalisierbar.

2. Fall: Es gibt zwei verschieden Eigenwerte: $\chi)A(x) = (x-\lambda)^2(x-\mu)$. Ist dann $m_A(x) = (x-\lambda)(x-\mu)$, so ist A diagonalisierbar. Ist $m_A(x) = (x-\lambda)^2(x-\mu)$, so hat A die JNF $= \begin{pmatrix} \lambda & 1 & \\ & \lambda & \\ & & \mu \end{pmatrix}$.

3. Fall: Es gibt nur einen Eigenwert, $\chi_A(x) = (x-\lambda)^3$. Ist dann $m_A(x) = (x-\lambda)$, dann ist A diagonalisierbar. Ist $m_A(x) = (x-\lambda)^2$, dann gibt es zwei Jordan-Blöcke der Längen 2 und 1. Ist schliesslich $m_A(x) = (x-\lambda)^3$ so gibt es einen Jordan-Block der Länge 3. □

Beispiel 3.8.8. Für $n = 4$ reicht diese Information nicht mehr aus, denn die Matrizen

$$\begin{pmatrix} 1 & 1 & & \\ & 1 & & \\ & & 1 & 1 \\ & & & 1 \end{pmatrix} \text{und} \begin{pmatrix} 1 & 1 & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

haben dasselbe charakteristische und dasselbe Minimalpolynom.

Proposition 3.8.9. *Sei $K = \bar{K}$ und $\lambda \in K$ ein Eigenwert der Matrix $A \in \mathrm{M}_n(K)$. Die geometrische Vielfachheit*

$$\dim \mathrm{Eig}(A, \lambda)$$

ist gleich der Anzahl der Jordan-Blöcke zum Eigenwert λ.

Beweis. Es reicht, anzunehmen, dass λ der einzige Eigenwert ist. Sei dann A in Jordan-Form und seien $J_{k_1}(\lambda), \dots, J_{k_s}(\lambda)$ die entsprechenden Jordan-Blöcke. Dann sind die Vektoren $e_1, e_{k_1+1}, \dots, e_{k_1+\dots+k_{s-1}+1}$ eine Basis des Eigenraums, also folgt $\dim \mathrm{Eig}(A, \lambda) = s$. □

Beispiel 3.8.10. Die JNF von $\begin{pmatrix} 3 & 1 & -1 & 1 \\ 0 & 2 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & -1 & 0 & 2 \end{pmatrix}$ ist $\begin{pmatrix} 2 & 1 & & \\ & 2 & & \\ & & 2 & 1 \\ & & & 2 \end{pmatrix}$.

Proposition 3.8.11. *Sei $K = \bar{K}$ und $A \in \mathrm{M}_n(K)$. Dann ist*

$$\dim\ker(A-\lambda)^{j+1} - \dim\ker(A-\lambda)^j$$

gleich der Anzahl der Jordan-Blöcke mit Eigenwert λ und Länge $\geq j+1$.

Beweis. Für einen Jordan-Block $A = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$ gilt

$$\ker(A-\lambda)^j = \mathrm{Spann}(e_1, \ldots, e_j).$$

Damit folgt die Behauptung. □

Damit ergibt sich der

Algorithmus zur Bestimmung der Jordan-Normalform

1. Bestimme $\chi_A(x)$.
2. Bestimme die Nullstellen, also die Eigenwerte $\lambda_1, \ldots, \lambda_n$. Algebraische Vielfachheit = Länge der Jordan-Matrix zum jeweiligen Eigenwert.
3. Bestimme die Zahlen $N_j = \dim\ker(A-\lambda)^j$. Dann gibt es N_1 Jordan-Blöcke. Davon haben $N_2 - N_1$ die Länge ≥ 2 und $N_3 - N_2$ die Länge ≥ 3 usf.

* * *

4 Analytische Geometrie

4.1 Skalarprodukte

Definition 4.1.1. Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Ein **Skalarprodukt** auf V ist eine Abbildung

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

mit folgenden Eigenschaften:

(a) für jedes $w \in V$ ist die Abbildung $v \mapsto \langle v, w \rangle$ linear,

(b) für alle $v, w \in V$ gilt
$$\langle w, v \rangle = \overline{\langle v, w \rangle},$$

(c) für jedes $v \in V$ gilt $\langle v, v \rangle \geq 0$ und
$$\langle v, v \rangle = 0 \qquad \Leftrightarrow \qquad v = 0.$$

Man sagt für (b) auch, ein Skalakprodukt ist **antisymmetrisch** und für (c), es ist **positiv definit**.

Eine Abbildung $h : V \times V \to \mathbb{K}$, die nur (a) und (b) erfüllt, heisst **Hermitesche Form**. Also ist ein Sklararprodukt dasselbe wie eine positiv definite Hermitesche Form.

Beispiel 4.1.2. Das **Standard Skalarprodukt** auf $\mathbb{R}^n$ ist

$$\langle x, y \rangle = x^t y = x_1 y_1 + \cdots + x_n y_n$$

und auf $\mathbb{C}^n$
$$\langle z, w \rangle = z_1 \overline{w_1} + \cdots + z_n \overline{w_n}.$$

Dieses Skalarprodukt kann auch in der Form

$$\langle v, w \rangle = v^t \, \overline{w}$$

geschrieben werden, wobei $\overline{w} = \begin{pmatrix} \overline{w_1} \\ \vdots \\ \overline{w_n} \end{pmatrix}$.

Definition 4.1.3. Für eine Matrix $A \in \mathrm{M}_n(\mathbb{K})$ setze

$$A^* = \overline{A}^t.$$

Diese Matrix heisst die **adjungierte** zu A. Für $n = 2$ gilt zum Beispiel

$\begin{pmatrix} z & w \\ u & v \end{pmatrix}^* = \begin{pmatrix} \overline{z} & \overline{u} \\ \overline{w} & \overline{v} \end{pmatrix}$. Die Matrix A heisst **selbstadjungiert**, falls

$$A = A^*.$$

Eine Matrix A ist also genau dann selbstadjungiert, wenn $A_{i,j} = \overline{A_{j,i}}$ für alle $1 \le i, j \le n$ gilt.

Lemma 4.1.4. *Sei $A \in \mathrm{M}_n(\mathbb{K})$. Dann sind äquivalent*

(a) *Die Abbidlung $b_A : V \times V \to \mathbb{K}$,*

$$b_A(v,w) = \langle Av, w\rangle = (Av)^t\overline{w}$$

ist ein Skalarprodukt.

(b) *Die Matrix A ist*

  (i) **selbstadjungiert** *und*

  (ii) **positiv***, d.h., es gilt $v^tA\overline{v} > 0$ für jeden Vektor $v \in V \smallsetminus \{0\}$.*

Beweis. (a)⇒(b): Sei b_A ein Skalarprodukt. Dann gilt

$$A_{i,j} = e_iAe_j = b_A(e_i,e_j) = \overline{b_A(e_j,e_i)} = \overline{A_{j,i}}.$$

Damit folgt (i). Die zweite Aussage (ii) ist nur eine Umformulierung der positiven Definitheit.

(b)⇒(a): Es gilt

$$b_A(w,v) = w^t\,A\,\overline{v} = (w^t\,A\,\overline{v})^t = \overline{v}^t\,A^t\,w = \overline{v}^t\,\overline{A}\,w = \overline{v^t\,A\,\overline{w}} = \overline{b_A(v,w)}.$$

Der Rest ist klar. □

Bemerkung 4.1.5. Ist $\langle .,.\rangle$ ein Skalarprodukt auf einem $\mathbb{C}$-Vektorraum V, dann ist für jedes $v \in V$ die Abbildung $T = T_v : w \mapsto \langle v,w\rangle$ eine **anti-lineare Abbildung**, d.h., es gilt

$$T(\lambda w + \mu w') = \overline{\lambda}T(w) + \overline{\mu}T(w').$$

Definition 4.1.6. Sei $\mathbb{K}$ der Körper $\mathbb{R}$ oder $\mathbb{C}$. Sei $\langle .,.\rangle$ ein Skalarprodukt auf dem $\mathbb{K}$-Vektorraum V. Die **euklidische Norm** eines Vektors $v \in V$ ist

$$\|v\| = \sqrt{\langle v,v\rangle}.$$

Man sagt: zwei Vektoren $v, w \in V$ stehen **senkrecht** aufeinander, falls $\langle v,w\rangle = 0$ ist. Man beachte, dass die Norm im $\mathbb{R}^3$ genau die geometrische Länge eines Vektors beschreibt.

Satz 4.1.7 (Satz des Pythagoras)**.** *Stehen die Vektoren v, w senkrecht aufeinander, dann ist*

$$\|v\|^2 + \|w\|^2 = \|v + w\|^2 .$$

Beweis. Es gilt

$$\begin{aligned}\|v + w\|^2 &= \langle v + w, v + w\rangle \\ &= \langle v, v\rangle + \underbrace{\langle v, w\rangle}_{=0} + \underbrace{\langle w, v\rangle}_{=0} + \langle w, w\rangle \\ &= \|v\|^2 + \|w\|^2 .\end{aligned}$$

□

Definition 4.1.8. Ein reeller Vektorraum mit einem Skalarprodukt heißt **euklidischer Raum**. Ein komplexer Vektorraum mit einem Skalarprodukt heißt **unitärer Raum**.

Satz 4.1.9 (Cauchy-Schwarz-Ungleichung)**.** *Sei V ein unitärer oder euklidischer Vektorraum. Dann gilt für alle $v, w \in V$:*

$$|\langle v, w\rangle| \le \|v\| \, \|w\| .$$

Beweis. Wir können $w \neq 0$ annehmen. Für jedes $t \in \mathbb{R}$ gilt

$$0 \le \|v - tw\|^2 = \langle v - tw, v - tw\rangle = \|v\|^2 - 2t \operatorname{Re}\langle w, v\rangle + t^2 \|w\|^2 .$$

Dieses quadratische Polynom in t nimmt sein Minimum in $t = \frac{\operatorname{Re}\langle v, w\rangle}{\|w\|^2}$ an. Setzen wir diesen Wert ein, folgt

$$0 \le \|v\|^2 - 2\frac{\operatorname{Re}\langle v, w\rangle^2}{\|w\|^2} + \frac{\operatorname{Re}\langle v, w\rangle^2}{\|w\|^4} \|w\|^2 ,$$

also

$$\operatorname{Re}\langle v, w\rangle^2 \le \|v\|^2 \|w\|^2 .$$

Es gibt nun ein $\theta \in \mathbb{K}$ mit $|\theta| = 1$ so dass $\theta \langle v, w\rangle = |\langle v, w\rangle|$ gilt. Indem wir v durch θv ersetzen, folgt aus dem obigen

$$\begin{aligned}|\langle v, w\rangle|^2 &= \operatorname{Re}\left(|\langle v, w\rangle|^2\right) = \operatorname{Re}\left(\theta^2 \langle v, w\rangle^2\right) \\ &= \operatorname{Re}\left(\langle \theta v, w\rangle^2\right) = \operatorname{Re}(\langle \theta v, w\rangle)^2 \le \|\theta v\|^2 \|w\|^2 = \|v\|^2 \|w\|^2 .\end{aligned}$$

□

Satz 4.1.10. *Es gilt*

- $\|v\| \geq 0$ *und* $\|v\| = 0 \;\Leftrightarrow\; v = 0$ *Definitheit*
- $\|\lambda v\| = |\lambda| \, \|v\|$ *Multiplikativität*
- $\|v + w\| \leq \|v\| + \|w\|$ *Dreiecksungleichung*

Eine Abbil;dung $V \to [0, \infty)$ mit diesen drei Eigenschaften nennt man eine **Normabbildung**.

Beweis. Die ersten beiden Eigenschaften sind klar. Zur Dreiecksungleichung rechne mit der Cauchy-Schwarz-Ungleichung:

$$\begin{aligned}
\|v + w\|^2 &= \langle v + w, v + w \rangle \\
&= \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle \\
&\leq \langle v, v \rangle + 2|\langle v, w \rangle| + \langle w, w \rangle \\
&\leq \|v\|^2 + \|w\|^2 + 2\,\|v\|\,\|w\| \\
&= (\|v\| + \|w\|)^2
\end{aligned}$$ □

* * *

4.2 Projektion und Orthonormalisierung

Definition 4.2.1. Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Sei V ein endlich-dimensionaler Vektorraum mit Skalarprodukt. Die Menge $\{e_1, \dots, e_n\}$ von V heißt **orthonormal**, wenn

$$\langle e_i, e_j \rangle = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases}$$

Ist die Menge überdies eine Basis, so heißt sie **Orthonormalbasis**. Als Beispiel betrachte das Standard-Skalarprodukt und die Standard-Basis im Fall $V = \mathbb{K}^n$.

Satz 4.2.2 (Gram-Schmidt-Verfahren)**.** *Jeder euklidische oder untäre Raum hat eine Orthonormalbasis. Darüberhinaus gilt: Sei $v_1, \dots, v_n$ eine gegebene Basis von V. Dann existiert eine Orthonormalbasis $e_1, \dots, e_n$ mit*

$$\mathrm{Spann}(v_1, \dots, v_k) = \mathrm{Spann}(e_1, \dots, e_k)$$

für jedes $1 \leq k \leq n$.

Beweis. Wir konstruieren die e_j induktiv. Als Induktionsanfang sei $e_1 = \frac{1}{\|v_1\|} v_1$. Im Induktionsschritt sei eine orthonormale Menge $e_1, \dots, e_k$ bereits konstruiert mit $\mathrm{Spann}(v_1, \dots, v_j) = \mathrm{Spann}(e_1, \dots, e_j)$ für jedes $1 \le j \le k$. Definiere

$$\tilde{v}_{k+1} = v_{k+1} - \langle v_{k+1}, e_1 \rangle e_1 - \dots - \langle v_{k+1}, e_k \rangle e_k.$$

Da v_{k+1} nicht im Spann der $e_1, \dots, e_k$ liegt, ist $\tilde{v}_{k+1} \neq 0$. Ferner gilt $\left\langle \tilde{v}_{k+1}, e_j \right\rangle = 0$ für jedes $1 \le j \le k$. Definiere nun

$$e_{k+1} = \frac{1}{\|\tilde{v}_{k+1}\|} \tilde{v}_{k+1}.$$

Dann ist $e_1, \dots, e_{k+1}$ orthonormal mit $\mathrm{Spann}(v_1, \dots, v_j) = \mathrm{Spann}(e_1, \dots, e_j)$ für jedes $1 \le j \le k+1$. □

Beispiel 4.2.3. Sei V der reelle Vektorraum der Polynome vom Grad ≤ 1 mit dem Skalarprodukt $\langle f, g \rangle = \int_0^1 f(x) g(x)\, dx$. Wir orthonormalisieren die natürliche Basis $(v_1, v_2) = (1, x)$. Zunächst ist

$$\|v_1\|^2 = \int_0^1 1\, dx = 1,$$

also ist $e_1 = v_1$. Dann ist

$$\langle v_1, v_2 \rangle = \int_0^1 x\, dx = \frac{1}{2}, \quad \langle v_2, v_2 \rangle = \int_0^1 x^2\, dx = \frac{1}{3}.$$

Wir erhalten

$$\tilde{v}_2 = v_2 - \langle v_2, e_1 \rangle e_1 = v_2 - \frac{1}{2} e_1 = x - \frac{1}{2}.$$

Wir rechnen

$$\|\tilde{v}_2\|^2 = \left\langle x - \frac{1}{2}, x - \frac{1}{2} \right\rangle = \int_0^1 x^2\, dx - \int_0^1 x\, dx + \int_0^1 \frac{1}{4}\, dx = \frac{1}{3} - \frac{1}{2} + \frac{1}{4} = \frac{1}{12}.$$

Es folgt $e_2 = \sqrt{12} \tilde{v}_2 = 2\sqrt{3}\left(x - \frac{1}{2}\right) = 2\sqrt{3}x - \sqrt{3}$.

Definition 4.2.4. Sei nun $U \subset V$ ein Unterraum. Definiere den **Orthogonalraum** $U^\perp$ durch

$$U^\perp = \left(v \in V : \langle v, u \rangle = 0\ \forall_{u \in U}\right).$$

Satz 4.2.5. *Es gilt*

$$V = U \oplus U^\perp.$$

Beweis. Wir zeigen zunächst $U \cap U^\perp = 0$. Sei also $u \in U \cap U^\perp$. Dann ist $\langle u, u \rangle = 0$, also $u = 0$.

Es bleibt zu zeigen, dass $V = U + U^\perp$. Sei $e_1, \dots, e_k$ eine ONB von U. Für beliebiges $v \in V$

sei

$$u = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_k \rangle e_k$$

und $w = v - u$. Dann gilt $v = u + w$, ferner ist $u \in U$ und wir müssen zeigen, dass $w \in U^\perp$ ist. Hierzu reicht es, zu zeigen, dass $\langle w, e_j \rangle = 0$ für $1 \le j \le k$ gilt. Es ist

$$\langle w, e_j \rangle = \langle v, e_j \rangle - \langle u, e_j \rangle = \langle v, e_j \rangle - \langle v, e_j \rangle = 0. \qquad \square$$

Definition 4.2.6. Eine lineare Abbildung $P : V \to V$ heisst **Projektion**, wenn $P^2 = P$. In einer Übungaaufgabe wurde gezeigt, dass dann $V = \ker(P) \oplus \mathrm{Bild}(P)$. Eine Projektion P heißt **Orthogonalprojektion**, falls

$$(\mathrm{Bild}\, P) \perp (\ker P).$$

Sei $U \subset V$ ein Unterraum, dann schreiben wir $P_U : V \to V$ für die Orthogonalprojektion mit Bild U und Kern $U^\perp$.

4.3 Selbstadjungierte und normale Abbildungen

Definition 4.3.1. Sei $(V, \langle .,. \rangle)$ ein endlich-dimensionaler $\mathbb{K}$ Vektorraum mit Skalarprodukt. Eine lineare Selbstabbildung $T : V \to V$ heißt auch **Endomorphismus**.

Satz 4.3.2 (Riesz). *Für jede lineare Abbildung $\alpha : V \to \mathbb{K}$ existiert genau ein Vektor v_α, so dass*

$$\alpha(v) = \langle v, v_\alpha \rangle$$

für jedes $v \in V$ gilt.

Beweis. Ist $\alpha = 0$ setzen wir $v_\alpha = 0$. Ist $\alpha \neq 0$, so gibt es wegen $V = U \oplus U^\perp$ ein $v_0 \in U^\perp$ mit $\alpha(v_0) = 1$. Für beliebiges $v \in V$ gilt dann $v - \mu v_0 \in U$ für $\mu = \alpha(v)$, also ist $v = u + \alpha(v)v_0$. Setze $v_\alpha = \frac{1}{\langle v_0, v_0 \rangle} v_0$. Dann gilt

$$\langle v, v_\alpha \rangle = \frac{\langle v, v_0 \rangle}{\langle v_0, v_0 \rangle} = \frac{\langle u + \alpha(v) v_0, v_0 \rangle}{\langle v_0, v_0 \rangle} = \alpha(v).$$

Schließlich zur Eindeutigkeit: ist w ein zweiter Vektor mit dieser Eigenschaft, dann ist $\langle v, v_\alpha - w \rangle = 0$ für jeden Vektor v, also insbesondere für $v = v_\alpha - w$, woraus $w = v_\alpha$ folgt. $\square$

Definition 4.3.3. Sei $T : V \to V$ linear. Für gegebenes $w \in V$ ist die Abbildung

$v \mapsto \langle Tv, w\rangle$ linear, also gibt es einen eindeutig bestimmten Vektor T^*w mit

$$\langle Tv, w\rangle = \langle v, T^*w\rangle$$

für jedes $v \in V$. Die Abbildung $w \mapsto T^*w$ heißt die zu T **adjungierte Abbildung**.

Satz 4.3.4. *Die adjungierte Abbildung T^* ist linear.*

Beweis. Seien $v, w, w' \in V$ und $\lambda \in \mathbb{K}$, so gilt

$$\begin{aligned}\langle v, T^*(\lambda w + w')\rangle &= \langle Tv, \lambda w + w'\rangle \\ &= \overline{\lambda}\,\langle Tv, w\rangle + \langle Tv, w'\rangle \\ &= \overline{\lambda}\,\langle v, T^*w\rangle + \langle v, T^*(w')\rangle \\ &= \langle v, \lambda T^*w + T^*(w')\rangle\,.\end{aligned}$$

Da die Skalarprodukte $\langle v, w\rangle$ für alle v den Vektor w eindeutig festlegen, folgt

$$T^*(\lambda w + w') = \lambda T^*w + T^*(w').$$ □

Proposition 4.3.5. *Seien $S, T : V \to V$ linear. Es gilt*

$$I^* = I, \quad (\lambda T + S)^* = \overline{\lambda}T^* + S^*, \quad (ST)^* = T^*S^*, \quad T^{**} = T.$$

Hier steht I für die Einheitsmatrix und T^{**} für $(T^*)^*$.

Beweis. Die Beweise laufen alle ähnlich. Man nimmt das Skalarprodukt beider Seiten der Gleichung mit einem beliebigen Vektor. Exemplarisch zeigen wir $(ST)^* = T^*S^*$. Für $v, w \in V$ gilt

$$\langle v, (ST)^*w\rangle = \langle STv, w\rangle = \langle Tv, S^*w\rangle = \langle v, T^*S^*w\rangle\,.$$

Da v und w beliebig sind, folgt die Aussage, denn wir haben ja gezeigt, dass der Vektor $(ST)^*w - T^*S^*w$ auf allen Vektoren senkrecht steht, also Null ist. □

Definition 4.3.6. Eine lineare Abbildung T heißt **selbstadjungiert**, falls $T = T^*$ gilt.

Satz 4.3.7. *Ist T selbstadjungiert, dann ist jeder Eigenwert reell.*

Beweis. Sei λ ein Eigenwert und v ein Eigenvektor. Dann gilt

$$\lambda\,\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Tv, v\rangle = \langle v, Tv\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\,\langle v, v\rangle\,.$$

Da $\langle v, v \rangle \neq 0$, folgt $\lambda = \overline{\lambda}$. $\square$

Beispiel 4.3.8. Sei $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \mathrm{M}_2(\mathbb{R})$. Dann muss jeder Eigenwert dieser Matrix reell sein. Wir rechnen

$$\begin{aligned} \chi_A(x) &= \det \begin{pmatrix} x-a & -b \\ -b & x-c \end{pmatrix} = (x-a)(x-c) - b^2 \\ &= x^2 - (a+c)x + ac - b^2 \\ &= \Big(x - \frac{a+c}{2}\Big)^2 - \Big(\frac{(a+c)^2}{4} - ac + b^2\Big). \end{aligned}$$

Damit die Eigenwerte reell sind, muss also $\Big(\frac{(a+c)^2}{4} - ac + b^2\Big) \geq 0$ sein. Es ist aber

$$\begin{aligned} \frac{(a+c)^2}{4} - ac + b^2 &= \frac{a^2 + 2ac + c^2 - 4ac}{4} + b^2 \\ &= \frac{a^2 - 2ac + c^2}{4} + b^2 = \frac{(a-c)^2}{4} + b^2 \geq 0. \end{aligned}$$

* * *

Definition 4.3.9. Ein Endomorphismus $T : V \to V$ eines Raums mit Skalarprodukt heißt **unitär**, falls $TT^* = I$, also $T^* = T^{-1}$, oder auch

$$\langle Tv, Tw \rangle = \langle v, w \rangle \tag{*}$$

für alle $v, w \in V$.

Beispiel 4.3.10. Sei $V = \mathbb{R}^2$, $\theta \in \mathbb{R}$ und sei

$$k(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Dann ist $k(\theta)^t k(\theta) = I$, also ist $k(\theta)$ unitär.

Definition 4.3.11. Eine Matrix $A \in \mathrm{M}_n(\mathbb{C})$ heißt **unitär**, falls A bezüglich des Standard-Skalarproduktes auf $\mathbb{C}^n$ unitär ist, falls also

$$A^* A = I$$

gilt, wobei $A^* = \overline{A}^t$. Sei $\mathrm{U}(n)$ die Menge aller unitären Matrizen in $\mathrm{M}_n(\mathbb{C})$.

Proposition 4.3.12. *Die Menge* $\mathrm{U}(n)$ *ist eine Untergruppe von* $\mathrm{GL}_n(C)$. *Man nennt sie die* **unitäre Gruppe**.

Beweis. Jeder unitäre Endomorphismus ist invertierbar, also gilt $\mathrm{U}(n) \subset \mathrm{GL}_n(\mathbb{C})$. Ist $A \in \mathrm{U}(n)$,so ist $A^{-1} = A^*$ ebenfalls in $\mathrm{U}(n)$. Sind $A, B \in \mathrm{U}(n)$, dann gilt

$$(AB)^*AB = B^*A^*AB = B^*B = I,$$

daher ist $AB \in \mathrm{U}(n)$. □

Korollar 4.3.13. *Eine Matrix $A \in \mathrm{M}_n(\mathbb{C})$ ist genau dann unitär, wenn die Zeilen oder die Spalten eine Orthonormalbasis von $\mathbb{C}^n$ bilden.*

Beweis. Seien $a_1, \dots, a_n$ die Spalten von A. Es gilt

$$(A^*A)_{i,j} = \overline{a}_i^t a_j = \overline{\langle a_i, a_j \rangle}.$$

Also gilt $A^*A = I \iff (a_1, \dots, a_n)$ ist eine ONB. Der Beweis fier die Zeilen geht ebenso indem man $AA^* = I$ benutzt. □

* * *

4.4 Der Spektralsatz über den komplexen Zahlen

Lemma 4.4.1. *Sei V unitär und $T : V \to V$ linear. Sei $U \subset V$ ein Unterraum. Dann gilt*

(a) $TU \subset U \quad \Rightarrow \quad T^*U^\perp \subset U^\perp$.

(b) *Ist $S : V \to V$ linear mit $ST = TS$ und ist $\lambda \in \mathbb{K}$, dann gilt*

$$S\Big(\mathrm{Eig}(T, \lambda)\Big) \subset \mathrm{Eig}(T, \lambda).$$

(c) *Ist T normal und $U = \mathrm{Eig}(T, \lambda)$, dann ist die Zerlegung*

$$V = U \oplus U^\perp$$

stabil unter T und T^*, d.h. beide Operatoren werfen beide Raeume jeweils in sich.

Beweis. (a) Es gelte $TU \subset U$. Sei $v \in U^\perp$. Fuer jedes $u \in U$ gilt dann $Tu \in U$ und damit

$$\langle T^*v, u \rangle = \langle v, Tu \rangle = 0$$

und damit $T^*v \in U^\perp$.

(b) Sei $v \in \mathrm{Eig}(T, \lambda)$, also $Tv = \lambda v$. Dann gilt

$$T\Big(Sv\Big) - S(Tv) = S(\lambda v) = \lambda\, Sv,$$

also $Sv \in \mathrm{Eig}(T, \lambda)$.

(c) Wende (a) und (b) auf T und T^* an. □

Definition 4.4.2. Ein Endomorphismus $T : V \to V$ heißt **normal**, falls

$$TT^* = T^*T$$

gilt. Insbesondere folgt

$$T \text{ selbstadjungiert oder orthogonal/unitär} \quad \Rightarrow \quad T \text{ normal.}$$

Satz 4.4.3 (Spektralsatz für normale Operatoren)**.** *Sei $\mathbb{K} = \mathbb{C}$ und sei $T : V \to V$ linear. Dann sind äquivalent*

(a) *T ist normal.*

(b) *Es gibt eine Orthonormalbasis von V, die aus Eigenvektoren von T besteht.*

(c) *T ist diagonalisierbar und die Eigenräume stehen senkrecht aufeinander.*

Beweis. (a)⇒(b): Induktion nach $n = \dim V$. Für $n = 0$ ist nichts zu zeigen. Sei also $n \geq 1$ und die Behauptung für alle kleineren Dimensionen gezeigt. Es gibt dann einen Eigenwert $\lambda \in \mathbb{C}$ von T. Sei $U = \mathrm{Eig}(T, \lambda)$ der zugehörige Eigenraum. Ist $U = V$, dann ist $T = \lambda \mathrm{Id}$ und jede Orthonormalbasis erfuellt die Behauptung. Sei also $U \neq V$. Nach Lemma 4.4.1 ist die Zerlegung $V = U \oplus U^\perp$ stabil unter T. Nach Induktionsvoraussetzung haben U und $U^\perp$ Orthonormalbasen $e_1, \dots, e_k$ und $e_{k+1}, \dots e_n$ aus Eigenvektoren. Dann ist $e_1, \dots, e_n$ eine Orthonormlabasis von V, bestehend aus Eigenvektoren.

(b)⇒(c) ist klar.

(c)⇒(a): Seien $\lambda_1, \dots, \lambda_s$ die Eigenwerte von T. Dann ist $V = \bigoplus_{j=1}^s \mathrm{Eig}(T, \lambda_j)$. Schreibe $V_j = \mathrm{Eig}(T, \lambda_j)$ und betrachte die lineare Abbildung $S : V \to V$, die jeden Eigenraum V_j in sich wirft und auf diesem Raum gleich $\overline{\lambda}_j \mathrm{Id}$ ist. Fuer $v \in V_i$ und $w \in V_j$ gilt dann

$$\langle Tv, w \rangle = \lambda_i \langle v, w \rangle = \delta_{i,j} \left\langle v, \overline{\lambda_i} w \right\rangle = \langle v, Sw \rangle .$$

Damit ist $S = T^*$ und nach Konstruktion gilt $ST = TS$. □

Korollar 4.4.4. *Ist T normal, so folgt für $\lambda \in \mathbb{C}$,*

$$\mathrm{Eig}(T, \lambda) = \mathrm{Eig}(T^*, \overline{\lambda}).$$

Beweis. Dies folgt aus dem Beweisschritt (c)⇒(a). □

Korollar 4.4.5. *Eine Matrix $A \in \mathrm{M}_n(\mathbb{C})$ ist genau dann normal, wenn es eine unitäre Matrix $S \in U(n)$ gibt, so dass $S^{-1}AS$ eine Diagonalmatrix ist.*

Beweis. Nach dem Satz ist A genau dann normal, wenn es eine ONB $e_1, \dots, e_n$ von $\mathbb{C}^n$ gibt, die aus Eigenvektoren besteht. Für eine beliebige Basis e_j sind diese Eigenschaften aber äquivalent dazu, dass für die Matrix S mit Spalten $e_1, \dots, e_n$ gilt: S ist unitär und $S^{-1}AS$ ist diagonal. □

Definition 4.4.6. Sei $A \in \mathrm{M}_n(\mathbb{C})$ selbstadjungiert. Dann heisst A **positiv definit**, wenn die hermitesche Form $b(v, w) = \langle Av, w \rangle$ positiv definit ist, wenn also gilt

$$v \neq 0 \quad \Rightarrow \quad \langle Av, v \rangle > 0.$$

Hierbei ist $\langle v, w \rangle = v^t \overline{w}$ das übliche Skalarprodukt auf $\mathbb{C}^n$.

Korollar 4.4.7. *Sei $A \in \mathrm{M}_n(\mathbb{C})$ positiv definit. Dann existiert ein $u \in \mathrm{U}(n)$, so dass*

$$uAu^{-1} = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix},$$

wobei $a_1, \dots, a_n > 0$ gilt.

Beweis. Nach dem Satz existiert ein u so dass uAu^{-1} Diagonalgestalt hat. Da A positiv definit ist, ist auch uAu^{-1} positiv definit und daher müssen alle Eigenwerte $a_1, \dots, a_n$ strikt positiv sein. □

* * *

4.5 Der Spektralsatz über den reellen Zahlen

In diesem Abschnitt ist $(V, \langle ., . \rangle)$ ein endlich-dimensionaler euklidischer Vektorraum, also $\mathbb{K} = \mathbb{R}$.

Lemma 4.5.1. *Ist $T : V \to V$ linear und selbstadjungiert, dann hat T einen Eigenvektor in V.*

Beweis. Nach Wahl einer ONB reicht es zu zeigen, dass eine Matrix $A \in \mathrm{M}_n(\mathbb{R})$ mit $A^t = A$ einen reellen Eigenwert zu einem Eigenvektor in $\mathbb{R}^n$ hat. Hierzu fasse A als Element von $\mathrm{M}_n(\mathbb{C})$ auf, dann ist A selbstadjungiert, und damit ist jeder Eigenwert von A reell, wir müssen zeigen, dass es einen Eigenvektor in $\mathbb{R}^n$ gibt. Sei also $\lambda \in \mathbb{R}$ ein Eigenwert, dann ist $\det(A - \lambda) = 0$, also ist der Kern von $A - \lambda$ in $\mathbb{R}^n$ nicht Null, also gibt es einen reellen Eigenvektor. □

Satz 4.5.2 (Spektralsatz über $\mathbb{R}$)**.** *Ist $\mathbb{K} = \mathbb{R}$ und $T : V \to V$ linear, so sind äquivalent*

(a) *T ist selbstadjungiert.*

(b) *Es gibt eine ONB aus Eigenvektoren.*

(c) *T ist diagonalisierbar und die Eigenräume stehen senkrecht aufeinander.*

Beweis. (a)⇒(c): Induktion nach der Dimension: Ist $\dim V = 1$, so ist die Behauptung klar. Sei also $\dim V \geq 2$. Nach dem Lemma existiert ein $\lambda \in \mathbb{R}$ mit $U = \mathrm{Eig}(T, \lambda) \neq 0$. Dann gilt $T(U^\perp) \subset U^\perp$, denn ist $v \in U^\perp$ und $u \in U$, so ist

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

Nach Induktionsvoraussetzung ist $U^\perp$ eine orthogonale Summe von Eigenräumen, also gilt dies auch für V.

(c)⇒(b) ist klar, da alle Eigenräume ONBs haben.

(b)⇒(a) Sei $e_1, \dots, e_n$ eine ONB mit $Te_j = \lambda_j e_j$. Dann ist

$$\begin{aligned} \left\langle Te_i, e_j \right\rangle &= \lambda_i \left\langle e_i, e_j \right\rangle = \lambda_i \delta_{i,j} = \lambda_j \delta_{i,j} \\ &= \lambda_j \left\langle e_i, e_j \right\rangle = \left\langle e_i, \lambda_j e_j \right\rangle = \left\langle e_i, Te_j \right\rangle. \end{aligned}$$ □

Korollar 4.5.3. *Ist $A \in \mathrm{M}_n(\mathbb{R})$ normal, dann sind äquivalent:*

(a) *Die Eigenwerte von A sind reell.*

(b) *A ist selbstadjungiert.*

Definition 4.5.4. Sei $\mathrm{O}(n) = \mathrm{M}_n(\mathbb{R}) \cap \mathrm{U}(n)$ die Menge aller reellen Matrizen A mit $AA^t = A^tA = I$. Die Menge $O(n)$ der orthogonalen Matrizen ist eine Untergruppe von $\mathrm{GL}_n(\mathbb{R})$.

Definition 4.5.5. Eine Matrix A heisst **symmetrisch**, wenn $A = A^t$ gilt.

Korollar 4.5.6 (Spektralsatz für symmetrische Matrizen)**.** *Eine Matrix $A \in \mathrm{M}_n(\mathbb{R})$ it genau dann symmetrisch, wenn es eine Matrix $S \in \mathrm{O}(n)$ gibt so dass $S^{-1}AS$ diagonal ist.*

Beweis. Da A reell ist, ist A genau dann symmetrisch, wenn die selbstadjungiert ist. Nach dem Satz ist A genau dann symmetrisch, wenn es eine ONB aus Eigenvektoren $(e_1, \dots, e_n)$ gibt. Die Matrix S mit den Spalten $e_1, \dots, e_n$ leistet das Gewünschte. Ist umgekehrt S gegeben, so bilden die Spalten von S eine ONB aus Eigenvektoren. □

Warnung. Eine symmetrische Matrix in $\mathrm{M}_n(\mathbb{R})$ ist immer über $\mathbb{R}$ diagonalisierbar. Für eine orthogonale Matrix gilt das nicht, wie das Beispiel $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ zeigt.

Definition 4.5.7. Eine symmetrische Matrix $A \in \mathrm{M}_n(\mathbb{R})$ heisst **positiv definit**, wenn die symmetrische Bilinearform $(v,w) \mapsto \langle Av, w\rangle = (Av)^t w = v^t Aw$ positiv definit ist. Wir schreiben in diesem Fall $A > 0$.

Lemma 4.5.8. *Sei $A \in \mathrm{M}_n(\mathbb{R})$ symmetrisch und sei $S \in \mathrm{GL}_n(\mathbb{R})$, dann gilt*

$$A > 0 \qquad \Leftrightarrow \qquad S^t AS > 0.$$

Proof. "$\Rightarrow$": Es gelte $A > 0$. Sei $0 \neq v \in \mathbb{R}^n$ und sei $w = Sv$. Dann gilt $0 < w^t Aw = (Sv)^t ASv = v^t \left(S^t AS\right) v$.

"$\Leftarrow$": Dieselbe Rechnung mit v und w vertauscht. □

Korollar 4.5.9. *Sei $A \in \mathrm{M}_n(\mathbb{R})$ positiv definit. Dann existiert ein $k \in \mathrm{O}(n)$, so dass*

$$kAk^{-1} = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix},$$

wobei $a_1, \dots, a_n > 0$ gilt.

Beweis. Nach Korollar 4.5.6 existiert ein $k \in \mathrm{O}(n)$ so dass kAk^{-1} Diagonalgestalt hat. Da A positiv definit ist, ist auch kAk^{-1} positiv definit und daher müssen alle Eigenwerte $a_1, \dots, a_n$ strikt positiv sein. □

Als Anwendung noch ein Psotitivitätskriterium, das in der Analysis 2 gebraucht wird.

Definition 4.5.10. Eine selbstadjungierte Matrix $A \in \mathrm{M}_n(\mathbb{K})$ heisst **positiv semidefinitdefinit**, falls die Hermitesche Form b_A positiv semidefinit ist, wenn also für jeden Vektor $v \in \mathbb{K}^n \setminus \{0\}$ gilt

$$b_A(v,v) = v^t A\overline{v} \geq 0.$$

Satz 4.5.11. *Für eine selbstadjungierte Matrix $A \in \mathrm{M}_n(\mathbb{R})$ sind äquivalent:*

(a) *A ist positiv definit,*

(b) *Es gilt* $\det(A_k) > 0$ *für jedes $k = 1, \dots, n$, wobei A_k die linke obere $k \times k$ Teilmatrix ist.*

Dasselbe gilt, wenn man "positiv definit" durch "positiv semidefinit" und > 0 durch ≥ 0 ersetzt.

Man nennt die Zahlen $\det(A_k)$ auch die **Hauptminoren** von A.

Beweis. (a)⇒(b): Jede positiv definite Matrix A hat positive Determinante, denn ist

$$S^{-1}AS = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix},$$

mit $S \in O(n)$, dann gilt

$$\alpha_j = e_j^t S^{-1} A S e_j = e_j^t S^t A S e_j = (Se_j)\ ASe_j = b_A(Se_j, Se_j) > 0.$$

Daher ist $\det(A) = \alpha_1 \cdots \alpha_n > 0$.

Die Matrix A_k beschreibt die Einschränkung der zu A gehoerenden Bilinearform auf den k-dimensionalen Unterraum

$$\mathrm{Spann}(e_1, \dots, e_k).$$

Also ist auch A_k positiv und $\det(A_k) > 0$.

(b)⇒(a): Induktion nach n. Der Fall $n = 1$ ist klar. Nach Induktionsannahme koennen wir A_{n-1} als positiv annehmen, also gibt es ein $S_1 \in O(n-1)$ mit

$$S_1^t A_{n-1} S_1 = \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_{n-1} \end{pmatrix}$$

mit $\alpha_j > 0$. Setze dann

$$S_2 = S_1 \begin{pmatrix} \sqrt{\alpha_1} & & \\ & \ddots & \\ & & \sqrt{\alpha_{n-1}} \end{pmatrix}.$$

Dann gilt $S_2^t A S_2 = I_{n-1}$, die $(n-1) \times (n-1)$ Einheitsmatrix. Sei dann

$$S_3 = \left(\begin{array}{ccc|c} & & & 0 \\ & S_2 & & \vdots \\ & & & 0 \\ \hline 0 & \dots & 0 & 1 \end{array}\right).$$

Dann gilt

$$B := S_3^t A S_3 = \left(\begin{array}{ccc|c} & & & \beta_1 \\ & I_{n-1} & & \vdots \\ & & & \beta_{n-1} \\ \hline \beta_1 & \dots & \beta_{n-1} & \beta_n \end{array}\right).$$

Mit der Matrix

$$S = \left(\begin{array}{ccc|c} & & & -\beta_1 \\ & I_{n-1} & & \vdots \\ & & & -\beta_{n-1} \\ \hline 0 & \dots & 0 & 1 \end{array}\right)$$

ergibt sich dann

$$D := S^t B S = \left(\begin{array}{ccc|c} & & & 0 \\ & I_{n-1} & & \vdots \\ & & & 0 \\ \hline 0 & \dots & 0 & \alpha \end{array}\right)$$

mit $\alpha = \beta_n - \beta_1^2 - \cdots - \beta_{n-1}^2$. Es bleibt zu zeigen, dass $\alpha > 0$ ist. Es gilt

$$\alpha = \det D = \det(S)^2 \det B = \det(S)^2 \det(S_3)^2 \; \det A > 0. \qquad \square$$