Mathematik für Physiker 2

Thomas Markwig
Fachbereich Mathematik
Eberhard Karls Universität Tübingen

Vorlesungsskript

Sommersemester 2016

Inhaltsverzeichnis

Einleitung

Die vorliegende Ausarbeitung zur Vorlesung Mathematik für Physiker 2 im Sommersemester 2016 wird im wesentlichen wiedergeben, was während der Vorlesung an die Tafel geschrieben wird. Einige wenige Abschnitte werden ausführlicher sein. Die Ausarbeitung ersetzt somit in keiner Weise ein Lehrbuch.

KAPITEL I

Vektorräume und lineare Abbildungen

Im folgenden wollen wir die Theorie der Vektorräume und der linearen Abbildungen studieren, unter anderem mit dem Ziel, die Lösungsmengen von linearen Gleichungssystemen zu verstehen und berechnen zu können. Der Analysis lagen die Körper der reellen und komplexen Zahlen zugrunde. Wesentlichster Baustein neben der Addition und Multiplikation waren der Absolutbetrag mit Werten in $\mathbb{R}$ und die Ordnungsrelation auf $\mathbb{R}$, die $\mathbb{R}$ zu einem vollständigen angeordneten Körper machte. Für die lineare Algebra spielen der Absolutbetrag und die Ordnungsrelation keine Rolle mehr. Wir kommen ohne ε's und δ's und komplizierte Abschätzungen aus. Deshalb können wir unser Arsenal an Grundstrukturen, mit denen wir arbeiten wollen, auch erweitern.

$\boxed{K \text{ wird im folgenden einen } \textit{beliebigen Körper} \text{ bezeichnen,}}$

etwa $\mathbb{Q}$, $\mathbb{R}$ oder $\mathbb{C}$ oder auch einen endlichen Körper wie etwa $\mathbb{F}_2$ in Beispiel B3.2. Manchmal können wir sogar auf die Division verzichten und statt einem Körper eine Struktur wie die ganzen Zahlen $\mathbb{Z}$ zugrunde legen.

§ 1 Rechnen mit Matrizen

Definition 1.1 (Matrizen und der K^n)
Es seien $m, n \geq 1$ zwei positive ganze Zahlen.

a. Eine $m \times n$-*Matrix* über K ist ein rechteckiges Schema A mit Einträgen aus K der Form
$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}.$$
Wenn keine Unklarheiten zu befürchten sind, schreiben wir verkürzt auch
$$A = (a_{ij})_{i=1,\ldots,m;j=1,\ldots,n} = (a_{ij}).$$

b. Die Menge aller $m \times n$-Matrizen über K wird mit
$$\mathrm{Mat}(m \times n, K)$$
bezeichnet, und falls $m = n$, dann auch kurz mit $\mathrm{Mat}_n(K) = \mathrm{Mat}(n, K)$ und man spricht von *quadratischen Matrizen.*

c. Ist $A = (a_{ij})$ eine $m \times n$-Matrix, dann bezeichnen wir

$$a_i := (a_{i1}, \ldots, a_{in})$$

als den i-ten *Zeilenvektor* von A und

$$a^j := \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

als den j-ten *Spaltenvektor* von A.

d. Ist $A = (a_{ij}) \in \mathrm{Mat}(m \times n, K)$, so heißt die $n \times m$-Matrix

$$A^t := \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{m1} \\ a_{12} & a_{22} & \ldots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \ldots & a_{mn} \end{pmatrix},$$

d. h. für $A^t = (a'_{ij})$ gilt $a'_{ij} = a_{ji}$, die *Transponierte* von A.

e. Schließlich definieren wir

$$K^n := \mathrm{Mat}(n \times 1, K) = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \;\middle|\; x_1, \ldots, x_n \in K \right\}.$$

Die Elemente von K^n heißen *Vektoren* oder *Punkte* im K^n. x_i heißt die i-te *Komponente* des Vektors x.

Definition 1.2 (Operationen mit Matrizen)

a. Es seien $A = (a_{ij}), B = (b_{ij}) \in \mathrm{Mat}(m \times n, K)$ und $\lambda \in K$. Dann definiert man

$$A + B := (a_{ij} + b_{ij}) = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \ldots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \ldots & a_{2n} + b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \ldots & a_{mn} + b_{mn} \end{pmatrix},$$

sowie

$$\lambda \cdot A := (\lambda a_{ij}) = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \ldots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \ldots & \lambda a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \ldots & \lambda a_{mn} \end{pmatrix}.$$

b. Sind $A = (a_{ij}) \in \mathrm{Mat}(m \times n, K)$ und $B = (b_{jk}) \in \mathrm{Mat}(n \times p, K)$ zwei Matrizen, wobei A genauso viele Spalten wie B Zeilen hat. Dann definieren wir das *Matrixprodukt* durch

$$A \circ B := C, \quad \text{mit } C = (c_{ik}) \in \mathrm{Mat}(m \times p, K) \text{ und } c_{ik} := \sum_{j=1}^{n} a_{ij} b_{jk}.$$

Beispiel 1.3
Folgende Matrizen $A, B \in \mathrm{Mat}(2 \times 3, \mathbb{K})$ und $C \in \mathrm{Mat}(3 \times 2, \mathbb{K})$ seien gegeben:

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 1 & 2 \end{pmatrix}, B = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 4 & 5 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 1 \end{pmatrix}.$$

Dann gilt:

$$A + B = \begin{pmatrix} 4 & 2 & 3 \\ 3 & 5 & 7 \end{pmatrix}, \ 3 \cdot A = \begin{pmatrix} 3 & 0 & 6 \\ 9 & 3 & 6 \end{pmatrix} \quad \text{und} \quad A \circ C = \begin{pmatrix} 7 & 2 \\ 11 & 3 \end{pmatrix}.$$

Bemerkung 1.4

a. Die in Definition 1.2 a. definierte Addition zweier Matrizen definiert auf $\mathrm{Mat}(m \times n, K)$ offensichtlich eine zweistellige Operation, bezüglich derer $\mathrm{Mat}(m \times n, K)$ eine abelsche Gruppe $(\mathrm{Mat}(m \times n, K), +)$ wird, wie man leicht nachprüft.

b. Wir werden meist kurz λA bzw. λx schreiben, statt $\lambda \cdot A$ bzw. $\lambda \cdot x$, wenn $\lambda \in K$, $A \in \mathrm{Mat}(m \times n, K)$ und $x \in K^n$.

c. Wir schreiben statt $A \circ B$ häufig kurz AB, insbesondere auch Ax statt $A \circ x$.

d. Spaltenvektoren nehmen im Skript sehr viel Raum ein. Um platzsparend arbeiten zu können, werden wir deshalb statt den Spaltenvektor $x \in K^n$ als

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

anzugeben, meist den *transponierten* Zeilenvektor

$$x = (x_1 \ \ldots \ x_n)^t$$

betrachten, und um Mißverständnissen vorzubeugen, fügen wir zudem meist Kommata als Trennsymbole ein

$$x = (x_1, \ldots, x_n)^t.$$

e. Man beachte, daß das Produkt nur dann definiert ist, wenn A soviele Spalten wie B Zeilen hat. Das Produkt $A \circ B$ hat dann soviele Zeilen wie A und soviele Spalten wie B.

Jede Matrix definiert wie folgt eine Abbildung.

Definition 1.5 (Die Abbildung f_A)
Ist $A \in \mathrm{Mat}(m \times n, K)$, so definieren wir

$$f_A : K^n \to K^m : x \mapsto Ax = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j} x_j \\ \vdots \\ \sum_{j=1}^n a_{mj} x_j \end{pmatrix}.$$

f_A heißt die *zu* A *assoziierte* oder *zu* A *gehörige Abbildung*.

Beispiel 1.6
Die Matrix

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

definiert die Abbildung

$$f_A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 : (x_1, x_2)^t \mapsto (x_1 + 2x_2, 3x_2)^t.$$

Bemerkung 1.7 (Einheitsvektoren)
Um den Zusammenhang zwischen A und f_A besser zu verstehen, betrachten wir für $i = 1, \ldots, n$ den i-ten *Einheitsvektor* $e_i = (\delta_{1i}, \cdots, \delta_{ni})^t \in K^n$, wobei

$$\delta_{ji} := \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

das *Kronecker Symbol* ist, d. h.

$$e_i := \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix},$$

wobei die Eins in der i-ten Komponente steht.

Es ist dann

$$f_A(e_i) = \begin{pmatrix} \sum_{j=1}^n a_{1j}\delta_{ji} \\ \vdots \\ \sum_{j=1}^n a_{mj}\delta_{ji} \end{pmatrix} = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix} = a^i,$$

d. h. die i-te Spalte von A ist das Bild des i-ten Einheitsvektors unter f_A.

Hieraus folgt insbesondere, daß A durch f_A eindeutig bestimmt ist.

Lemma 1.8 (Einfache Rechenregeln für Matrizen)
Für $x, y \in K^n$, $A, B \in \mathrm{Mat}(m \times n, K)$, $C \in \mathrm{Mat}(n \times p, K)$ *und* $\lambda \in K$ *gelten:*

a. $A(x + y) = Ax + Ay$ *und* $A(\lambda x) = \lambda Ax$,

b. $\lambda \cdot (A \circ C) = (\lambda \cdot A) \circ C = A \circ (\lambda \cdot C)$,

c. $f_{A+B} = f_A + f_B$, *und*

d. $f_{\lambda A} = \lambda f_A$.

Beweis: Der Beweis ist eine Übungsaufgabe. □

Wir wollen jetzt sehen, wie sich die Multiplikation von Matrizen mit den zugehörigen Abbildungen verträgt.

Satz 1.9 (Matrixmultiplikation und Komposition)
Für $A \in \mathrm{Mat}(m \times n, K)$ *und* $B \in \mathrm{Mat}(n \times p, K)$ *gilt:*

$$f_{A \circ B} = f_A \circ f_B.$$

Beweis: Da Definitionsbereich und Wertebereich von beiden Abbildungen übereinstimmen, reicht es zu zeigen:

$$(f_{A \circ B})(x) = (f_A \circ f_B)(x), \quad \text{für alle } x \in K^p.$$

Seien $A = (a_{ij})$ und $B = (b_{jk})$, und sei $x = (x_1, \ldots, x_p)^t \in K^p$ gegeben.

$$(f_{A \circ B})(x) = (A \circ B)x = \begin{pmatrix} \sum_{j=1}^n a_{1j} b_{j1} & \ldots & \sum_{j=1}^n a_{1j} b_{jp} \\ \vdots & & \vdots \\ \sum_{j=1}^n a_{mj} b_{j1} & \ldots & \sum_{j=1}^n a_{mj} b_{jp} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{k=1}^p \sum_{j=1}^n a_{1j} b_{jk} x_k \\ \vdots \\ \sum_{k=1}^p \sum_{j=1}^n a_{mj} b_{jk} x_k \end{pmatrix}.$$

Ferner gilt:

$$(f_A \circ f_B)(x) = f_A(Bx) = A(Bx) = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} \sum_{k=1}^p b_{1k} x_k \\ \vdots \\ \sum_{k=1}^p b_{nk} x_k \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{j=1}^n a_{1j} \sum_{k=1}^p b_{jk} x_k \\ \vdots \\ \sum_{j=1}^n a_{mj} \sum_{k=1}^p b_{jk} x_k \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n \sum_{k=1}^p a_{1j} b_{jk} x_k \\ \vdots \\ \sum_{j=1}^n \sum_{k=1}^p a_{mj} b_{jk} x_k \end{pmatrix}.$$

Beide Ausdrücke stimmen (bis auf die Reihenfolge der Summation) überein, was zu zeigen war. □

Korollar 1.10 (Die Matrixmultiplikation ist assoziativ.)
Für $A \in \mathrm{Mat}(m \times n, K)$, $B \in \mathrm{Mat}(n \times p, K)$ *und* $C \in \mathrm{Mat}(p \times q, K)$ *gilt*

$$A \circ (B \circ C) = (A \circ B) \circ C.$$

Beweis: Dies folgt aus Satz 1.9, da die Komposition von Abbildungen assoziativ ist und da eine Matrix A durch die Abbildung f_A eindeutig bestimmt ist. □

Man kann die Aussage des Korollars natürlich auch direkt nachweisen, was auf die gleiche Rechnung wie in 1.9 führt - statt des einen Vektors x hat man die q Spaltenvektoren von C zu multiplizieren, was etwas mehr Schreibarbeit bedeutet.

Lemma 1.11 (Distributivgesetze)
Sind $A, B \in \mathrm{Mat}(m \times n, K)$ *und* $C, D \in \mathrm{Mat}(n \times p, K)$*, so gelten die Distributivgesetze:*

$$A \circ (C + D) = A \circ C + A \circ D,$$

sowie

$$(A + B) \circ C = A \circ C + B \circ C.$$

Beweis: Die Aussage kann wie Korollar 1.10 aus Lemma 1.8 und Satz 1.9 abgeleitet werden und sei dem Leser als Übung anempfohlen. □

Definition 1.12 (Invertierbare Matrizen)
Eine Matrix $A \in \mathrm{Mat}_n(K)$ heißt *invertierbar*, falls es eine Matrix $A^{-1} \in \mathrm{Mat}_n(K)$ gibt, so daß

$$A \circ A^{-1} = A^{-1} \circ A = \mathbb{1}_n,$$

wobei die Matrix $\mathbb{1}_n = (\delta_{ij})_{i,j=1,\ldots,n} \in \mathrm{Mat}_n(K)$ die *Einheitsmatrix* ist, die auf der Diagonalen Einsen und außerhalb der Diagonalen Nullen als Einträge hat. Eine Matrix mit der Eigenschaft von A^{-1} nennt man eine *Inverse* zu A.

Satz 1.13 (Die allgemeine lineare Gruppe $\mathrm{Gl}_n(K)$)
Die Menge der invertierbaren $n \times n$-Matrizen

$$\mathrm{Gl}_n(K) = \{A \in \mathrm{Mat}_n(K) \mid A \text{ ist invertierbar}\}$$

ist eine Gruppe mit neutralem Element $\mathbb{1}_n$, die für $n > 1$ nicht kommutativ ist.

Insbesondere ist die Inverse zu A eindeutig bestimmt und es gelten für $A, B \in \mathrm{Gl}_n(K)$

$$(AB)^{-1} = B^{-1}A^{-1} \quad \textit{und} \quad (A^{-1})^{-1} = A.$$

Beweis: Der Beweis ist eine Übungsaufgabe. □

Aufgaben

Aufgabe 1.14
Zeige, daß die Matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Mat}_2(K)$$

genau dann invertierbar ist, wenn $ad - bc \neq 0$ ist. Die Inverse ist dann

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Aufgabe 1.15
Es sei $A \in \mathrm{Mat}(m \times n, K)$ und $B \in \mathrm{Mat}(n \times p, K)$. Zeige, $(AB)^t = B^t A^t$.

Aufgabe 1.16 (Nilpotente Matrizen)
Es sei $N = (n_{ij}) \in \mathrm{Mat}_n(K)$ eine Matrix, für die die Einträge auf der oberen Nebendiagonale alle 1 sind und für die alle anderen Einträge 0 sind, d.h. $n_{ij} = \delta_{j-i,1}$.

Zeige für $k = 1, \ldots, n$, daß die Einträge der Matrix $N^k = \left(n_{ij}^{(k)}\right)$ auf der k-ten oberen Nebendiagonale alle 1 und alle anderen Einträge 0 sind, d.h. $n_{ij}^{(k)} = \delta_{j-i,k}$. Insbesondere ist $N^n = 0$ und $N^k \neq 0$ für $k < n$.

§ 2 Vektorräume und lineare Abbildungen

A) Vektorräume

Definition 2.1 (Vektorräume)
Ein K-*Vektorraum* (oder Vektorraum über K) besteht aus einer nicht-leeren Menge V sowie einer zweistelligen Operation

$$+ : V \times V \to V : (x, y) \mapsto x + y,$$

die *Vektoraddition* genannt wird, und einer zweistelligen Operation

$$\cdot : K \times V \to V : (\lambda, x) \mapsto \lambda \cdot x = \lambda x,$$

die *Skalarmultiplikation* genannt wird, so daß die folgenden Gesetze gelten:

a. $(V, +)$ ist eine abelsche Gruppe,

b. für $\lambda, \mu \in K$ und $x, y \in V$ gelten:
   (i) $(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$, ("*verallgemeinertes Distributivgesetz*")
   (ii) $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$, und ("*verallgemeinertes Distributivgesetz*")
   (iii) $(\lambda \cdot \mu) \cdot x = \lambda \cdot (\mu \cdot x)$. ("*verallgemeinertes Assoziativgesetz*")
   (iv) $1 \cdot x = x$.

Die Elemente aus V nennt man *Vektoren* und die aus K *Skalare*. Der *Nullvektor*, d. h. das neutrale Element aus V bezüglich der Addition, wird mit 0 bzw. mit 0_V bezeichnet und das neutrale Element von $(K, +)$ ebenfalls mit 0 bzw. mit 0_K.

Beispiel 2.2

a. Der *Nullraum* $V = \{0\}$ mit $\lambda \cdot 0 = 0$ für alle $\lambda \in K$ ist für jeden Körper K ein K-Vektorraum. Man bezeichnet den Nullraum auch mit K^0.

b. Der Körper K selbst mit der Körperaddition als Vektoraddition und der Körpermultiplikation als Skalarmultiplikation ist ein K-Vektorraum.

c. $\mathbb{R}$ ist ein $\mathbb{Q}$-Vektorraum und $\mathbb{C}$ ist ein $\mathbb{R}$-Vektorraum, jeweils mit der üblichen Addition und Multiplikation.

d. Die Menge $\mathrm{Mat}(m \times n, K)$ der $m \times n$-Matrizen über K mittels der in Definition 1.2 definierten Addition und Skalarmultiplikation ist ein K-Vektorraum mit der *Nullmatrix*

$$0 := \begin{pmatrix} 0 & \ldots & 0 \\ \vdots & & \vdots \\ 0 & \ldots & 0 \end{pmatrix}$$

als Nullvektor.

e. Damit ist insbesondere K^n mit der komponentenweisen Addition und Skalarmultiplikation ein K-Vektorraum mit $0_{K^n} = (0, \ldots, 0)^t$.

Speziell sind $\mathbb{R}^n$, $\mathbb{C}^n$ und $\mathbb{F}_2^n$ Vektorräume über $\mathbb{R}$, $\mathbb{C}$ bzw. $\mathbb{F}_2$ (für die Definition des Körpers $\mathbb{F}_2$ verweisen wir auf Beispiel B3.2.

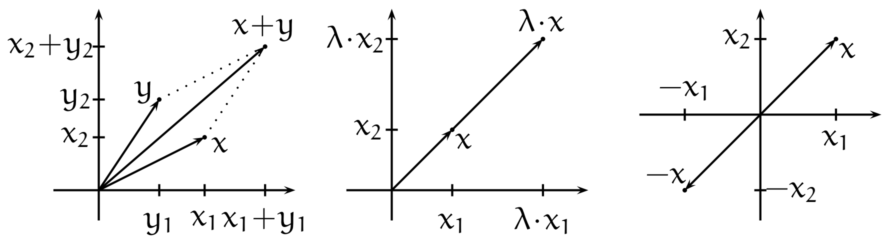


Abbildung 1. Addition und Skalarmultiplikation in $\mathbb{R}^2$

f. Ist M eine Menge und V ein K-Vektorraum, so wird die Menge

$$V^M = \{f : M \longrightarrow V \mid f \text{ ist eine Abbildung}\}$$

durch die Operationen

$$+ : V^M \times V^M \to V^M : (f, g) \mapsto \big(f + g : M \to V : x \mapsto f(x) + g(x)\big)$$

und

$$\cdot : K \times V^M \to V^M : (\lambda, f) \mapsto \big(\lambda \cdot f : M \to V : x \mapsto \lambda \cdot f(x)\big)$$

zu einem K-Vektorraum, wie man leicht nachrechnet.
Ist z.B. $M = \mathbb{N}$ und $K = V = \mathbb{R}$, so ist

$$\mathbb{R}^{\mathbb{N}} = \{a : \mathbb{N} \longrightarrow \mathbb{R} \mid a \text{ ist eine Abbildung}\} = \big\{(a_n)_{n\in\mathbb{N}} \;\big|\; a_n \in \mathbb{R}\big\}$$

der Vektorraum der Folgen in $\mathbb{R}$. Unsere Definitionen sorgen dafür, daß Folgen komponentenweise addiert und mit Skalaren multipliziert werden.

g. Ist t eine Veränderliche und sind $a_0, \ldots, a_n \in K$, so nennen wir einen Ausdruck der Form

$$\sum_{k=0}^{n} a_k \cdot t^k = a_n \cdot t^n + a_{n-1} \cdot t^{n-1} + \ldots + a_1 \cdot t + a_0$$

ein *Polynom* in der Veränderlichen t mit Koeffizienten in K. Ist $a_n \neq 0$, so heißt

$$\deg\left(\sum_{k=0}^{n} a_k \cdot t^k\right) := n$$

der *Grad* des Polynoms, und wir setzen zudem $\deg(0) := -\infty$. Die Menge

$$K[t] := \left\{ \sum_{k=0}^{n} a_k \cdot t^k \;\middle|\; n \in \mathbb{N}, a_k \in K \right\}$$

der Polynome mit Koeffizienten in K wird durch die Addition

$$\sum_{k=0}^{m} a_k \cdot t^k + \sum_{k=0}^{n} b_k \cdot t^k := \sum_{k=0}^{\max\{m,n\}} (a_k + b_k) \cdot t^k,$$

wobei $a_k = 0$ für $k > m$ und $b_k = 0$ für $k > n$ gelten soll, und durch die Skalarmultiplikation

$$\lambda \cdot \sum_{k=0}^{n} a_k \cdot t^k := \sum_{k=0}^{n} (\lambda \cdot a_k) \cdot t^k$$

zu einem K-Vektorraum, wie man leicht nachprüft.

h. Da man für $M = \{1, \ldots, n\}$ eine Abbildung $f : M \to K$ in eindeutiger Weise durch das Tupel der Bildelemente $\big(f(1), \ldots, f(n)\big)$ beschreiben kann, sieht man leicht, daß die Zuordnung

$$K^M \to K^n : f \mapsto \big(f(1), \ldots, f(n)\big)^t$$

in diesem Falle eine Bijektion ist. Man prüft überdies leicht nach, daß diese Abbildung ein Vektorraumhomomorphismus im Sinne von Definition 2.19 ist. K^M und K^n sind dann also isomorph.

Lemma 2.3 (Einfache Rechenregeln für Vektoren)
In einem K-Vektorraum gelten folgende Rechenregeln:

a. $0_K \cdot x = 0_V$ *und* $\lambda \cdot 0_V = 0_V$ *für alle* $x \in V$, $\lambda \in K$.

b. *Für* $\lambda \in K$ *und* $x \in V$ *gilt:*

$$\lambda \cdot x = 0_V \;\Rightarrow\; \lambda = 0 \quad \textit{oder} \quad x = 0.$$

c. $(-1) \cdot x = -x$ *für alle* $x \in V$.

Beweis: Es seien $x \in V$ und $\lambda \in K$ gegeben.

a. Es gilt:

$$0_V + 0_K \cdot x = 0_K \cdot x = (0_K + 0_K) \cdot x = 0_K \cdot x + 0_K \cdot x,$$

also $0_V = 0_K \cdot x$, wie aus den Kürzungsregeln für $(V, +)$ folgt. Analog gilt:

$$0_V + \lambda \cdot 0_V = \lambda \cdot 0_V = \lambda \cdot (0_V + 0_V) = \lambda \cdot 0_V + \lambda \cdot 0_V,$$

und damit $0_V = \lambda \cdot 0_V$.

b. Ist $\lambda \in K$ mit $\lambda \neq 0$, dann gibt es ein Inverses $\lambda^{-1} \in K$. Aus $\lambda \cdot x = 0$ folgt dann aber wegen a. und den Vektorraumaxiomen

$$0_V = \lambda^{-1} \cdot 0_V = \lambda^{-1} \cdot (\lambda \cdot x) = \big(\lambda^{-1} \cdot \lambda\big) \cdot x = 1 \cdot x = x.$$

c. Für $x \in K$ gilt:

$$x + (-1) \cdot x = 1 \cdot x + (-1) \cdot x = \big(1 + (-1)\big) \cdot x = 0_K \cdot x = 0_V.$$

Also ist $(-1) \cdot x$ das (eindeutig bestimmte) additive Inverse zu x.

□

B) Unterräume

Definition 2.4 (Unterräume)
Es sei V ein Vektorraum über K. Eine nicht-leere Teilmenge $U \subseteq V$ von V heißt *Unterraum*, wenn für alle $\lambda \in K$ und $x, y \in U$ gilt

$$\lambda \cdot x \in U \quad \text{und} \quad x + y \in U. \tag{1}$$

Man sagt, U sei *abgeschlossen* bezüglich der Addition und der Skalarmultiplikation. Wir schreiben $U \leq V$, um auszudrücken, daß U ein Unterraum von V ist.

Proposition 2.5 (Unterräume sind Vektorräume.)
Jeder Unterraum eines K-Vektorraums ist selbst ein K-Vektorraum.

Beweis: Als Unterraum eines K-Vektorraums ist U eine nicht-leere Menge. Für die Addition und Skalarmultiplikation, die U von V erbt, gilt nach Voraussetzung

$$U \times U \longrightarrow U : (x, y) \mapsto x + y$$

und

$$K \times U \longrightarrow U : (\lambda, x) \mapsto \lambda \cdot x,$$

sie sind also zweistellige Operationen mit Werten in U!

Da U nicht leer ist, gibt es ein $y \in U$ und mithin folgt aus (1)

$$0_V = 0_K \cdot y \in U.$$

Damit besitzt U offenbar ein neutrales Element der Addition, da für alle $x \in U$ auch

$$0_V + x = x$$

gilt. Ist $x \in U$ beliebig, so ist zudem

$$-x = (-1) \cdot x \in U,$$

so daß U wegen

$$-x + x = 0_V = 0_U$$

auch das Inverse von x enthält. Da das Assoziativgesetz für alle Elemente in V gilt und die Elemente aus U auch in V sind, gilt es automatisch für alle Elemente in U. Wir haben also gezeigt, daß U eine abelsche Gruppe bezüglich $+$ ist.

Die verbleibenden Axiome aus Teil b. in Definition 2.1 vererben sich analog zum Assoziativgesetz unmittelbar von V auf U. Damit ist U also ein K-Vektorraum. □

Beispiel 2.6

a. Ist V ein K-Vektorraum, so ist $\{0_V\}$ stets ein Unterraum von V. Ferner ist V selbst ein Unterraum. Man nennt diese beiden auch die *trivialen Unterräume*.

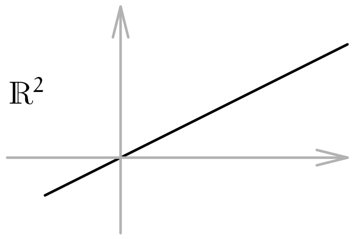
b. Eine Gerade G durch den Ursprung in der Ebene $\mathbb{R}^2$ mit Steigung m genügt der Gleichung $y = m \cdot x$ und ist deshalb

$$G = \left\{(x, mx)^t \;\middle|\; x \in \mathbb{R}\right\}.$$

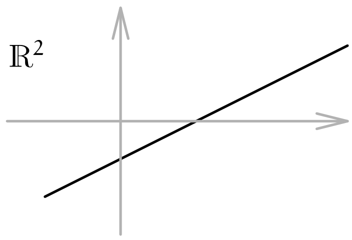
Für $v = (x, mx)^t, w = (x', mx')^t \in G$ und $\lambda \in \mathbb{R}$ gilt dann

$$v + w = \big(x + x', m \cdot (x + x')\big)^t,\ \lambda \cdot v = (\lambda x, m\lambda x)^t \in G.$$

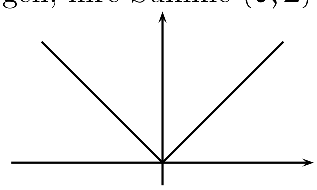
Mithin ist G ein Unterraum von $\mathbb{R}^2$.



c. Eine Gerade in $\mathbb{R}^2$, die nicht durch den Ursprung geht, kann kein Unterraum von $\mathbb{R}^2$ sein, da ein Unterraum den Nullvektor enthalten muß.



d. Der Graph der Betragsfunktion ist *kein* Unterraum von $\mathbb{R}^2$, da $(-1, 1)^t$ und $(1, 1)^t$ auf dem Graphen liegen, ihre Summe $(0, 2)^t$ aber nicht.



e. Die Menge

$$U := \{(a_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid (a_n)_{n\in\mathbb{N}} \text{ ist konvergent}\}$$

der *konvergenten Folgen* in $\mathbb{R}$ ist ein Unterraum des $\mathbb{R}$-Vektorraums $\mathbb{R}^{\mathbb{N}}$ aller Folgen. Dies folgt aus den Grenzwertsätzen für Folgen.

f. Ist I ein Intervall in $\mathbb{R}$, so sind die Menge $\mathcal{C}(I, \mathbb{R})$ aller auf I stetigen Abbildungen sowie die Menge $\mathcal{C}^k(I, \mathbb{R})$ aller auf I k-fach stetig differenzierbaren Abbildungen Unterräume des Vektorraums $\mathbb{R}^I$ aller Abbildungen von I nach $\mathbb{R}$.

Solche Funktionenräume spielen in der Analysis eine große Rolle. Sie sind für kein n isomorph zu $\mathbb{R}^n$, und sie sind ein gutes Beispiel für den Wert der abstrakten Theorie der Vektorräume.

g. Ist $n \in \mathbb{N}$ fest vorgegeben, so bilden die Polynome vom Grad höchstens n

$$P_n := \left\{ \sum_{k=0}^{n} a_k \cdot t^k \;\middle|\; a_k \in K \right\}$$

einen Unterraum des Vektorraums der Polynome $K[t]$.

Lemma 2.7 (Durchschnitt von Unterräumen)
Der Durchschnitt beliebig vieler Unterräume ist wieder ein Unterraum.

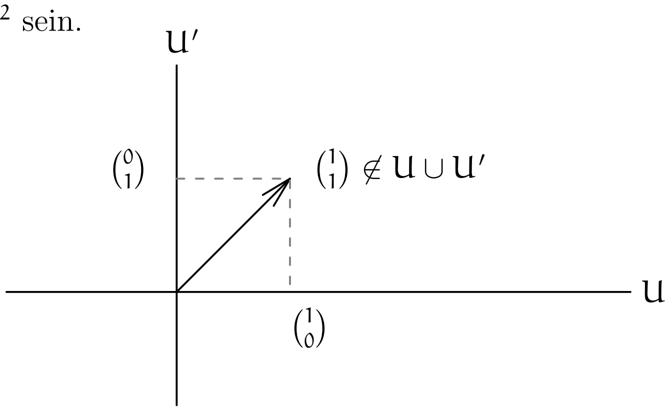
Beweis: Es seien U_i für $i \in I$ Unterräume eines K-Vektorraums V. Da $0_V \in U_i$ für alle $i \in I$, ist $U := \bigcap_{i \in I} U_i$ nicht die leere Menge. Es bleibt also zu zeigen, daß für $x, y \in U$ und $\lambda \in K$ gilt:

$$x + y \in U \quad \text{und} \quad \lambda x \in U.$$

Für ein beliebiges $i \in I$ gilt, da U_i ein Unterraum von V ist und da $x, y \in U \subseteq U_i$, daß $x + y \in U_i$ und $\lambda x \in U_i$. Also liegen die Vektoren im Durchschnitt U. $\square$

Bemerkung 2.8
Die Vereinigung von zwei Unterräumen ist i. a. kein Unterraum mehr!

Sei etwa U die x-Achse und U' die y-Achse im $\mathbb{R}^2$. Beides sind offenbar Unterräume von $\mathbb{R}^2$. Dann liegt $(1,1)^t = e_1 + e_2$ nicht in $U \cup U'$, und mithin kann $U \cup U'$ kein Unterraum von $\mathbb{R}^2$ sein.



Definition 2.9 (Linearkombination und lineare Hülle)
Es sei V ein K-Vektorraum.

a. Wir nennen $x \in V$ eine *Linearkombination* von $x_1, \ldots, x_r \in V$, falls es $\lambda_1, \ldots, \lambda_r \in K$ gibt mit

$$x = \lambda_1 x_1 + \ldots + \lambda_r x_r.$$

Ist eines der λ_i ungleich Null, so nennen wir die Linearkombination *nicht-trivial.*

b. Ist $M \subseteq V$, so nennen wir den Durchschnitt

$$\mathrm{Lin}\,(M) := \langle M \rangle := \bigcap_{M \subseteq U \leq V} U$$

aller Unterräume von V, die M enthalten, die *lineare Hülle* oder das *Erzeugnis* von M.

Bemerkung 2.10

a. Man beachte, daß die lineare Hülle von M wegen Lemma 2.7 ein Unterraum von V ist. Aufgrund der Definition ist es *der kleinste* Unterraum, der M enthält.

b. Ist $M = \emptyset$, so ist $\mathrm{Lin}\,(M) = \mathrm{Lin}\,(\emptyset) = \{0\}$.

c. Eine Linearkombination ist immer eine *endliche* Summe von Vielfachen von Vektoren aus V. In der linearen Algebra wird es *nie* unendliche Summen geben.

d. Mit Induktion nach der Anzahl der Summanden folgt aus (1) unmittelbar, daß ein Unterraum U abgeschlossen bezüglich endlicher Linearkombinationen von Vektoren aus U ist.

Proposition 2.11 (Lineare Hülle = Menge der Linearkombinationen)
Ist V ein K-Vektorraum und $\emptyset \neq M \subseteq V$, so ist die lineare Hülle von M

$$\mathrm{Lin}\,(M) = \{\lambda_1 \cdot x_1 + \ldots + \lambda_r \cdot x_r \mid r \in \mathbb{N}, x_1, \ldots, x_r \in M, \lambda_1, \ldots, \lambda_r \in K\} \leq V$$

die Menge aller Linearkombinationen von Elementen in M.

Beweis: Wir setzen $U := \{\lambda_1 \cdot x_1 + \ldots + \lambda_r \cdot x_r \mid r \in \mathbb{N}, x_1, \ldots, x_r \in M, \lambda_1, \ldots, \lambda_r \in K\}$.

Als Unterraum, der M enthält, enthält $\mathrm{Lin}\,(M)$ auch jede endliche Linearkombination von Elementen in M, also auch U.

Wir wollen nun zeigen, daß U ein Unterraum von V ist, der M enthält, da er aufgrund der Definition des Erzeugnisses dann auch das Erzeugnis von M enthält. Dazu beachten wir zunächst, daß für $x \in M$ auch $x = 1 \cdot x \in U$ gilt. Also ist $M \subseteq U$ und somit ist U auch nicht leer. Seien nun $x = \sum_{i=1}^{r} \lambda_i \cdot x_i$ und $y = \sum_{j=1}^{s} \mu_j \cdot y_j$ mit $x_i, y_j \in M$ und $\lambda_i, \mu_i \in K$ sowie $\lambda \in K$ gegeben. Dann ist

$$x + y = \lambda_1 \cdot x_1 + \ldots + \lambda_r \cdot x_r + \mu_1 \cdot y_1 + \ldots + \mu_s \cdot y_s \in U,$$

weil es eine endliche Linearkombination von Elementen in M ist, und ebenso ist

$$\lambda \cdot x = \sum_{i=1}^{r} (\lambda \cdot \lambda_i) \cdot x_i \in U.$$

Also ist U ein Unterraum von V. □

Beispiel 2.12

a. Ist $M = \{x_1, \ldots, x_n\} \subseteq V$ endlich, so ist das Erzeugnis von M

$$\mathrm{Lin}\,(x_1, \ldots, x_n) := \mathrm{Lin}\,(M) = \{\lambda_1 \cdot x_1 + \ldots + \lambda_n \cdot x_n \mid \lambda_1, \ldots, \lambda_n \in K\}.$$

Insbesondere gilt $\mathrm{Lin}\,(x) := \mathrm{Lin}\,(\{x\}) = \{\lambda \cdot x \mid \lambda \in K\}$.

b. Die Lineare Hülle der Vektoren $x_1 = (1,0)^t$ und $x_2 = (0,1)^t$ in $\mathbb{R}^2$ ist

$$\mathrm{Lin}\,(x_1, x_2) = \{\lambda_1 \cdot x_1 + \lambda_2 \cdot x_2 = (\lambda_1, \lambda_2) \mid \lambda_1, \lambda_2 \in \mathbb{R}\} = \mathbb{R}^2.$$

c. Die Lineare Hülle von $x = (1, m)^t \in \mathbb{R}^2$ ist die Gerade

$$\mathrm{Lin}\,(x) = \{(\lambda, \lambda m)^t \mid \lambda \in \mathbb{R}\}.$$

d. Es gilt offenbar $\operatorname{Lin}\left(t^0, t^1, \ldots, t^n\right) = P_n$.

Proposition 2.13 (Summe zweier Unterräume)
Es seien U *und* U' *Unterräume des* K*-Vektorraums* V*. Dann gilt*

$$U + U' := \{u + u' \mid u \in U, u' \in U'\} = \operatorname{Lin}(U \cup U') \leq V,$$

und wir nennen diesen Unterraum von V *die* Summe *von* U *und* U'.

Beweis: Wegen Proposition 2.11 ist $U + U'$ in $\operatorname{Lin}(U \cup U')$ enthalten, da die Elemente von $U + U'$ Linearkombinationen von Elementen in $U \cup U'$ sind.

Umgekehrt ist jede Linearkombination von Elementen in $U \cup U'$ von der Form $\sum_{i=1}^r \lambda_i \cdot u_i + \sum_{j=1}^s \mu_j \cdot u_j'$ mit $u_i \in U$, $u_j' \in U'$ und $\lambda_i, \mu_j \in K$. Da U und U' aber Unterräume sind, ist

$$u := \sum_{i=1}^{r} \lambda_i \cdot u_i \in U$$

und

$$u' := \sum_{j=1}^{s} \mu_j \cdot u_j' \in U'.$$

Deshalb ist die Linearkombination

$$\sum_{i=1}^{r} \lambda_i \cdot u_i + \sum_{j=1}^{s} \mu_j \cdot u_j' = u + u' \in U + U'$$

in $U + U'$, und mit Proposition 2.11 enthält $U + U'$ auch $\operatorname{Lin}(U \cup U')$. □

Bemerkung 2.14 (Summen von Unterräumen)

a. Die Summe zweier Unterräume ersetzt ihre Vereinigung in der Theorie der Vektorräume. Sie ist der kleinste Unterraum, der beide enthält. Im Beispiel aus Bemerkung 2.8 ist die Summe der beiden Unterräume ganz $\mathbb{R}^2$.

b. Analog zu Proposition 2.13 zeigt man allgemeiner: Sind $U_1, \ldots, U_n$ Unterräume des K-Vektorraums V, so gilt

$$U_1 + \ldots + U_n := \{u_1 + \ldots + u_n \mid u_i \in U_i\} = \operatorname{Lin}(U_1 \cup \ldots \cup U_n).$$

Beispiel 2.15
Jeder Vektor x in $U + U'$ läßt sich schreiben als $x = u + u'$ mit $u \in U$ und $u' \in U'$, diese Darstellung muß aber nicht eindeutig sein.

Sind z.B. $U = \operatorname{Lin}((1,0,0)^t, (0,1,1)^t)$ und $U' = \operatorname{Lin}((1,1,0)^t, (1,0,1)^t)$ als Unterräume von $\mathbb{R}^3$ gegeben, so können wir den Vektor $x = (1, 0, -1)^t$ auf folgende beiden Weisen als Summe zweier Vektoren in U und U' schreiben:

$$x = (0, -1, -1)^t + (1, 1, 0)^t = (2, 0, 0)^t + (-1, 0, -1)^t.$$

Definition 2.16 (Direkte Summe)
Es seien $U_1, \ldots, U_n$ Unterräume des K-Vektorraums V. Wir nennen die Summe $U = U_1 + \ldots + U_n$ eine *direkte Summe*, wenn sich jeder Vektor $x \in U_1 + \ldots + U_n$

auf eindeutige Weise als Summe $x = u_1 + \ldots + u_n$ mit $u_i \in U_i$ schreiben läßt. Wir schreiben dann $U = U_1 \oplus \ldots \oplus U_n$.

Proposition 2.17 (Direkte Summe zweier Unterräume)
Es seien U, U' *und* W *Unterräume des* K*-Vektorraums* V.
Genau dann gilt $W = U \oplus U'$, *wenn* $W = U + U'$ *und* $U \cap U' = \{0\}$.

Beweis: Ist die Summe $W = U \oplus U'$, so gilt insbesondere $W = U + U'$. Für $x \in U \cap U'$, gilt zudem

$$x = x + 0 = 0 + x \in U + U',$$

und wegen der Eindeutigkeit der Darstellung in $U + U'$ muß $x = 0$ sein.

Ist umgekehrt $W = U+U'$ und $U\cap U' = \{0\}$ und sind $x_1+x_1' = x_2+x_2' \in U+U' = W$ mit $x_i \in U$ und $x_i' \in U'$, $i = 1, 2$, so gilt:

$$x_1 - x_2 = x_2' - x_1' \in U \cap U' = \{0\}.$$

Also ist $x_1 = x_2$ und $x_1' = x_2'$, d. h. die Darstellung ist eindeutig. □

Beispiel 2.18
Betrachte die Unterräume $U = \text{Lin}\left((1, 1, 1)^t\right)$ und $U' = \text{Lin}\left((1, 0, 1)^t\right)$ von $\mathbb{R}^3$. Ein Vektor x liegt genau dann im Durchschnitt $U \cap U'$, wenn es $\lambda, \mu \in K$ gibt mit

$$x = \lambda \cdot (1, 1, 1)^t = (\lambda, \lambda, \lambda)^t$$

und

$$x = \mu \cdot (1, 0, 1)^t = (\mu, 0, \mu)^t.$$

Gleichsetzen der beiden Ausdrücke liefert die Bedingungen $\lambda = \mu$ und $\lambda = 0$, also gilt $x = (0, 0, 0)^t$, d.h.

$$U \cap U' = \{(0, 0, 0)^t\}.$$

Damit ist die Summe $U + U'$ eine direkte Summe.

C) Lineare Abbildungen

Zu jeder Struktur gehören die strukturerhaltenden Abbildungen.

Definition 2.19 (Lineare Abbildungen)
Es seien V und W zwei K-Vektorräume.

a. Eine Abbildung $f : V \to W$ heißt K-*lineare Abbildung* oder *Vektorraumhomomorphismus*, wenn für alle $\lambda \in K$ und $x, y \in V$ gilt

$$f(x + y) = f(x) + f(y) \quad \text{und} \quad f(\lambda \cdot x) = \lambda \cdot f(x).$$

b. Eine injektive (bzw. surjektive bzw. bijektive) K-lineare Abbildung heißt auch *Monomorphismus* (bzw. *Epimorphismus* bzw. *Isomorphismus*). Gilt $V = W$, so nennen wir eine K-lineare Abbildung auch einen *Endomorphismus*, und ist sie zudem bijektiv, so sprechen wir von einem *Automorphismus*.

c. Existiert ein Isomorphismus von V nach W, so nennen wir V und W *isomorph* und schreiben $V \cong W$.

d. Die Menge aller K-linearen Abbildungen von V nach W bezeichnen wir mit $\mathrm{Hom}_K(V, W)$ und die Menge aller Endomorphismen von V mit $\mathrm{End}_K(V)$.

Bemerkung 2.20
Die beiden Bedingungen in Definition 2.19 a. lassen sich zusammenfassen zu der Bedingung $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$ für alle $\lambda, \mu \in K$ und $x, y \in V$.

Beispiel 2.21 a. Die Abbildung $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x_1, x_2)^t \mapsto x_1 - x_2$ ist $\mathbb{R}$-linear. Denn für $x = (x_1, x_2)$ und $y = (y_1, y_2)$ sowie $\lambda \in \mathbb{R}$ gilt

$$\begin{aligned} f(x+y) &= f\big((x_1 + y_1, x_2 + y_2)\big) = x_1 + y_1 - x_2 - y_2 \\ &= x_1 - x_2 + y_1 - y_2 = f(x) + f(y) \end{aligned}$$

und

$$f(\lambda x) = f\big((\lambda x_1, \lambda x_2)\big) = \lambda x_1 - \lambda x_2 = \lambda \cdot (x_1 - x_2) = \lambda \cdot f(x).$$

b. Ist I ein Intervall, so ist die Abbildung

$$D : \mathcal{C}^1(I, \mathbb{R}) \longrightarrow \mathcal{C}(I, \mathbb{R}) : f \mapsto f'$$

$\mathbb{R}$-linear, da aus der Linearität der Ableitung folgt

$$D(\lambda f + \mu g) = (\lambda f + \mu g)' = \lambda \cdot f' + \mu \cdot g' = \lambda \cdot D(f) + \mu \cdot D(g).$$

c. Die Abbildung

$$f : \mathbb{R}[t] \longrightarrow \mathbb{R}^{\mathbb{R}} : \sum_{k=0}^{n} a_k \cdot t^k \mapsto (a_k)_{k \in \mathbb{N}},$$

wobei $a_k = 0$ für $k > n$ gelten soll, ist eine injektive $\mathbb{R}$-lineare Abbildung, also ein Monomorphismus. Ihr Bild, der Unterraum

$$\mathrm{Im}(f) = \{(a_k)_{k \in \mathbb{N}} \mid \text{ nur endlich viele } a_k \text{ sind nicht } 0\}$$

der *abbrechenden Folgen* in $\mathbb{R}$, ist mithin isomorph zu $K[t]$.

d. Die formale Ableitung

$$d : K[t] \longrightarrow K[t] : \sum_{k=0}^{n} a_k \cdot t^k \mapsto \sum_{k=1}^{n} k \cdot a_k \cdot t^{k-1}$$

ist eine K-lineare Abbildung, wie man leicht nachrechnet.

Lemma 2.22 (Einfache Eigenschaften linearer Abbildungen)
Seien U, V und W K-Vektorräume und $f : U \longrightarrow V$ und $g : V \longrightarrow W$ seien K-linear. Ferner seien $x, x_1, \ldots, x_n \in U$ und $\lambda_1, \ldots, \lambda_n \in K$. Dann gelten:

a. $f(0_U) = 0_V$ *und* $f(-x) = -f(x)$.

b. $f(\lambda_1 x_1 + \ldots + \lambda_n x_n) = \lambda_1 f(x_1) + \ldots + \lambda_n f(x_n)$.

c. *Ist f bijektiv, so ist $f^{-1} : V \longrightarrow U$ K-linear.*

d. $g \circ f : U \longrightarrow W$ *ist* K*-linear.*

e. $\mathrm{Hom}_K(U, V)$ *ist ein Unterraum von* V^U.

Beweis: a. Aus der Verträglichkeit mit der Skalarmultiplikation folgen

$$f(0_U) = f(0_K \cdot 0_U) = 0_K \cdot f(0_U) = 0_V$$

und

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

b. Die Aussage folgt mittels Induktion aus den beiden Bedingungen für Linearität.

c. Seien $y, y' \in V$ und $\lambda, \lambda' \in K$ sowie $x = f^{-1}(y)$ und $x' = f^{-1}(y')$. Wegen der Linearität von f gilt

$$f(\lambda \cdot x + \lambda' \cdot x') = \lambda \cdot f(x) + \lambda' \cdot f(x').$$

Wenden wir auf beiden Seiten f^{-1} an, so erhalten wir

$$\begin{aligned} \lambda \cdot f^{-1}(y) + \lambda' \cdot f^{-1}(y') &= \lambda \cdot x + \lambda' \cdot x' = f^{-1}\big(f(\lambda \cdot x + \lambda' \cdot x')\big) \\ &= f^{-1}\big(\lambda \cdot f(x) + \lambda' \cdot f(x')\big) = f^{-1}\big(\lambda \cdot y + \lambda' \cdot y'\big). \end{aligned}$$

Mithin ist f^{-1} eine lineare Abbildung.

d. Seien $\lambda, \mu \in K$ und $x, y \in U$, so gelten

$$\begin{aligned} (g \circ f)(\lambda x + \mu y) &= g\big(f(\lambda x + \mu y)\big) = g(\lambda f(x) + \mu f(y)) \\ &= \lambda g\big(f(x)\big) + \mu g(\big(f(y)\big) = \lambda (g \circ f)(x) + \mu (g \circ f)(y). \end{aligned}$$

e. Dies folgt aus Aufgabe 2.44.

□

Proposition 2.23 (f_A ist linear.)
Für $A \in \mathrm{Mat}(m \times n, K)$ *ist* $f_A : K^n \longrightarrow K^m$ *eine* K*-lineare Abbildung.*

Beweis: Aus Lemma 1.8 folgt für $x, y \in K^n$ und $\lambda \in K$

$$f_A(x + y) = A(x + y) = Ax + Ay = f_A(x) + f_A(y)$$

und

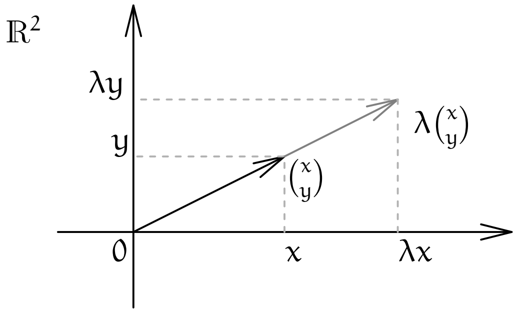
$$f_A(\lambda x) = A(\lambda x) = \lambda(Ax) = \lambda f_A(x).$$

□

Beispiel 2.24

a. Im Fall $n = 1$ und $A = (a)$ ist die K-lineare Abbildung $f_A : K \to K : x \mapsto a \cdot x$ gerade die Multiplikation mit a.

b. Die lineare Abbildung $f_A : \mathbb{R}^2 \mapsto \mathbb{R}^2 : (x, y)^t \mapsto (\lambda x, \lambda y)$ zu $A = \lambda \mathbb{1}_2$ ist eine *Streckung* um den Faktor λ.



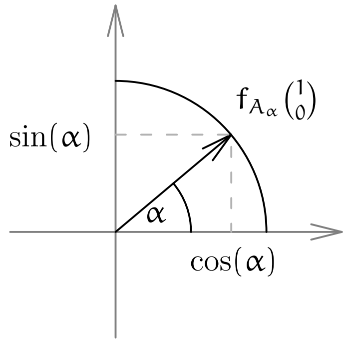
c. Für $\alpha \in \mathbb{R}$ setzen wir

$$A_\alpha := \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

Dann ist die lineare Abbildung $f_{A_\alpha} : \mathbb{R}^2 \to \mathbb{R}^2$ eine Drehung um den Winkel α. Beachte dazu, daß

$$A_\alpha e_1 = \big(\cos(\alpha), \sin(\alpha)\big)^t \quad \text{und} \quad A_\alpha e_2 = \big(-\sin(\alpha), \cos(\alpha)\big)^t,$$

woraus die Aussage für die *Einheitsvektoren* e_1 und e_2 unmittelbar folgt.



Daraus leitet sich die Aussage für einen beliebigen Vektor $(x, y)^t$ mittels der Linearität von f_{A_α} ab: $f_{A_\alpha}\big((x, y)^t\big) = x f_{A_\alpha}(e_1) + y f_{A_\alpha}(e_2)$.

d. Ist $n \geq m$, so ist die Abbildung

$$\mathrm{pr} : K^n \to K^m : (x_1, \ldots, x_n)^t \mapsto (x_1, \ldots, x_m)^t$$

eine K-lineare Abbildung, genannt die kanonische *Projektion*.
Ist $m \geq n$, dann ist die kanonische *Inklusion*

$$i_{K^n} : K^n \to K^m : (x_1, \ldots, x_n)^t \mapsto (x_1, \ldots, x_n, 0, \ldots, 0)^t$$

ebenfalls K-linear. Beides prüft man leicht nach.

Proposition 2.25 (Kern und Bild sind Unterräume)
Es seien V und W K-Vektorräume und $f : V \longrightarrow W$ sei K-linear.

a. *Ist U ein Unterraum von V, so ist $f(U)$ ein Unterraum von W.*
b. *Ist U ein Unterraum von W, so ist $f^{-1}(U)$ ein Unterraum von V.*
c. *Das* Bild $\mathrm{Im}(f) = f(V)$ *von f ist ein Unterraum von W.*
d. *Der* Kern *von f,* $\mathrm{Ker}(f) = \{x \in V \mid f(x) = 0\}$*, ist ein Unterraum von V.*

Beweis:

a. Es sei U ein Unterraum von V. Dann ist $0_V \in U$ und somit $0_W = f(0_V) \in f(U)$, so daß $f(U)$ nicht leer ist. Sind $\lambda \in K$ und $u = f(x), v = f(y) \in f(U)$ mit $x, y \in U$, so gilt

$$u + v = f(x) + f(y) = f(x + y) \in f(U)$$

und

$$\lambda u = \lambda f(x) = f(\lambda x) \in f(U).$$

Also ist $f(U)$ ein Unterraum von W.

b. Es sei U ein Unterraum von W. Dann ist $0_W \in U$ und wegen $f(0_V) = 0_W$ ist dann $0_V \in f^{-1}(U)$, so daß $f^{-1}(U)$ nicht leer ist. Sind $\lambda \in K$ und $x, y \in f^{-1}(U)$, so gilt $f(x), f(y) \in U$ und somit

$$f(x + y) = f(x) + f(y) \in U$$

und

$$f(\lambda x) = \lambda f(x) \in U.$$

Also auch $x + y \in f^{-1}(U)$ und $\lambda x \in f^{-1}(U)$, so daß $f^{-1}(U)$ ein Unterraum von V ist.

c. Dies folgt aus a. mit $U = V$.

d. Dies folgt aus b. mit $U = \{0_W\}$.

□

Beispiel 2.26
Die Abbildung $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x_1, x_2) \mapsto x_1 - x_2$ aus Beispiel 2.21 hat den Kern

$$\mathrm{Ker}(f) = \{x \in \mathbb{R}^2 \mid f(x) = x_1 - x_2 = 0\} = \left\{(\lambda, \lambda)^t \;\middle|\; \lambda \in K\right\}$$

und das Bild ist $\mathrm{Im}(f) = \mathbb{R}$.

Proposition 2.27 (Injektivität linearer Abbildungen)
Eine lineare Abbildung $f : V \longrightarrow W$ *ist genau dann injektiv, wenn* $\mathrm{Ker}(f) = \{0_V\}$.

Beweis: Ist f injektiv und $x \in \mathrm{Ker}(f)$, so folgt aus $f(x) = 0_W = f(0_V)$ auch $x = 0_V$. Also ist $\mathrm{Ker}(f) = \{0_V\}$.

Sei nun umgekehrt $\mathrm{Ker}(f) = \{0_V\}$ und seien $x, y \in V$ mit $f(x) = f(y)$. Dann folgt $f(x - y) = f(x) - f(y) = 0_W$ und damit $x - y \in \mathrm{Ker}(f) = \{0_V\}$. Also ist $x = y$, und somit ist f injektiv. □

D) Faktorräume

Definition 2.28 (Faktorraum)
Es sei V ein K-Vektorraum und U ein Unterraum von V.

a. Für $x \in V$ nennen wir

$$\overline{x} := x + U := \{x + u \mid u \in U\}$$

die *Restklasse* oder *Nebenklasse* von x modulo U und x einen *Vertreter* der Restklasse. Man nennt $x+U$ auch einen *affinen Raum* parallel zum Unterraum U mit Aufpunkt x.

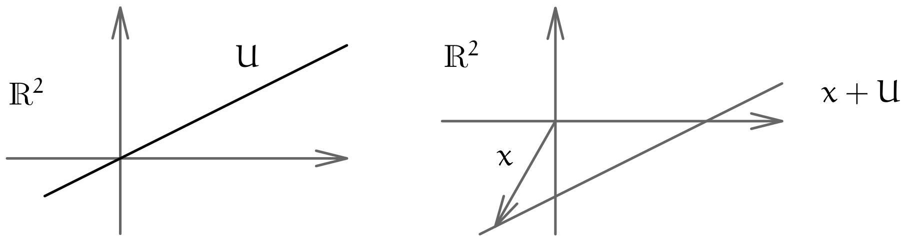


ABBILDUNG 2. Ein affiner Raum $x + U$ zum Unterraum U.

Man beachte, daß aus der Notation $\overline{x}$ nicht mehr abzulesen ist, modulo welchem Unterraum man rechnet. Die Notation ist deshalb mit Vorsicht zu verwenden.

b. Wir nennen die Menge der Restklassen modulo U

$$V/U := \{x + U \mid x \in V\} = \{\overline{x} \mid x \in V\}$$

auch den *Faktorraum* von V modulo U.

Bemerkung 2.29 (Restklassen als Äquivalenzklassen)
Ist U ein Unterraum des K-Vektorraums V, so wird durch

$$x \sim y \;:\Longleftrightarrow\; x - y \in U$$

für $x, y \in V$ eine *Äquivalenzrelation* auf V definiert, wie man leicht nachprüft. Die Äquivalenzklasse von x ist dann gerade $x + U$, und V/U ist die Menge aller Äquivalenzklassen der Äquivalenzrelation. Dies ist konzeptionell der richtige Weg, die Restklassen von V modulo U sowie den Faktorraum einzuführen.

Beweis: Wir wollen zunächst nachweisen, daß $\sim$ eine Äquivalenzrelation ist. Dazu wählen wir $x, y, z \in V$. Wegen $x - x = 0 \in U$ gilt $x \sim x$ und $\sim$ ist reflexiv. Gilt $x \sim y$, d.h. $x - y \in U$, so gilt auch $y - x = -(x - y) \in U$ und somit $y \sim x$, d.h. $\sim$ ist auch symmetrisch. Gilt schließlich $x \sim y$ und $y \sim z$, d.h. $x - y, y - z \in U$, so gilt ebenfalls $x - z = (x - y) + (y - z) \in U$ und somit $x \sim z$, d.h. $\sim$ ist auch transitiv. Wir haben also gezeigt, daß $\sim$ eine Äquivalenzrelation ist, und hierbei haben wir die Unterraumeigenschaften von U ausgenutzt.

Es bleibt noch für $x \in V$ zu zeigen, daß $x+U$ die zu x gehörige Äquivalenzklasse ist. Sei dazu zunächst y in der Äquivalenzklasse von x gegeben, d.h. $y \sim x$ und damit

$y-x \in U$. Dann ist aber auch $y = x+(y-x) \in x+U$. Sei nun umgekehrt $y \in x+U$ gegeben. Dann gibt es ein $u \in U$ mit $y = x+u$ und somit ist $y-x = u \in U$ und $y \sim x$, was zur Folge hat, daß y zur Äquivalenzklasse von x gehört. □

Lemma 2.30 (Rechnen mit Restklassen)
Es sei V *ein* K*-Vektorraum,* U *ein Unterraum,* $x, x', y, y' \in V$ *und* $\lambda \in K$*. In* V/U *gelten dann die folgenden Aussagen:*

a. *Entweder* $\overline{x} = \overline{y}$ *oder* $\overline{x} \cap \overline{y} = \emptyset$.

b. *Es gilt:*

$$\overline{x} = \overline{y} \quad \Longleftrightarrow \quad x-y \in U.$$

Insbesondere, $\overline{x} = \overline{0} = U$ *genau dann, wenn* $x \in U$.

c. *Gilt* $\overline{x} = \overline{x'}$ *und* $\overline{y} = \overline{y'}$*, so gelten auch*

$$\overline{x+y} = \overline{x'+y'} \quad \textit{und} \quad \overline{\lambda x} = \overline{\lambda x'}.$$

Beweis:

a. Dies folgt unmittelbar aus Proposition A6.10, da $\overline{x}$ und $\overline{y}$ Äquivalenzklassen sind.

b. Dies folgt aus Bemerkung 2.29.

c. Wir wollen hier zur Verdeutlichung $x+U$ und $y+U$ statt $\overline{x}$ und $\overline{y}$ schreiben. Aus $x+U = x'+U$ sowie $y+U = y'+U$ folgt nach b.

$$x-x', y-y' \in U.$$

Damit gilt dann auch

$$(x+y)-(x'+y') = (x-x')+(y-y') \in U$$

und

$$(\lambda x - \lambda x') = \lambda \cdot (x-x') \in U.$$

Wegen b. gilt dann wieder $(x+y)+U = (x'+y')+U$ und $\lambda x + U = \lambda x' + U$.

□

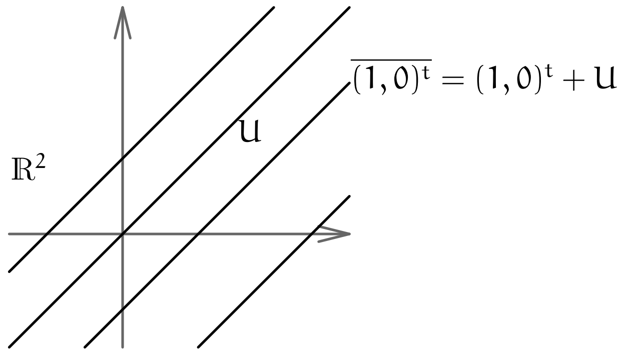
Beispiel 2.31
Sei $V = \mathbb{R}^2$ und $U = \text{Lin}\left((1,1)^t\right)$. Dann gilt

$$\overline{(1,0)^t} = (1,0)^t + U = (2,1)^t + U = \overline{(2,1)^t}$$

und

$$\overline{(1,0)^t} = (1,0)^t + U \neq (0,1)^t + U = \overline{(0,1)^t}.$$

Die Restklassen in V/U sind in diesem Fall genau die Geraden, die parallel zur Geraden U sind.



Satz 2.32 (Der Faktorraum ist ein Vektorraum.)
Es sei V *ein* K-*Vektorraum und* U *ein Unterraum. Dann definiert*

$$\overline{x} + \overline{y} := \overline{x+y} \tag{2}$$

und

$$\lambda \cdot \overline{x} := \overline{\lambda x} \tag{3}$$

für $\overline{x}, \overline{y} \in V/U$ *und* $\lambda \in K$ *eine Addition und eine Skalarmultiplikation auf* V/U *bezüglich derer der Faktorraum* V/U *ein* K-*Vektorraum ist.*

Zudem ist die Abbildung

$$\pi : V \longrightarrow V/U : x \mapsto \overline{x}$$

eine surjektive K-*lineare Abbildung mit* $\operatorname{Ker}(\pi) = U$, *die wir die* Restklassenabbildung *nennen.*

Beweis: Bevor wir uns den Vektorraumaxiomen widmen können, müssen wir zeigen, daß durch (2) und (3) überhaupt Operationen auf V/U definiert werden. Das Problem dabei ist, daß wir zur Definition der Summe und der Skalarmultiplikation Vertreter der Restklassen verwendet haben. Diese sind aber nicht eindeutig bestimmt. Wir müssen also sicherstellen, daß wir das gleiche Ergebnis erhalten, wenn wir andere Vertreter wählen. Man nennt dies die *Wohldefiniertheit* der Operationen. Dazu ist zu zeigen, daß aus für $\overline{x} = \overline{x'}$ und $\overline{y} = \overline{y'}$ auch $\overline{x+y} = \overline{x'+y'}$ und $\overline{\lambda x} = \overline{\lambda x'}$ gelten. Dies folgt aber aus Lemma 2.30.

Wir wollen nun zeigen, daß V/U den Vektorraumaxiomen genügt. Dazu seien $\overline{x}, \overline{y}, \overline{z} \in V/U$ und $\lambda, \mu \in K$ gegeben. Dann gilt

$$\big(\overline{x} + \overline{y}\big) + \overline{z} = \overline{x+y} + \overline{z} = \overline{(x+y)+z} = \overline{x+(y+z)} = \overline{x} + \overline{y+z} = \overline{x} + \big(\overline{y} + \overline{z}\big).$$

Die Addition ist also assoziativ. Zudem gilt

$$\overline{0} + \overline{x} = \overline{0+x} = \overline{x}$$

und

$$\overline{-x} + \overline{x} = \overline{-x+x} = \overline{0},$$

so daß $\overline{0}$ der Nullvektor ist und $\overline{x}$ ein Inverses besitzt. Da zudem

$$\overline{x} + \overline{y} = \overline{x+y} = \overline{y+x} = \overline{y} + \overline{x}$$

gilt, ist V/U eine abelsche Gruppe bezüglich $+$. Ähnlich wie sich die Axiome für die Addition von V auf V/U vererbt haben, vererben sich auch die Axiome für die Skalarmultiplikation.

$$(\lambda+\mu)\cdot\overline{x}=\overline{(\lambda+\mu)\cdot x}=\overline{\lambda x+\mu x}=\overline{\lambda x}+\overline{\mu x}=\lambda\cdot\overline{x}+\mu\cdot\overline{x}$$

und

$$\lambda\cdot\overline{x+y}=\overline{\lambda\cdot(x+y)}=\overline{\lambda x+\lambda y}=\overline{\lambda x}+\overline{\lambda y}=\lambda\cdot\overline{x}+\lambda\cdot\overline{y}$$

und

$$(\lambda\cdot\mu)\cdot\overline{x}=\overline{(\lambda\cdot\mu)\cdot x}=\overline{\lambda\cdot(\mu\cdot x)}=\lambda\overline{\mu\cdot x}=\lambda\cdot\big(\mu\cdot\overline{x}\big)$$

und

$$1\cdot\overline{x}=\overline{1\cdot x}=\overline{x}.$$

Also ist V/U ein K-Vektorraum.

Es bleibt, die Aussagen zur Restklassenabbildung π zu zeigen. Die Linearität von π folgt unmittelbar aus der Definition der Operationen auf V/U. Sind $x,y\in V$ und $\lambda\in K$, dann gilt

$$\pi(x+y)=\overline{x+y}=\overline{x}+\overline{y}=\pi(x)+\pi(y)$$

und

$$\pi(\lambda x)=\overline{\lambda x}=\lambda\cdot\overline{x}=\lambda\cdot\pi(x).$$

Außerdem ist π surjektiv, da jedes $\overline{x}\in V/U$ sich schreiben läßt als $\overline{x}=\pi(x)$. Und es gilt

$$x\in\operatorname{Ker}(\pi)\iff\overline{x}=\pi(x)=\overline{0}\iff x\in U.$$

Damit sind alle Aussagen des Satzes bewiesen. □

Bemerkung 2.33 (Die vier Rechenregeln für den Faktorraum)
Um mit dem Faktorraum rechnen zu können, braucht man nur die Rechenregeln:

a. $\overline{0}=0+U=U$ ist der Nullvektor.

b. $\overline{x}+\overline{y}=\overline{x+y}$.

c. $\lambda\cdot\overline{x}=\overline{\lambda x}$.

d. $\overline{x}=\overline{y}\quad\iff\quad x-y\in U$.

Satz 2.34 (Homomorphiesatz)
Ist $f:V\longrightarrow W$ *eine* K*-lineare Abbildung, so ist*

$$\overline{f}:V/\operatorname{Ker}(f)\longrightarrow\operatorname{Im}(f):\overline{x}\mapsto f(x)$$

ein Isomorphismus. Insbesondere gilt also $V/\operatorname{Ker}(f)\cong\operatorname{Im}(f)$.

Beweis: Da wir für die Definition von $\overline{f}(\overline{x})$ wieder den Restklassenvertreter x verwendet haben, müssen wir wieder zeigen, daß unsere Definition nicht von der speziellen Wahl des Vertreters abhängt. Man sagt wieder, wir müssen die *Wohldefiniertheit* von $\overline{f}$ zeigen.

Seien dazu $\overline{x} = x + \mathrm{Ker}(f) = y + \mathrm{Ker}(f) = \overline{y}$ gegeben. Dann gilt

$$x - y \in \mathrm{Ker}(f),$$

und mithin $0 = f(x-y) = f(x) - f(y)$, oder alternativ $f(x) = f(y)$. Die Abbildung $\overline{f}$ ist also wohldefiniert.

Die Linearität von $\overline{f}$ folgt dann aus der Linearität von f. Seien dazu $\overline{x}, \overline{y} \in V/\mathrm{Ker}(f)$ und $\lambda \in K$, dann gilt

$$\overline{f}\big(\overline{x} + \overline{y}\big) = \overline{f}\big(\overline{x+y}\big) = f(x+y) = f(x) + f(y) = \overline{f}\big(\overline{x}\big) + \overline{f}\big(\overline{y}\big)$$

und

$$\overline{f}\big(\lambda \cdot \overline{x}\big) = \overline{f}\big(\overline{\lambda \cdot x}\big) = f(\lambda \cdot x) = \lambda \cdot f(x) = \lambda \cdot \overline{f}\big(\overline{x}\big).$$

Es bleibt noch zu zeigen, daß $\overline{f}$ surjektiv und injektiv ist.

Ist $y \in \mathrm{Im}(f)$, so gibt es ein $x \in V$ mit $f(x) = y$, und somit gilt

$$y = f(x) = \overline{f}\big(\overline{x}\big).$$

Also ist $\overline{f}$ surjektiv.

Für die Injektivität nutzen wir Proposition 2.27. Es gilt

$$\overline{x} \in \mathrm{Ker}\left(\overline{f}\right) \iff 0 = \overline{f}\big(\overline{x}\big) = f(x) \iff x \in \mathrm{Ker}(f) \iff \overline{x} = \overline{0}.$$

Also enthält der Kern von $\overline{f}$ nur den Nullvektor, und somit ist $\overline{f}$ injektiv. □

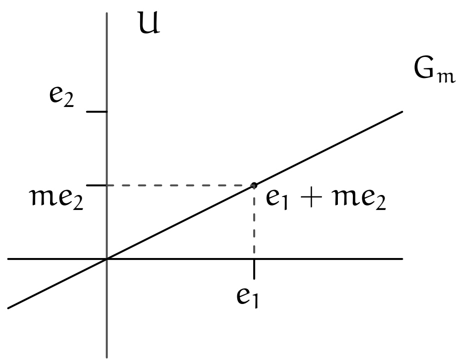
Definition 2.35 (Direkte Komplemente)
Es sei V ein K-Vektorraum und U und U' seien Unterräume von V. Dann heißt U' ein (direktes) *Komplement* von U, falls $V = U \oplus U'$.

Beispiel 2.36 (Komplemente sind nicht eindeutig.)
Ist $V = \mathbb{R}^2$ und $U = \mathrm{Lin}\,(e_2)$ die y-Achse, dann ist die Ursprungsgerade mit Steigung m

$$G_m := \mathrm{Lin}\,(e_1 + me_2)$$

für jedes $m \in \mathbb{R}$ ein Komplement von U. Beachte dazu nur, daß sich die Geraden U und G_m nur im Ursprung schneiden, d.h. $U \cap G_m = \{0\}$, und daß ein beliebiger Vektor $(x, y)^t \in \mathbb{R}^2$ sich schreiben läßt als

$$(x, y)^t = x \cdot (e_1 + me_2) + (-xm + y) \cdot e_2 \in G_m + U.$$



Proposition 2.37 (Der Faktorraum als Komplementersatz)
Sei V ein K-Vektorraum, U ein Unterraum von V und U' ein Komplement von U. Dann ist die Einschränkung der Restklassenabbildung

$$\pi_| : U' \to V/U : x \mapsto \overline{x}$$

auf U' ein Isomorphismus. Insbesondere sind je zwei Komplemente von U isomorph.

Beweis: Es ist klar, daß $\pi_|$ als Einschränkung einer K-linearen Abbildung wieder K-linear ist.

Wir zeigen zunächst, daß $\pi_|$ surjektiv ist. Sei dazu $\overline{x} \in V/U$ gegeben. Wegen $V = U \oplus U'$ läßt sich x als $x = y + z$ mit $y \in U$ und $z \in U'$ schreiben. Damit gilt:

$$\overline{x} = \overline{z} = \pi_|(z) \in \operatorname{Im}\left(\pi_|\right).$$

Also ist $\pi_|$ surjektiv.

Es bleibt zu zeigen, daß $\pi_|$ injektiv ist, d. h. $\operatorname{Ker}\left(\pi_|\right) = \{0\}$. Sei dazu $z \in \operatorname{Ker}\left(\pi_|\right)$, dann gilt

$$\overline{0} = \pi_|(z) = \overline{z}.$$

D. h. $z \in U$. Damit gilt aber $z \in U \cap U' = \{0\}$, also $z = 0$.

Seien schließlich U' und U'' zwei Komplemente von U, dann ist die Komposition

$$U' \xrightarrow{\pi_{|U'}} V/U \xrightarrow{\pi_{|U''}^{-1}} U''$$

ein Isomorphismus von U' nach U''. Die beiden Komplemente sind also isomorph zueinander. □

Bemerkung 2.38
Daß V/U isomorph zu jedem Komplement von U ist, heißt im wesentlichen, daß man bei Betrachtungen, bei denen man ein Komplement von U benötigt, stattdessen auch mit V/U arbeiten kann. Während es sehr viele Komplemente von U geben kann, gibt es nur *einen* Faktorraum. Dieser ist durch U eindeutig bestimmt. Das ist unter Umständen ein großer Vorteil!

E) Ringe und Moduln

Bemerkung 2.39 (Kommutative Ringe mit Eins)
Eine Menge K mit zwei zweistelligen Operationen $+$ und $\cdot$, die allen Axiomen eines Körpers genügt, außer eventuell der Existenz von multiplikativen Inversen, nennt man einen *kommutativen Ring mit Eins*.

Ein typisches Beispiel dafür sind die ganzen Zahlen $\mathbb{Z}$, und in der Vorlesung algebraische Strukturen lernt man den Polynomring, z.B.

$$\mathbb{R}[t] = \left\{ \sum_{k=0}^{n} a_k \cdot t^k \;\middle|\; a_k \in \mathbb{R}, n \in \mathbb{N} \right\},$$

als ein weiteres typisches Beispiel kennen.

Der Nachteil von Ringen im Vergleich zu Körpern ist, daß man im allgemeinen nicht durch die Elemente teilen darf. Deshalb sind nicht alle Aussagen, die für Körper gelten auch gleichermaßen für Ringe richtig. Aber immer dann, wenn man ohne Division auskommt, geht alles gut. Das trifft auf die meisten Definitionen und Aussagen in diesem Kapitel zu!

Bemerkung 2.40 (Matrizen über Ringen)
Setzen wir in Abschnitt 1 nur voraus, daß K ein kommutativer Ring mit Eins ist, so bleiben alle Definitionen und Aussagen korrekt.

Bemerkung 2.41 (Moduln und lineare Abbildungen)
Setzen wir in Definition 2.1 nur voraus, daß K ein kommutativer Ring mit Eins ist, so nennen wir V einen K-*Modul*. Der Name ist dabei aber auch das einzige, was sich ändert. Entsprechend wird man in Definition 2.4 dann von einem *Untermodul* statt von einem Unterraum reden. Alle anderen Begriffe, Beispiele und Aussagen dieses Abschnitts bleiben ohne Änderung und mit dem jeweils gleichen Beweis korrekt, bis auf eine einzige Ausnahme:

$$\lambda \cdot v = 0 \quad \not\Longrightarrow \quad \lambda = 0 \text{ oder } v = 0.$$

Im Beweis dieser Aussage mußten wir durch λ teilen können, wenn λ nicht Null war!

Wir werden an einigen Stellen der Vorlesung lineare Abbildungen oder Matrizen über dem Polynomring benötigen. Deshalb werde ich immer mal wieder anmerken, welche Aussagen auch für Ringe und Moduln gelten. Wenn ohnehin vieles ohne Änderung im Beweis korrekt bleibt, hätte man die Aussagen natürlich auch gleich für Ringe und Moduln formulieren können. Die Erfahrung zeigt aber, daß die Mehrzahl der Studenten sich mit Körpern und Vektorräumen als Begriffen wohler fühlt.

Aufgaben

Aufgabe 2.42
Welche der folgenden Teilmengen von K^3 sind Unterräume des K^3? Begründe Deine Anworten.

a. $\{(x_1, x_2, x_3)^t \in \mathbb{R}^3 \mid x_1 \cdot x_2 = 2x_3\}$ für $K = \mathbb{R}$.
b. $\{(x_1, x_2, x_3)^t \in \mathbb{R}^3 \mid ax_1 + x_2 + x_3 = a + 1\}$ für ein festes $a \in \mathbb{R}$ für $K = \mathbb{R}$.
c. $\{(x_1, x_2, x_3)^t \in \mathbb{R}^3 \mid x_1 \leq 0\}$ für $K = \mathbb{R}$.
d. $\{(1, 0, 0)^t, (0, 1, 0)^t, (1, 1, 0)^t, (0, 0, 0)^t\}$ für $K = \mathbb{R}$ oder $K = \mathbb{F}_2$.

Aufgabe 2.43
Gegeben seien die folgenden Teilmengen des $\mathbb{Q}$–Vektorraums $V = \mathbb{Q}^4$:

a. $U_1 = \{(x, x + 1, x + 2, x + 4)^t \mid x \in \mathbb{Q}\}$,
b. $U_2 = \{(x, 2x, 3x, 4x)^t \mid x \in \mathbb{Q}\}$,

c. $U_3 = \{(x_1, x_2, x_3, x_4)^t \mid x_1, x_2, x_3, x_4 \in \mathbb{Q},\ x_3 = x_1 + x_2,\ x_4 = x_2 + x_3\}$,

d. $U_4 = \{(x_1, x_2, x_3, x_4)^t \mid x_1, x_2, x_3, x_4 \in \mathbb{Q},\ x_2 \geq 0\}$,

e. $U_5 = \{(x_1, x_2, x_3, x_4)^t \mid x_1, x_2, x_3, x_4 \in \mathbb{Z}\}$.

Welche dieser Mengen sind Unterräume von V? Begründe Deine Aussage.

Aufgabe 2.44
Seien U, V und W K-Vektorräume, $\lambda, \lambda' \in K$ und $f, f' \in \mathrm{Hom}_K(U, V)$ und $g, g' \in \mathrm{Hom}_K(V, W)$. Dann gelten:

a. $f + f', \lambda \cdot f \in \mathrm{Hom}_K(U, V)$, d.h. $\mathrm{Hom}_K(U, V)$ ist ein Unterraum von V^U.

b. $g \circ \big(\lambda f + \lambda' f'\big) = \lambda(g \circ f) + \lambda'\big(g \circ f'\big)$ und $\big(\lambda g + \lambda' g'\big) \circ f = \lambda(g \circ f) + \lambda'\big(g' \circ f\big)$.

c. $\lambda(g \circ f) = (\lambda g) \circ f = g \circ (\lambda f)$.

Aufgabe 2.45
Seien V, W K–Vektorräume und $f, g \in \mathrm{Hom}_K(V, W)$. Zeige

$$\mathrm{Ker}(f + g) \supseteq \mathrm{Ker}(f) \cap \mathrm{Ker}(g)$$

und

$$\mathrm{Im}(f + g) \subseteq \mathrm{Im}(f) + \mathrm{Im}(g).$$

Finde außerdem Beispiele, so dass die Inklusionen strikt sind.

Aufgabe 2.46 (f-invariante Unterräume)
Ist V ein K-Vektorraum, $f : V \longrightarrow V$ K-linear und $U \leq V$ ein Unterraum von V mit $f(U) \subseteq U$, so nennen wir U einen f-*invarianten* Unterraum von V.

Zeige, daß durch

$$f_U : U \longrightarrow U : x \mapsto f(x)$$

und

$$f_{V/U} : V/U \longrightarrow V/U : \overline{x} \mapsto \overline{f(x)}$$

K-lineare Abbildungen definiert werden.

Aufgabe 2.47
Es sei $V = \mathbb{R}^{\mathbb{R}}$ der $\mathbb{R}$-Vektorraum aller Abbildungen von $\mathbb{R}$ nach $\mathbb{R}$ und

$$U = \{f : \mathbb{R} \longrightarrow \mathbb{R} \mid f(-x) = -f(x)\ \forall\, x \in \mathbb{R}\}$$

sowie

$$U' = \{f : \mathbb{R} \longrightarrow \mathbb{R} \mid f(-x) = f(x)\ \forall\, x \in \mathbb{R}\}.$$

Zeige, daß U und U' Unterräume von V sind mit $V = U \oplus U'$.

Aufgabe 2.48 (Erster Isomorphiesatz)
Sei V ein K–Vektorraum und $U, U' \leq V$. Zeige

$$U/(U \cap U') \cong (U + U')/U'.$$

Aufgabe 2.49 (Projektionen)
Es sei V ein K-Vektorraum. $\mathsf{f} \in \mathrm{End}_{\mathsf{K}}(\mathsf{V})$ heißt *Projektion*, falls $\mathsf{f}^2 = \mathsf{f}$ gilt. Zeige, die folgenden Aussagen sind äquivalent:

a. f ist eine Projektion,

b. $\mathrm{id}_{\mathsf{V}} - \mathsf{f}$ ist eine Projektion,

c. $\mathrm{Im}\big(\mathrm{id}_{\mathsf{V}} - \mathsf{f}\big) = \mathrm{Ker}(\mathsf{f})$,

d. $\mathrm{Ker}\big(\mathrm{id}_{\mathsf{V}} - \mathsf{f}\big) = \mathrm{Im}(\mathsf{f})$.

Zeige auch, sind obige Bedingungen erfüllt, so gilt zudem $\mathsf{V} = \mathrm{Ker}(\mathsf{f}) \oplus \mathrm{Im}(\mathsf{f})$.

§ 3 Basen von Vektorräumen

In diesem Abschnitt ist V stets ein K-Vektorraum.

Das wesentlichste Konzept im Zusammenhang mit Vektorräumen ist das der Basis. Mit Hilfe einer Basis können die Elemente eines Vektorraums effizient auf eindeutige Weise dargestellt werden. Wir führen in diesem Kapitel Basen als linear unabhängige Erzeugendensysteme ein.

A) Linear unabhängige Familien von Vektoren

Definition 3.1 (Familien)
Es seien I und X zwei Mengen.

a. Wir nennen ein Tupel der Form $F = (x_i)_{i\in I}$ mit $x_i \in X$ eine *Familie* von Elementen in X. Ist I endlich, so nennen wir die Familie *endlich* und setzen $|F| := |I|$.

b. Ist $F = (x_i)_{i\in I}$ eine Familie von Elementen in X, schreiben wir $x \in F$, wenn es ein $i \in I$ gibt mit $x = x_i$.

c. Sind $F = (x_i)_{i\in I}$ und $G = (y_i)_{i\in I}$ zwei Familien von Elementen in X, so schreiben wir $F = G$, falls $x_i = y_i$ für alle $i \in I$.

d. Ist $F = (x_i)_{i\in I}$ eine Familie von Elementen in X und $J \subseteq I$, so nennen wir $F' = (x_j)_{j\in J}$ eine *Teilfamilie* von F und F eine *Oberfamilie* von F', und wir schreiben $F' \subseteq F$. Ebenso schreiben wir $x \in F$, um auszudrücken, daß $x = x_i$ für ein $i \in I$ gilt.

e. Wir schreiben kurz $\operatorname{Lin}(F)$ für die lineare Hülle $\operatorname{Lin}(\{x_i \mid i \in I\})$ und nennen $\operatorname{Lin}(F)$ die *lineare Hülle* von F.

Beispiel 3.2
Ist $I = \{1, 2, 3\}$ und $X = \mathbb{R}^2$, so wird durch $x_1 = (1, 0)^t$, $x_2 = (1, 1)^t$, $x_3 = (1, 0)^t$ eine endliche Familie $(x_1, x_2, x_3) = \big((1, 0)^t, (1, 1)^t, (1, 0)^t\big)$ definiert. $(x_1, x_3) = \big((1, 0)^t, (1, 0)^t\big)$ ist eine Teilfamilie.

Bemerkung 3.3 (Familien von Vektoren)

a. In einer Familie können Elemente auch *mehrfach* auftreten, in einer Menge geht das nicht. Z.B. $F = \big((1, 0)^t, (1, 1)^t, (1, 0)^t\big)$.

b. Ist die Menge I geordnet, so ordnen wir die Mitglieder der Familie F in der gleichen Weise. Z.B. $\big((1, 0)^t, (1, 1)^t, (1, 0)^t\big) \neq \big((1, 0)^t, (1, 0)^t, (1, 1)^t\big)$.

c. In unseren Anwendungen wird die Menge I meist $\{1, \ldots, n\}$ für eine positive natürliche Zahl n sein, und wir ordnen die Elemente dann in der naheliegenden Weise.

d. Formal korrekt sollte man die Familie $F = (x_i)_{i\in I}$ als Abbildung $F : I \longrightarrow X : i \mapsto x_i$ angeben. Die Tupelschreibweise ist aber suggestiver als die Schreibweise als Abbildung.

Definition 3.4 (Lineare Unabhängigkeit)
Es sei V ein K-Vektorraum und F eine Familie von Vektoren in V.

a. Eine endliche Familie $(x_1, \ldots, x_n)$ von Vektoren in V heißt *linear unabhängig*, wenn aus

$$\lambda_1 \cdot x_1 + \ldots + \lambda_n \cdot x_n = 0$$

stets

$$\lambda_1 = \ldots = \lambda_n = 0$$

folgt, d.h. wenn nur die triviale Linearkombination der x_i Null ergibt.
Wir sagen dann oft einfach, die Vektoren $x_1, \ldots, x_n$ seien linear unabhängig.

b. Eine endliche Familie $(x_1, \ldots, x_n)$ von Vektoren in V heißt *linear abhängig*, wenn es Skalare $\lambda_1, \ldots, \lambda_n \in K$ gibt, so daß

$$\lambda_1 \cdot x_1 + \ldots + \lambda_n \cdot x_n = 0,$$

aber nicht alle λ_i sind Null, d.h. wenn eine nicht-triviale Linearkombination der x_i Null ergibt.
Wir nennen oft einfach die Vektoren $x_1, \ldots, x_n$ linear abhängig.

c. F heißt *linear abhängig*, wenn es eine endliche linear abhängige Teilfamilie gibt.

d. F heißt *linear unabhängig*, wenn jede endliche Teilfamilie linear unabhängig ist.

Beispiel 3.5 (Lineare Unabhängigkeit)

a. Die Einheitsvektoren $e_1, \ldots, e_n \in K^n$ sind linear unabhängig. Denn aus

$$(0, \ldots, 0)^t = \lambda_1 \cdot e_1 + \ldots + \lambda_n \cdot e_n = (\lambda_1, \ldots, \lambda_n)^t$$

folgt unmittelbar $\lambda_1 = \ldots = \lambda_n = 0$.

b. Die Familie $\big((1,0)^t, (0,1)^t, (1,1)^t\big)$ ist linear abhängig, da

$$(1,0)^t + (0,1)^t - (1,1)^t = (0,0)^t.$$

c. Wir betrachten die Folge $e_k = (\delta_{n,k})_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$, die als k-ten Eintrag eine Eins hat und ansonsten konstant Null ist. Dann ist die Familie $F = (e_k)_{k\in\mathbb{N}}$ linear unabhängig in $\mathbb{R}^{\mathbb{N}}$.
Um das zu sehen, betrachten wir die endliche Teilfamilie $(e_{k_1}, \ldots, e_{k_l})$ für $0 \leq k_1 < \ldots < k_l$. Dann folgt aus

$$\lambda_{k_1} \cdot e_{k_1} + \ldots + \lambda_{k_l} \cdot e_{k_l} = (0)_{n\in\mathbb{N}}$$

unmittelbar $\lambda_{k_1} = \ldots = \lambda_{k_l} = 0$, da die linke Folge als Folgenglied k_i den Wert λ_{k_i} hat. Also ist jede endliche Teilfamilie von F linear unabhängig, und somit ist auch F linear unabhängig.

Lemma 3.6 (Kriterien für lineare Abhängigkeit)
Es sei $F = (x_i)_{i\in I}$ *eine Familie von Vektoren im* K*-Vektorraum* V.

a. *Ist* $0 \in F$*, so ist* F *linear abhängig.*

b. *Gibt es ein* $i \neq j$ *mit* $x_i = x_j$*, so ist* F *linear abhängig.*

c. F *ist genau dann linear abhängig, wenn es ein* $x \in F$ *gibt, das Linearkombination anderer Vektoren in* F *ist.*

Beweis: Im ersten Fall ist $1 \cdot 0_V = 0_V$ eine nicht-triviale Linearkombination, im zweiten Fall ist $1 \cdot x_i - 1 \cdot x_j = 0_V$ eine solche. In jedem Fall ist F also linear abhängig, weil F eine endliche linear abhängige Teilfamilie enthält. Damit sind a. und b. gezeigt.

Ist F linear abhängig, so gibt es eine nicht-triviale Linearkombination

$$\sum_{j\in J} \lambda_j \cdot x_j = 0$$

mit $J \subseteq I$ endlich und nicht alle λ_j sind Null. Sei also $i \in J$ mit $\lambda_i \neq 0$, dann ist

$$x_i = \sum_{i\neq j\in J} -\frac{\lambda_j}{\lambda_i} \cdot x_j$$

Linearkombination anderer Vektoren in F.

Ist umgekehrt $x_i = \sum_{j\in J} \lambda_j \cdot x_j$ mit $J \subseteq I$ endlich und $i \in I \setminus J$, so ist

$$-x_i + \sum_{j\in J} \lambda_j \cdot x_j = 0$$

eine nicht-triviale Linearkombination, die Null ergibt. Mithin ist F linear abhängig.

□

Beispiel 3.7
In Beispiel 3.5 b. gilt

$$(1, 0)^t = -(0, 1)^t + (1, 1)^t,$$

woraus ebenfalls die lineare Abhängigkeit der Familie folgt.

Notation 3.8 (Linearkombination)
Sei $F = (x_i)_{i\in I}$ eine Familie von Vektoren in V und I sei nicht notwendigerweise endlich. Wir werden des öfteren

$$x = \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i \cdot x_i \tag{4}$$

schreiben, wenn wir sagen wollen, daß x eine Linearkombination von Vektoren in F ist. Formal korrekt müßte es lauten: es gibt eine endliche Teilfamilie $(x_j)_{j\in J}$ von F und Skalare $\lambda_j \in K$ für $j \in J$, so daß

$$x = \sum_{j\in J} \lambda_j \cdot x_j.$$

Wir interpretieren dies so, daß in (4) nur endlich viele der λ_i nicht Null sind, und daß somit die Summe auf der rechten Seite doch eine endliche Summe ist.

Mit dieser neuen Notation ist F genau dann linear unabhängig, wenn aus

$$\sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i \cdot x_i = 0$$

stets $\lambda_i = 0$ für alle $i \in I$ folgt; und analog ist F linear abhängig, wenn es eine Linearkombination

$$\sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i \cdot x_i = 0$$

gibt, bei der nicht alle λ_i Null sind.

Lemma 3.9 (Ergänzung linear unabhängiger Familien)
Ist $B = (x_i)_{i\in I}$ *eine linear unabhängige Familie in* V *mit* $\mathrm{Lin}\,(B) \subsetneqq V$*, so ist die Familie* $(x, x_i \mid i \in I)$ *für jedes* $x \in V \setminus \mathrm{Lin}\,(B)$ *linear unabhängig.*

Beweis: Seien dazu $\lambda, \lambda_i \in K$, $i \in I$, mit

$$\lambda x + \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i x_i = 0.$$

Wäre $\lambda \neq 0$, so wäre

$$x = \sum_{\substack{i\in I\\ \text{endlich}}} -\frac{\lambda_i}{\lambda} \cdot x_i \in \mathrm{Lin}\,(B)$$

im Widerspruch zur Wahl von x. Also ist $\lambda = 0$, und somit folgt aus

$$\sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i x_i = \lambda x + \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i x_i = 0$$

und der linearen Unabhängigkeit von B, daß auch alle anderen λ_i Null sind. Also ist $(x, x_i \mid i \in I)$ linear unabhängig. □

B) Erzeugendensysteme und Basen

Definition 3.10 (Erzeugendensystem und Basis)
Es sei V ein K-Vektorraum und F eine Familie von Vektoren in V.

a. F heißt ein *Erzeugendensystem* von V, wenn $V = \mathrm{Lin}\,(F)$, d.h. wenn jeder Vektor in V eine Linearkombination von Vektoren in F ist.

b. F heißt eine *Basis* von V, wenn F ein linear unabhängiges Erzeugendensystem von V ist.

c. V heißt *endlich erzeugt*, wenn V ein endliches Erzeugendensystem besitzt.

Beispiel 3.11 (Erzeugendensystem und Basis)

a. Die Familie $B = (e_1, \ldots, e_n)$ der Einheitsvektoren im K^n ist eine Basis des K^n, die wir auch die *Standardbasis* oder die *kanonische Basis* des K^n nennen. Denn nach Beispiel 3.5 ist B linear unabhängig und zudem ist ein beliebiger Vektor $x = (x_1, \ldots, x_n)^t \in K^n$ eine Linearkombination

$$x = x_1 \cdot e_1 + \ldots + x_n \cdot e_n$$

der Vektoren in B.

b. Analog sieht man, daß für $n, m \geq 1$ die Familie

$$\left(E_i^j \mid i = 1, \ldots, m; j = 1, \ldots, n\right),$$

wobei $E_i^j = (e_{lk})_{l=1,\ldots,m;k=1,\ldots,n}$ mit

$$e_{lk} = \delta_{il} \cdot \delta_{jk} = \begin{cases} 1, & \text{falls } l = i \text{ und } k = j, \\ 0, & \text{sonst} \end{cases}$$

die Matrix ist, die in Zeile i und Spalte j eine Eins als Eintrag hat und sonst nur Nullen, eine Basis des K-Vektorraums $\operatorname{Mat}(m \times n, K)$ ist.

c. Die Familie $\left((1,0)^t, (0,1)^t, (1,1)^t\right)$ in Beispiel 3.5 b. ist ein Erzeugendensystem von $\mathbb{R}^2$, aber keine Basis, da sie linear abhängig ist.

d. Die Familie $(e_k)_{k \in \mathbb{N}}$ im Vektorraum der Folgen $\mathbb{R}^{\mathbb{N}}$ aus Beispiel 3.5 c. ist *kein* Erzeugendensystem von $\mathbb{R}^{\mathbb{N}}$. Es scheint zwar, als gelte für eine beliebige Folge $(a_n)_{n \in \mathbb{N}}$

$$(a_n)_{n \in \mathbb{N}} = (a_0, a_1, a_2, \ldots) = \sum_{n \in \mathbb{N}} a_n \cdot e_n,$$

aber diese Summe ist *nicht* endlich und mithin keine zulässige Linearkombination! Die konstante Folge $(1)_{n \in \mathbb{N}}$ ist sicher keine endliche Linearkombination der e_k, da eine solche nur endlich viele Folgenglieder ungleich Null haben kann.

e. Die Familie $(1, i)$ ist eine Basis von $\mathbb{C}$ als $\mathbb{R}$-Vektorraum, da jede komplexe Zahl von der Gestalt $x + iy$ mit $x, y \in \mathbb{R}$ ist und da eine solche Zahl nur dann Null ist, wenn x und y beide Null sind.

f. Die Familie $B = (t^0, t^1, t^2, \ldots)$ ist eine Basis von $K[t]$.

Proposition 3.12 (Eindeutige Darstellbarkeit bezüglich einer Basis)
Eine Familie B von Vektoren in V ist genau dann eine Basis von V, wenn jeder Vektor in V in eindeutiger Weise als Linearkombination von Elementen in B geschrieben werden kann.

Beweis: Sei zunächst $B = (x_i)_{i \in I}$ eine Basis von V und $x \in V$. Nach Voraussetzung ist B ein Erzeugendensystem von V und mithin ist

$$x = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot x_i$$

eine Linearkombination von Vektoren in B. Ist nun

$$x = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i' \cdot x_i$$

eine zweite Linearkombination von Vektoren in B, die x ergibt, so ist

$$0 = x - x = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot x_i - \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i' \cdot x_i = \sum_{\substack{i \in I \\ \text{endlich}}} (\lambda_i - \lambda_i') \cdot x_i$$

eine Linearkombination von Vektoren in B, die Null ergibt. Da B linear unabhängig ist, muß dann aber stets

$$\lambda_i - \lambda_i' = 0$$

gelten. Die Darstellung ist also eindeutig.

Sie nun umgekehrt jeder Vektor x in V auf eindeutige Weise als Linearkombination der Vektoren in B darstellbar. Dann ist offenbar B ein Erzeugendensystem von V, und 0_V kann nur auf die triviale Weise als Linearkombination von Vektoren in B dargestellt werden, so daß B auch linear unabhängig ist. □

Beispiel 3.13
$B = \big((1,1)^t, (1,-1)^t\big)$ ist eine Basis von $\mathbb{R}^2$, da sich ein beliebiger Vektor $(\lambda_1, \lambda_2)^t \in \mathbb{R}^2$ in eindeutiger Weise als

$$(\lambda_1, \lambda_2)^t = \frac{\lambda_1 + \lambda_2}{2} \cdot (1,1)^t + \frac{\lambda_1 - \lambda_2}{2} \cdot (1,-1)^t$$

schreiben läßt, wie man leicht sieht.

Satz 3.14 (Existenz- und Eindeutigkeitssatz für lineare Abbildungen)
Es seien V und W zwei K-Vektorräume, $B = (x_i)_{i \in I}$ eine Basis von V und $F = (y_i)_{i \in I}$ eine Familie von Vektoren in W.

Dann gibt es genau eine K-lineare Abbildung $f : V \longrightarrow W$, so daß für alle $i \in I$

$$f(x_i) = y_i.$$

Insbesondere, zwei lineare Abbildungen sind gleich, sobald sie auf einer Basis übereinstimmen.

Beweis: Jeder Vektor $x \in V$ läßt sich nach Proposition 3.12 in eindeutiger Weise als Linearkombination

$$x = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot x_i$$

schreiben. Wir definieren die Abbildung f dann durch

$$f(x) := \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot y_i. \tag{5}$$

Wir wollen nun zeigen, daß $f : V \longrightarrow W$ dann K-linear ist. Seien dazu

$$x = \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i \cdot x_i, \quad x' = \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i' \cdot x_i \in V$$

und $\lambda, \lambda' \in K$ gegeben, dann gilt für die eindeutige Darstellung von $\lambda x + \lambda' x'$ offenbar

$$\lambda x + \lambda' x' = \sum_{\substack{i\in I\\ \text{endlich}}} (\lambda \cdot \lambda_i + \lambda' \cdot \lambda_i') \cdot x_i,$$

und mithin erhalten wir

$$\begin{aligned} f\big(\lambda x + \lambda' x'\big) &= f\left(\sum_{\substack{i\in I\\ \text{endlich}}} (\lambda \cdot \lambda_i + \lambda' \cdot \lambda_i') \cdot x_i\right) \\ &= \sum_{\substack{i\in I\\ \text{endlich}}} (\lambda \cdot \lambda_i + \lambda' \cdot \lambda_i') \cdot y_i \\ &= \lambda \cdot \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i \cdot y_i + \lambda' \cdot \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i' \cdot y_i \\ &= \lambda \cdot f(x) + \lambda' \cdot f(x'). \end{aligned}$$

Die Abbildung f ist also K-linear, und nach Definition gilt auch $f(x_i) = y_i$.

Es bleibt zu zeigen, daß es keine zweite K-lineare Abbildung geben kann, die diese Eigenschaft hat. Sei dazu $g : V \longrightarrow W$ eine K-lineare Abbildung mit $g(x_i) = y_i$ für alle $i \in I$. Ein beliebiges $x \in V$ läßt sich wieder schreiben als

$$x = \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i \cdot x_i$$

und dann gilt

$$f(x) = \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i \cdot f(x_i) = \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i \cdot y_i = \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i \cdot g(x_i) = g(x).$$

Mithin stimmt f mit g überein. □

Bemerkung 3.15 (Existenz- und Eindeutigkeitssatz für lineare Abbildungen)
Satz 3.14 besagt, daß man die Werte einer linearen Abbildung auf einer Basis beliebig vorschreiben kann. Egal welche Vektoren im Zielbereich man als Bilder wählt, es gibt eine und nur eine lineare Abbildung, die den Basiselementen genau diese Vektoren zuordnet!

Wegen der Formel für $f(x)$ in (5) sagt man auch, daß sich f aus der Vorschrift $f(x_i) = y_i$, $i \in I$, durch *lineare Fortsetzung* ergibt.

Beispiel 3.16
Setzen wir $x_1 = (1,1)^t$ und $x_2 = (1,-1)^t$, so ist $B = (x_1, x_2)$ eine Basis von $\mathbb{R}^2$.

Wählen wir nun zudem $y_1 = (1,1)^t$ und $y_2 = (3,1)^t$, so muß es genau eine $\mathbb{R}$-lineare Abbildung $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ geben mit

$$f\big((1,1)^t\big) = f(x_1) = y_1 = (1,1)^t \quad \text{und} \quad f\big((1,-1)^t\big) = f(x_2) = y_2 = (3,1)^t.$$

Diese besitzt die Abbildungsvorschrift

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 : (x,y)^t \mapsto (2x-y, x)^t.$$

Wir werden später sehen, wie man die Abbildungsvorschrift systematisch bestimmen kann.

Korollar 3.17 (Alle linearen Abbildungen $K^n \to K^m$ sind von der Form f_A.)
Jede lineare Abbildung $f : K^n \longrightarrow K^m$ *ist von der Form* $f = f_A$ *für eine eindeutig bestimmte Matrix* $A \in \mathrm{Mat}(m \times n, K)$.

Beweis: Ist $f : K^n \longrightarrow K^m$ eine lineare Abbildung, so setzen wir $a^i := f(e_i) \in K^m$ für $i = 1, \ldots, n$ und bilden eine Matrix A mit den a^i als Spaltenvektoren. Dann ist f_A eine lineare Abbildung, mit

$$f_A(e_i) = Ae_i = a^i = f(e_i),$$

so daß aus der Eindeutigkeitsaussage in 3.14 unmittelbar $f_A = f$ folgt. Die Eindeutigkeit der Matrix A folgt aus der Tatsache, daß A die Abbildung f_A eindeutig festlegt (siehe Bemerkung 1.7). □

Proposition 3.18 (Charakterisierung von Basen)
Für eine Familie B *von Vektoren in* V *sind die folgenden Aussagen gleichwertig:*

a. B *ist eine Basis von* V.

b. B *ist ein minimales Erzeugendensystem von* V.

c. B *ist eine maximale linear unabhängige Familie in* V.

Bemerkung 3.19
Ein Erzeugendensystem B von V heißt *minimal*, wenn keine echte Teilfamilie von B ein Erzeugendensystem ist. Dies heißt *nicht*, daß sie in jedem anderen Erzeugendensystem enthalten ist! Es gibt nicht *das* minimale Erzeugendensystem.

Eine linear unabhängige Familie B in V heißt *maximal*, wenn keine echte Oberfamilie linear unabhängig ist. Dies heißt *nicht*, daß sie jede andere linear unabhänige Familie enthält! Es gibt nicht *die* maximale linear unabhängige Familie.

Beweis von Proposition 3.18: Es sei $B = (x_i)_{i \in I}$.

a. ⇒ b.: Ist B eine Basis, so erzeugt B den Vektorraum V per definitionem. Ist $(x_j \mid j \in J)$ eine echte Teilfamilie von B und ist $i \in I \setminus J$, so gibt es wegen der linearen Unabhängigkeit von B keine Darstellung

$$x_i - \sum_{\substack{j \in J \\ \text{endlich}}} \lambda_j x_j = 0$$

also ist $x_i \notin \operatorname{Lin}(x_j \mid j \in J)$.

b. $\Rightarrow$ c.: Wir zeigen zunächst, daß B linear unabhängig ist. Angenommen, dies sei nicht der Fall, dann läßt sich nach Lemma 3.6 ein x_i als Linearkombination

$$x_i = \sum_{\substack{i \neq j \in I \\ \text{endlich}}} \lambda_j x_j$$

der übrigen Vektoren in B darstellen. Damit gilt dann aber

$$\operatorname{Lin}(x_j \mid j \in I \setminus \{i\}) = \operatorname{Lin}(x_j \mid j \in I) = V,$$

im Widerspruch zur Minimalität des Erzeugendensystems B.

Sei nun $(x_j \mid j \in J)$ mit $I \subsetneqq J$ eine echte Oberfamilie von B und $j \in J \setminus I$, so ist x_j eine Linearkombination

$$x_j = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i x_i$$

der Elemente in B, da B ein Erzeugendensystem ist. Folglich ist $(x_j \mid j \in J)$ linear abhängig nach Lemma 3.6.

c. $\Rightarrow$ a.: Da B linear unabhängig ist, bleibt zu zeigen, daß $\operatorname{Lin}(B) = V$. Gäbe es ein $x \in V \setminus \operatorname{Lin}(B)$, so wäre wegen Lemma 3.9 auch $(x, x_i \mid i \in I)$ linear unabhängig, im Widerspruch zur Maximalität von B.

$\square$

Beispiel 3.20

Kommen wir zu unserem Beispiel $B = \big((1,1)^t, (1,-1)^t\big)$ zurück. Da sich ein beliebiger Vektor $(\lambda_1, \lambda_2)^t \in \mathbb{R}^2$ als

$$(\lambda_1, \lambda_2)^t = \frac{\lambda_1 + \lambda_2}{2} \cdot (1,1)^t + \frac{\lambda_1 - \lambda_2}{2} \cdot (1,-1)^t$$

schreiben läßt, ist B ein Erzeugendensystem von $\mathbb{R}^2$, und offenbar kann man weder $(1,1)^t$ noch $(1,-1)^t$ weglassen. B ist also ein minimales Erzeugendensystem und mithin eine Basis von $\mathbb{R}^2$. Damit sparen wir uns, die Eindeutigkeit obiger Darstellung von $(\lambda_1, \lambda_2)^t$ zu zeigen, von der wir in Beispiel 3.13 nur gesagt haben, daß man sie leicht nachweisen könne!

C) Existenz von Basen

Da Basen für das Rechnen in Vektorräumen von großer Bedeutung sind, stellt sich unmittelbar die Frage nach ihrer Existenz in einem beliebigen Vektorraum. Für endlich erzeugte Vektorräume ergibt sich aus Lemma 3.9 unmittelbar, daß man aus einem endlichen Erzeugendensystem sukzessive eine maximale linear unabhängige Familie, sprich eine Basis, aufbauen kann (siehe Beweis von Satz 3.21). Wir werden uns damit noch mal ausführlich zu Beginn von Abschnitt 4 beschäftigen. Für nicht endlich erzeugte Vektorräume ist der Beweis der Existenz einer Basis ungleich schwerer. Man benötigt dazu das sogenannte *Zornsche Lemma*, eine Aussage, die zu den logischen Grundlagen der Mathematik gehört. Grob gesprochen gehört es zu den (im Rahmen einer formalen Mengenlehre) nicht aus anderen Axiomen herleitbaren Axiomen. Man kann aber zeigen, daß das Zornsche Lemma äquivalent zum *Wohlordnungssatz* und zum *Auswahlaxiom* ist, vgl. [**Moo82, Sze50**]. Ohne diese Axiome läßt sich der Existenzsatz über Basen nicht für beliebige Vektorräume beweisen. Wir beweisen den folgenden Satz deshalb nur für endliche Erzeugendensysteme.

Satz 3.21 (Basisexistenzsatz)
Sei E *ein Erzeugendensystem des* K*-Vektorraums* V *und sei* F *eine linear unabhängige Teilfamilie von* E*, dann gibt es eine Basis* B*, die Teilfamilie von* E *ist und* F *als Teilfamilie enthält, d.h.* $F \subseteq B \subseteq E$*.*

Insbesondere gelten die folgenden Aussagen:

a. *Jede linear unabhängige Familie in* V *kann zu einer Basis ergänzt werden.*
b. *Jedes Erzeugendensystem von* V *enthält eine Basis.*
c. *Jeder Vektorraum besitzt eine Basis.*

Beweis für den Fall eines endlichen Erzeugendensystems E**:** Ist F schon ein Erzeugendensystem von V, so ist F eine Basis von V und wir sind mit $B := F$ fertig. Ist F noch kein Erzeugendensystem, dann muß es einen Vektor x in E geben, der in $V \setminus \mathrm{Lin}(B)$ liegt, da E ja ein Erzeugendensystem von V ist. Nach Lemma 3.9 ist $F \cup (x)$ dann linear unabhängig. Auf diese Weise können wir fortfahren, solange die durch Erweiterung aus F konstruierte linear unabhängige Familie noch kein Erzeugendensystem von V ist. Da E nur endlich viele Elemente enthält, muß das Verfahren nach endlich vielen Schritten abbrechen, d. h. nach endlich vielen Schritten ist die aus F konstruierte linear unabhängige Familie auch ein Erzeugendensystem, also eine Basis, die F enthält und in E enthalten ist. □

Beispiel 3.22
Das Erzeugendensystem $E = \big((1,0)^t, (1,1)^t, (0,1)^t\big)$ von $\mathbb{R}^2$ enthält die kanonische Basis, und die linear unabhängige Familie $F = \big((1,1)^t\big)$ kann zur Basis $B = \big((1,1)^t, (1,-1)^t\big)$ von $\mathbb{R}^2$ ergänzt werden.

Korollar 3.23 (Existenz von Komplementen)
Jeder Unterraum U *von* V *besitzt ein direktes Komplement.*

Beweis: Wähle eine Basis B von U und ergänze sie durch eine linear unabhängige Familie B' zu einer Basis $B \cup B'$ von V gemäß dem Basisergänzungssatz 3.21. Dann ist $U' := \mathrm{Lin}\,(B')$ ein Komplement von U, denn

$$U + U' = \mathrm{Lin}\,(U \cup U') \supseteq \mathrm{Lin}\,(B \cup B') = V$$

und aus

$$x = \sum_{\substack{y \in B \\ \text{endlich}}} \lambda_y \cdot y = \sum_{\substack{z \in B' \\ \text{endlich}}} \lambda_z \cdot z \in U \cap U'$$

folgt wegen der linearen Unabhängigkeit von $B \cup B'$ mit

$$\sum_{\substack{y \in B \\ \text{endlich}}} \lambda_y \cdot y + \sum_{\substack{z \in B' \\ \text{endlich}}} -\lambda_z \cdot z = 0,$$

daß alle λ_y und λ_z Null sein müssen, so daß auch $U \cap U' = \{0\}$. $\square$

Beispiel 3.24 ($\mathbb{R}$ als $\mathbb{Q}$-Vektorraum)
Auch wenn jeder Vektorraum eine Basis besitzt, kann nicht notwendigerweise für jeden Vektorraum eine Basis angegeben werden. $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum ist ein gutes Beispiel für einen Vektorraum, bei dem man keine Basis angeben kann.

Behauptung: Eine Basis von $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum ist überabzählbar.

Hierzu argumentieren wir wie folgt, wobei wir eine Menge *höchstens abzählbar* nennen, wenn sie endlich oder abzählbar ist.

a. $\mathbb{R}$ *ist überabzählbar* nach Proposition A5.6.

b. *Die abzählbare Vereinigung höchstens abzählbarer Mengen ist wieder höchstens abzählbar.* Seien dazu $M_i = \{a_{i0}, a_{i1}, a_{i2}, \ldots\}$, $i \in \mathbb{N}$, (ohne Einschränkung) abzählbare Mengen, dann schreiben wir sie wie folgt auf:

$$\begin{array}{lcccccccccccc}
M_0 : & a_{00} & & a_{01} & \rightarrow & a_{02} & & a_{03} & \rightarrow & a_{04} & & \ldots \\
 & \downarrow & \nearrow & & \swarrow & & \nearrow & & \swarrow & & & \\
M_1 : & a_{10} & & a_{11} & & a_{12} & & a_{13} & & a_{14} & & \ldots \\
 & & \swarrow & & \nearrow & & \swarrow & & & & & \\
M_2 : & a_{20} & & a_{21} & & a_{22} & & a_{23} & & a_{24} & & \ldots \\
 & \downarrow & \nearrow & & \swarrow & & & & & & & \\
M_3 : & a_{30} & & a_{31} & & a_{32} & & a_{33} & & a_{34} & & \ldots \\
\vdots & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots
\end{array}$$

Abzählen der Elemente wie angedeutet, wobei man Elemente, die mehrfach vorkommen, nur beim ersten Mal berücksichtigt, liefert eine Bijektion von $\mathbb{N} \rightarrow \bigcup_{i \in \mathbb{N}} M_i$, mithin ist die Vereinigung abzählbar.

c. Es gilt also $\mathbb{Z} = \mathbb{N} \cup (-\mathbb{N})$ und

$$\mathbb{Q} = \bigcup_{0 \neq q \in \mathbb{N}} \tfrac{1}{q}\mathbb{Z} = \bigcup_{0 \neq q \in \mathbb{N}} \left\{ \tfrac{p}{q} \mid p \in \mathbb{Z} \right\}$$

sind abzählbar.

d. *Das kartesische Produkt zweier höchstens abzählbarer Mengen ist wieder höchstens abzählbar.* Seien dazu M und N zwei höchstens abzählbare Mengen, dann gilt

$$M \times N = \bigcup_{m \in M} \{m\} \times N,$$

wobei $N \rightarrow \{m\} \times N : n \mapsto (m, n)$ eine Bijektion ist, $\{m\} \times N$ also höchstens abzählbar ist.

e. *Ein Vektorraum* V *über einem höchstens abzählbaren Körper* K *mit höchstens abzählbarer Basis ist höchstens abzählbar.* Sei dazu (ohne Einschränkung) $B = (x_i \mid i \in \mathbb{N})$, eine abzählbare Basis von V. Für $n \in \mathbb{N}$ setzen wir

$$V_n := \operatorname{Lin}(x_1, \ldots, x_n).$$

Dann gilt $V_n \cong K^n$, also ist V_n nach d. mit Induktion über n abzählbar. Aber dann ist $V = \bigcup_{n \in \mathbb{N}} V_n$ nach b. abzählbar.

f. Da $\mathbb{Q}$ abzählbar ist, $\mathbb{R}$ aber überabzählbar, folgt aus e. die Behauptung.

Bemerkung 3.25 (Ringe und Moduln)
In diesem Abschnitt haben wir in vielen Beweisen verwendet, daß man durch Körperelemente teilen darf, sobald sie nicht Null sind. Deshalb gelten viele Aussagen nicht mehr im allgemeinen für Moduln über Ringen. Die Definitionen lassen sich aber dennoch in der gleichen Weise geben, und wir wollen hier zusammenstellen, welche Aussagen für Moduln über Ringen letztlich wahr bleiben.

Die Definitionen 3.1, 3.4, 3.8 und 3.10 können für Moduln über Ringen in der gleichen Weise gegeben werden. Beispiele 3.5 und 3.11 bleiben dann ebenso richtig wie die Aussagen in Lemma 3.6 a. und c. und in den wichtigen Sätzen 3.12, 3.14 und 3.17.

Die Aussagen in Lemma 3.6 und 3.9, Proposition 3.18 und Satz 3.21 gelten für Moduln über Ringen im allgemeinen *nicht* mehr. In ihren Beweisen wird durch Skalare geteilt, von denen nur bekannt ist, daß sie nicht Null sind.

Aufgaben

Aufgabe 3.26
Welche der folgenden Familien sind linear unabhängig / Erzeugendensysteme / Basen von $\mathbb{R}^2$?

a. $\big((1,0)^t, (0,1)^t, (1,1)^t\big)$.
b. $\big((1,1)^t, (2,2)^t\big)$.
c. $\big((1,3)^t\big)$.
d. $\big((1,1)^t, (1,-2)^t\big)$.
e. $\big((1,1)^t, (0,0)^t\big)$.
f. $\big((1,1)^t, (0,0)^t, (1,-2)^t\big)$.
g. $\big((1,2)^t, (2,1)^t\big)$.

Aufgabe 3.27
Es sei V ein K-Vektorraum, $U \subset V$ ein Unterraum, $0 \neq x \in U$ und $y \in V \setminus U$. Zeige, daß (x, y) linear unabhängig ist.

Aufgabe 3.28
Ist $f : V \longrightarrow W$ eine K-lineare Abbildung, F eine Familie von Vektoren in V, so ist

$$f\big(\mathrm{Lin}\,(F)\big) = \mathrm{Lin}\,(f(x) \mid x \in F).$$

Aufgabe 3.29
Es seien V und W zwei K-Vektorräume, $f : V \longrightarrow W$ eine K-lineare Abbildung und B eine Basis von V.

a. Genau dann ist f surjektiv, wenn $f(B)$ ein Erzeugendensystem von W ist.

b. Genau dann ist f injektiv, wenn $f(B)$ linear unabhängig ist.

c. Genau dann ist f bijektiv, wenn $f(B)$ eine Basis von W ist.

Aufgabe 3.30
Es seien $U_1, \ldots, U_k$ Unterräume eines K-Vektorraums V mit Basen $B_1, \ldots, B_k$. Zeige, genau dann ist $V = U_1 \oplus \ldots \oplus U_k$ die direkte Summe der U_i, wenn $B = B_1 \cup \ldots \cup B_k$ eine Basis von V ist.

Aufgabe 3.31
Sei V ein $\mathbb{C}$–Vektorraum, dann ist V offensichtlich auch ein $\mathbb{R}$–Vektorraum, und seien $x_1, \ldots, x_n \in V$. Zeige, daß $(x_1, \ldots, x_n)$ genau dann linear unabhängig über $\mathbb{C}$ ist, wenn $(x_1, ix_1, \ldots, x_n, ix_n)$ linear unabhängig über $\mathbb{R}$ ist.

Aufgabe 3.32
Es sei $(V, +, \cdot)$ ein K-Vektorraum, und $x_1, \ldots, x_n \in V$ seien linear abhängige Vektoren mit der Eigenschaft, daß je $n-1$ der Vektoren linear unabhängig sind. Zeige:

a. Es gibt $\lambda_1, \ldots, \lambda_n \in K \setminus \{0\}$ mit der Eigenschaft

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

b. Gilt für $\mu_1, \ldots, \mu_n \in K$ ebenfalls $\sum_{i=1}^{n} \mu_i x_i = 0$, so gibt es ein $\nu \in K$ mit $\mu_i = \lambda_i \cdot \nu$ für alle $i = 1, \ldots, n$.

Aufgabe 3.33
Es sei $A \in \mathrm{Mat}_n(K)$.

a. Genau dann ist f_A bijektiv, wenn $A \in \mathrm{Gl}_n(K)$.

b. Ist $A \in \mathrm{Gl}_n(K)$, so gilt $(f_A)^{-1} = f_{A^{-1}}$.

Aufgabe 3.34 (Im Lemma von Zorn reichen abzählbare Ketten nicht!)
Finde ein Beispiel für eine teilgeordnete Menge $(M, \leq)$, so daß jede abzählbare Kette

$$K_1 \leq K_2 \leq K_3 \leq \ldots$$

von Elementen $K_i \in M$ eine obere Schranke in M besitzt, daß aber M selbst kein maximales Element hat.

§ 4 Endlich-dimensionale Vektorräume

Wir betrachten jetzt *endlich erzeugte Vektorräume* V, d. h. Vektorräume, die ein endliches Erzeugendensystem besitzen. Nach Satz 3.21 besitzt V dann auch eine endliche Basis. Für solche Vektorräume kann man die Sätze des vorigen Abschnitts teilweise verschärfen und vor allem kann man in diesen Vektorräumen mit Hilfe von Basen und Matrizen effizient rechnen.

> In diesem Abschnitt ist V stets ein endlich erzeugter K-Vektorraum.

A) Austauschsatz von Steinitz

Lemma 4.1 (Austauschlemma)
Sei $B = (x_1, \ldots, x_n)$ *eine Basis von* V, $y = \sum_{i=1}^n \lambda_i x_i$ *und* $\lambda_j \neq 0$ *für ein* j.
Dann ist $(x_1, \ldots, x_{j-1}, y, x_{j+1}, \ldots, x_n)$ *eine Basis von* V, *d.h. man kann in der Basis* B *den Vektor* x_j *gegen* y *austauschen.*

Beweis: Wegen $\lambda_j \neq 0$ gilt

$$x_j = \frac{1}{\lambda_j} \cdot y - \sum_{i \neq j} \frac{\lambda_i}{\lambda_j} \cdot x_i,$$

und somit

$$V = \operatorname{Lin}(x_1, \ldots, x_n) = \operatorname{Lin}(y, x_1, \ldots, x_n) = \operatorname{Lin}(x_1, \ldots, x_{j-1}, y, x_{j+1}, \ldots, x_n).$$

Bleibt also zu zeigen, daß $(x_1, \ldots, x_{j-1}, y, x_{j+1}, \ldots, x_n)$ linear unabhängig ist. Seien dazu $\mu_i \in K$, $i = 1, \ldots, n$, gegeben mit

$$\begin{aligned} 0 = \mu_j y + \sum_{i \neq j} \mu_i x_i &= \sum_{i=1}^n \mu_j \lambda_i x_i + \sum_{i \neq j} \mu_i x_i \\ &= \mu_j \lambda_j x_j + \sum_{i \neq j} (\mu_j \lambda_i + \mu_i) x_i. \end{aligned}$$

Dann folgt aus der linearen Unabhängigkeit von $x_1, \ldots, x_n$

$$\mu_j \lambda_j = 0 \quad \text{und} \quad \mu_i = -\mu_j \lambda_i, \text{ für } i \neq j.$$

Wegen $\lambda_j \neq 0$, ist also $\mu_j = 0$ und damit auch

$$\mu_i = 0 \quad \text{für } i \neq j.$$

Damit ist die lineare Unabhängigkeit von $(x_1, \ldots, x_{j-1}, y, x_{j+1}, \ldots, x_n)$ gezeigt. □

Beispiel 4.2
Ist zum Beispiel $E = (e_1, \ldots, e_n)$ die kanonische Basis des K^n und $x = (\lambda_1, \ldots, \lambda_n)^t \in K^n$ mit $\lambda_j \neq 0$, so können wir e_j gegen x austauschen und erhalten wieder eine Basis.

Konkret kann man in der Basis $E = (e_1, e_2, e_3)$ von $\mathbb{R}^3$ den Vektor $(1, 2, 0)^t$ gegen e_1 oder e_2 austauschen, nicht aber gegen e_3.

Das Austauschlemma wird benutzt, um den wichtigen Steinitzschen Austauschsatz zu beweisen.

Satz 4.3 (Austauschsatz von Steinitz)
Sei $(x_1, \ldots, x_n)$ *eine Basis von* V *und* $(y_1, \ldots, y_r)$ *sei linear unabhängig in* V. *Dann lassen sich die* $x_1, \ldots, x_n$ *so umnumerieren, daß* $(y_1, \ldots, y_r, x_{r+1}, \ldots, x_n)$ *eine Basis von* V *ist. Insbesondere gilt* $r \leq n$.

Beweis von Satz 4.3: Wir führen den Beweis mittels Induktion über r.

Für $r = 0$ ist die Behauptung offensichtlich richtig. Nehmen wir also an, daß $r > 0$ und daß die Behauptung bereits richtig ist für $r-1$. D. h. nach evt. Umnumerieren ist $(y_1, \ldots, y_{r-1}, x_r, \ldots, x_n)$ eine Basis von V. Dann besitzt y_r eine Darstellung der Form

$$y_r = \lambda_1 y_1 + \ldots + \lambda_{r-1} y_{r-1} + \lambda_r x_r + \ldots + \lambda_n x_n,$$

mit $\lambda_i \in K$. Angenommen, $\lambda_r = \ldots = \lambda_n = 0$, dann wäre $(y_1, \ldots, y_r)$ linear abhängig, im Widerspruch zur Voraussetzung. Also gibt es ein $j \in \{r, \ldots, n\}$ mit $\lambda_j \neq 0$. Durch Umnumerieren können wir annehmen, daß $j = r$ gilt. Dann können wir aber nach dem Austauschlemma 4.1 y_r gegen x_r austauschen, und die Behauptung ist bewiesen. □

Bemerkung 4.4

a. Der Austauschsatz von Steinitz besagt also, daß man - nach eventuellem Umnumerieren - die linear unabhängigen Vektoren $x_1, \ldots, x_r$ durch $y_1, \ldots, y_r$ ersetzen kann.

b. Im Austauschsatz tauschen wir nacheinander x_{i_1} durch y_1, x_{i_2} durch y_2, etc. und schließlich x_{i_r} durch y_r für geeignete $i_1, \ldots, i_r$ aus. Im j-ten Schritt wissen wir, daß wir eine Darstellung

$$y_j = \sum_{l=1}^{j-1} \lambda_{i_l} y_l + \sum_{l \notin \{i_1, \ldots, i_{j-1}\}} \lambda_l x_l$$

haben mit $\lambda_l \neq 0$ für ein $l \notin \{i_1, \ldots, i_{j-1}\}$, und setzen wir dann $i_j := l$, so können wir x_{i_j} durch y_j ersetzen.
Wie wir eine solche Darstellung von y_j mit Hilfe des Gauß'schen Algorithmus berechnen können, werden wir später sehen. Damit haben wir dann ein konstruktives Verfahren für die Anwendung des Steinitzschen Austauschsatzes.

B) Die Dimension eines endlich-erzeugten Vektorraums

Als Folgerung des Steinitzschen Austauschsatzes erhalten wir den folgenden Satz.

Satz 4.5 (Alle Basen sind gleichmächtig.)

a. *Ist* V *endlich erzeugt, so ist jede Basis von* V *endlich und alle Basen haben gleich viele Elemente.*

b. *Ist* V *nicht endlich erzeugt, so hat jede Basis unendlich viele Elemente.*

Beweis: a. Nach Voraussetzung besitzt V ein endliches Erzeugendensystem E und nach Satz 3.21 folgt dann auch, daß V eine endliche Basis $B = (x_1, \ldots, x_n)$ besitzt. Dabei können wir o. E. annehmen, daß n die minimale Mächtigkeit einer Basis ist. Sei nun B' eine weitere Basis von V. Angenommen, $|B'| > n$. Dann gibt es eine linear unabhängige Teilfamilie $(y_1, \ldots, y_{n+1})$ in B', im Widerspruch zum Austauschsatz von Steinitz, der verlangt $n + 1 \leq n$.

b. Dies ist offensichtlich, da jede Basis V erzeugt.

□

Satz 4.5 rechtfertigt die folgende Definition.

Definition 4.6 (Dimension eines Vektorraums)
Für einen (nicht notwendig endlich erzeugten) K-Vektorraum V definieren wir die *Dimension* von V durch

$$\dim_K(V) := \begin{cases} n, & \text{falls } V \text{ eine Basis mit } n < \infty \text{ Elementen besitzt,} \\ \infty, & \text{falls } V \text{ nicht endlich erzeugt ist.} \end{cases}$$

Ist $\dim_K(V) < \infty$, so nennen wir V einen *endlich-dimensionalen* K-Vektorraum.

Aus Satz 4.5 und Definition 4.6 folgt unmittelbar das folgende Korollar.

Korollar 4.7
Sei $\dim_K(V) = n$*,* E *ein Erzeugendensystem von* V *und* F *linear unabhängig in* V*. Dann gelten*

$$|E| \geq n \quad \textit{und} \quad |F| \leq n.$$

Zudem gilt die Gleichheit genau dann, wenn die jeweilige Familie eine Basis ist.

Beweis: Nach Satz 3.21 ist F in einer Basis von V enthalten und E enthält eine Basis von V. Die Ungleichungen folgen dann aus Satz 4.5, und derselbe Satz liefert Gleichheit, wenn die Familien Basen sind. Gilt umgekehrt die Gleichheit, so muß E bzw. F ein minimales Erzeugendensystem bzw. eine maximal linear unabhängige Familie sein und somit nach Proposition 3.18 eine Basis. □

Beispiel 4.8

a. Es gilt:

$$\dim_K(V) = 0 \;\Leftrightarrow\; V = \operatorname{Lin}(\emptyset) \;\Leftrightarrow\; V = \{0\}.$$

b. $\dim_K\left(K^n\right) = n$, da die kanonische Basis $E = (e_1, \ldots, e_n)$ genau n Elemente enthält.

c. $\dim_{\mathbb{Q}}(\mathbb{Q}) = \dim_{\mathbb{R}}(\mathbb{R}) = \dim_{\mathbb{C}}(\mathbb{C}) = 1$, aber $\dim_{\mathbb{Q}}(\mathbb{R}) = \infty$ und $\dim_{\mathbb{R}}(\mathbb{C}) = 2$. Für letzteres beachte man, daß die Familie $(1, i)$ eine $\mathbb{R}$-Basis von $\mathbb{C}$ ist.

d. Die Dimension des Vektorraums P_n der Polynome vom Grad höchstens n ist $\dim_K(P_n) = n + 1$, da $B = (t^0, t^1, \ldots, t^n)$ eine Basis ist.

Satz 4.9 (Karten eines Vektorraums)
Es sei $B = (x_1, \ldots, x_n)$ *eine Basis von* V *und* $E = (e_1, \ldots, e_n)$ *die kanonische Basis des* K^n*. Dann bestimmt* B *einen Isomorphismus*

$$\phi_B : V \to K^n : x_i \mapsto e_i, \text{ für } i = 1, \ldots, n,$$

durch lineare Fortsetzung*. Man nennt* ϕ_B *die* Karte *von* V *zur Basis* B.

Beweis: Nach Satz 3.14 bestimmen die Zuordnungen

$$x_i \mapsto e_i, \; i = 1, \ldots, n, \quad \text{und} \quad e_i \mapsto x_i, \; i = 1, \ldots, n,$$

zwei lineare Abbildungen $\phi_B : V \to K^n$ und $\phi^B : K^n \to V$. Es bleibt zu zeigen, daß

$$\phi_B \circ \phi^B = \mathrm{id}_{K^n} \quad \text{und} \quad \phi^B \circ \phi_B = \mathrm{id}_V.$$

Dazu reicht es wegen Satz 3.14, daß die beiden Seiten jeweils auf einer Basis übereinstimmen, und das tun sie offenbar. □

Insbesondere haben wir das folgende Korollar gezeigt.

Korollar 4.10
Ist $\dim_K(V) = n$*, so gilt* $V \cong K^n$.

Korollar 4.11 (Die Dimension ist die einzige Invariante eines Vektorraums.)
Für zwei endlich-dimensionale K*-Vektorräume* V *und* W *sind gleichwertig:*

a. $V \cong W$.

b. $\dim_K(V) = \dim_K(W)$.

Beweis: Es seien $n = \dim_K(V)$ und $m = \dim_K(W)$.

Ist $f : V \to W$ ein Isomorphismus, so überführt er laut Aufgabe 3.29 eine Basis von V, die n Elemente enthält, in eine Basis von W, die m Elemente enthält. Mithin gilt $n = m$.

Ist umgekehrt $n = m$, so gibt es nach Korollar 4.10 Isomorphismen $f : V \to K^n$ und $g : K^n \to W$. Dann ist $g \circ f : V \to W$ ebenfalls ein Isomorphismus. □

Beispiel 4.12
Die Abbildungen $\sin, \cos \in \mathbb{R}^{\mathbb{R}}$ sind linear unabhängig, da aus

$$\lambda \cdot \sin + \mu \cdot \cos = 0$$

insbesondere

$$0 = \lambda \cdot \sin(0) + \mu \cdot \cos(0) = \mu$$

und

$$0 = \lambda \cdot \sin\left(\frac{\pi}{2}\right) + \mu \cdot \cos\left(\frac{\pi}{2}\right) = \lambda$$

folgt. Aber dann hat der Vektorraum $\mathrm{Lin}(\sin, \cos)$ die Dimension zwei, da $(\sin, \cos)$ eine Basis ist, und deshalb gilt

$$\mathrm{Lin}(\sin, \cos) \cong \mathbb{R}^2.$$

C) Dimensionsformeln

Lemma 4.13
Ist $\dim_K(V) < \infty$ *und* $U \subseteq V$ *ein Unterraum. Dann gelten*

$$\dim_K(U) \leq \dim_K(V),$$

und

$$U = V \iff \dim_K(U) = \dim_K(V).$$

Beweis: Ist U ein Unterraum, so kann eine Basis B von U zu einer Basis B' von V ergänzt werden, so daß $\dim_K(U) = |B| \leq |B'| = \dim_K(V)$ gelten muß.

Ist $U = V$, so ist offenbar auch $\dim_K(U) = \dim_K(V)$. Gilt umgekehrt $\dim_K(U) = \dim_K(V)$ und ist $B = (x_1, \ldots, x_n)$ eine Basis von U, so können wir sie zu einer Basis B' von V ergänzen nach dem Basisergänzungssatz 3.21. Wegen $B \subseteq B'$ und $|B| = n = |B'|$ folgt dann aber notwendigerweise $B = B'$, und somit $U = \text{Lin}(B) = \text{Lin}(B') = V$. □

Satz 4.14 (Dimensionsformel für Unterräume)
Ist $\dim_K(V) < \infty$ *und sind* U *und* U' *Unterräume von* V*, dann gilt:*

$$\dim_K\left(U + U'\right) = \dim_K(U) + \dim_K\left(U'\right) - \dim_K\left(U \cap U'\right).$$

Beweis: Wir beweisen mehr, nämlich wie wir geeignete Basen von U, U' und $U \cap U'$ wählen können. Sei $B_{U\cap U'} := (x_1, \ldots, x_r)$ eine Basis von $U \cap U'$. Wir ergänzen $B_{U\cap U'}$ zu einer Basis $B_U := (x_1, \ldots, x_r, y_1, \ldots, y_s)$ von U, und zu einer Basis $B_{U'} := (x_1, \ldots, x_r, z_1, \ldots, z_t)$ von U'. Das geht nach dem Basisergänzungssatz 3.21.

Behauptung: $B_{U+U'} := (x_1, \ldots, x_r, y_1, \ldots, y_s, z_1, \ldots, z_t)$ ist Basis von $U + U'$.

Dazu zeigen wir zunächst, daß jedes Element von $U + U'$ eine Linearkombination von Elementen aus $B_{U+U'}$ ist. Sei also $x + x' \in U + U'$ mit $x \in U$ und $x' \in U'$. Dann gilt:

$$x = \sum_{i=1}^{r} \lambda_i x_i + \sum_{j=1}^{s} \mu_j y_j \quad \text{und} \quad x' = \sum_{i=1}^{r} \lambda_i' x_i + \sum_{k=1}^{t} \mu_k' z_k,$$

mit $\lambda_i, \lambda_i', \mu_j, \mu_k' \in K$, $i = 1, \ldots, r$, $j = 1, \ldots, s$, $k = 1, \ldots, t$. Daraus folgt:

$$x + x' = \sum_{i=1}^{r} \left(\lambda_i + \lambda_i'\right) x_i + \sum_{j=1}^{s} \mu_j y_j + \sum_{k=1}^{t} \mu_k' z_k \in \text{Lin}\left(B_{U+U'}\right).$$

Dann müssen wir noch zeigen, daß $B_{U+U'}$ linear unabhängig ist. Sei dazu

$$\sum_{i=1}^{r} \lambda_i x_i + \sum_{j=1}^{s} \mu_j y_j + \sum_{k=1}^{t} \nu_k z_k = 0 \tag{6}$$

eine Linearkombination der Null. Dann ist

$$\underbrace{\sum_{i=1}^{r} \lambda_i x_i + \sum_{j=1}^{s} \mu_j y_j}_{\in U} = \underbrace{\sum_{k=1}^{t} -\nu_k z_k}_{\in U'} \in U \cap U'.$$

Da $B_{U\cap U'}$ eine Basis von $U \cap U'$ ist, gibt es also λ_i', so daß

$$\sum_{i=1}^{r} \lambda_i x_i + \sum_{j=1}^{s} \mu_j y_j = \sum_{i=1}^{r} \lambda_i' x_i = \sum_{i=1}^{r} \lambda_i' x_i + \sum_{j=1}^{s} 0 \cdot y_j$$

gilt. Da B_U linear unabhängig ist, ergibt ein Koeffizientenvergleich auf beiden Seiten insbesondere $\mu_j = 0$ für alle $j = 1, \ldots, s$. Damit erhalten wir aus (6) dann

$$\sum_{i=1}^{r} \lambda_i x_i + \sum_{k=1}^{t} \nu_k z_k = 0,$$

und da $B_{U'}$ linear unabhängig ist, müssen notwendigerweise alle λ_i und ν_k Null sein. Damit haben wir dann auch gezeigt, daß $B_{U+U'}$ linear unabhängig ist.

Aus der Behauptung folgt,

$$\dim_K \big(U+U'\big) = r+s+t = (r+s)+(r+t)-r = \dim_K(U)+\dim_K \big(U'\big)-\dim_K \big(U\cap U'\big).$$

□

Beispiel 4.15
Für die Unterräume $U = \operatorname{Lin}\left((1,0,0)^t, (1,1,1)^t\right)$ und $U' = \operatorname{Lin}\left((1,1,1)^t, (0,0,1)^t\right)$ von $\mathbb{R}^3$ sieht man leicht, daß

$$U \cap U' = \operatorname{Lin}\big((1,1,1)^t\big)$$

ein Vektorraum von Dimension eins ist, während U und U' jeweils Dimension zwei haben, da die angegebenen Erzeugendensysteme auch linear unabhängig sind. Mithin erhalten wir

$$\dim_{\mathbb{R}}(U + U') = 2 + 2 - 1 = 3 = \dim_{\mathbb{R}}(\mathbb{R}^3),$$

so daß

$$U + U' = \mathbb{R}^3$$

gelten muß. Da zudem

$$U + U' = \operatorname{Lin}\big((1,0,0)^t, (1,1,1)^t, (0,0,1)^t\big)$$

gilt, sehen wir, daß dieses Erzeugendensystem eine Basis des $\mathbb{R}^3$ ist.

Korollar 4.16 (Dimensionsformel für Komplemente)
Ist $\dim_K(V) < \infty$*, dann sind für Unterräume* U *und* U' *von* V *die folgenden Aussagen äquivalent:*

a. $V = U \oplus U'$.

b. $V = U + U'$ *und* $U \cap U' = \{0\}$.

c. $V = U + U'$ *und* $\dim_K(V) = \dim_K(U) + \dim_K\left(U'\right)$.

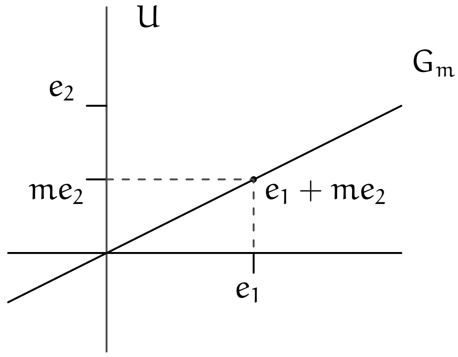
d. $U \cap U' = \{0\}$ *und* $\dim_K(V) = \dim_K(U) + \dim_K\left(U'\right)$.

Beweis: Dies ist eine direkte Folgerung aus Lemma 2.17, Lemma 4.13 und Satz 4.14. □

Beispiel 4.17
Wir erinnern uns an Beispiel 2.36. Dort haben wir in $V = \mathbb{R}^2$ den Unterraum $U = \operatorname{Lin}(e_2)$, die y-Achse, betrachtet und gezeigt, daß jede Ursprungsgerade mit Steigung m

$$G_m := \operatorname{Lin}(e_1 + me_2)$$

ein Komplement von U ist. Dies können wir nun mit weniger Aufwand begründen, denn die beiden Geraden schneiden sich offenbar nur im Ursprung und ihre Dimensionen addieren sich zu zwei.



Korollar 4.18 (Dimensionsformel für Faktorräume)
Es sei V ein endlich-dimensionaler K-Vektorraum und U ein Unterraum von V. Dann gilt

$$\dim_K(V/U) = \dim_K(V) - \dim_K(U).$$

Beweis: Nach Korollar 3.23 besitzt U ein Komplement U', und nach Proposition 2.37 gilt $U' \cong V/U$. Aus Korollar 4.11 und Korollar 4.16 folgt dann

$$\dim_K(V) = \dim_K(U) + \dim_K(U') = \dim_K(U) + \dim_K(V/U).$$

□

Bemerkung 4.19 (Basis von V/U)
Es sei V ein K-Vektorraum und U ein Unterraum von V mit Basis $(x_1, \ldots, x_r)$. Dann sind die folgenden Aussagen für $y_1, \ldots, y_s \in V$ gleichwertig:

a. $(x_1, \ldots, x_r, y_1, \ldots, y_s)$ ist eine Basis von V.

b. $(y_1, \ldots, y_s)$ ist Basis eines Komplementes von U.

c. $\left(\overline{y_1}, \ldots, \overline{y_s}\right)$ ist eine Basis von V/U.

Beweis: Die Äquivalenz von b. und c. folgt aus Proposition 2.37 und Aufgabe 3.29. Die Äquivalenz von a. und b. folgt aus dem Beweis von 3.23 und Korollar 4.16. □

Satz 4.20 (Dimensionsformel für lineare Abbildungen)
Es sei $f : V \longrightarrow W$ *eine* K*-lineare Abbildung und* $\dim_K(V) < \infty$. *Dann gilt*

$$\dim_K(V) = \dim_K\big(\mathrm{Ker}(f)\big) + \dim_K\big(\mathrm{Im}(f)\big).$$

Beweis: Aus dem Homomorphiesatz 2.34 erhalten wir den Isomorphismus

$$V/\mathrm{Ker}(f) \cong \mathrm{Im}(f),$$

so daß die Formel dann aus Korollar 4.18 folgt. □

Beispiel 4.21
Die lineare Abbildung $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x_1, x_2)^t \mapsto x_1 - x_2$ hat den Kern $\mathrm{Lin}\left((1,1)^t\right)$ von Dimension eins und ist surjektiv. Wir erhalten also die Formel

$$\dim_{\mathbb{R}}(\mathbb{R}^2) = 2 = 1 + 1 = \dim_{\mathbb{R}}(\mathrm{Ker}(f)) + \dim_{\mathbb{R}}\big(\mathrm{Im}(f)\big).$$

D) Bijektivität linearer Abbildungen

Korollar 4.22 (Injektiv = surjektiv = bijektiv)
Es seien V *und* W *zwei endlich-dimensionale* K*-Vektorräume gleicher Dimension und* $f : V \longrightarrow W$ *sei* K*-linear. Dann sind gleichwertig:*

a. f *ist bijektiv,*

b. f *ist injektiv,*

c. f *ist surjektiv.*

Beweis: Ist f injektiv, so ist $\mathrm{Ker}(f) = \{0\}$, und wir erhalten aus der Dimensionsformel für lineare Abbildungen 4.20

$$\dim_K(W) = \dim_K(V) = \dim_K\big(\mathrm{Ker}(f)\big) + \dim_K\big(\mathrm{Im}(f)\big) = \dim_K\big(\mathrm{Im}(f)\big).$$

Wegen Lemma 4.13 gilt dann $W = \mathrm{Im}(f)$ und f ist surjektiv.

Ist f surjektiv, so ist $W = \mathrm{Im}(f)$ und wir erhalten aus der Dimensionsformel für lineare Abbildungen 4.20

$$\dim_K\big(\mathrm{Ker}(f)\big) = \dim_K(V) - \dim_K\big(\mathrm{Im}(f)\big) = \dim_K(V) - \dim_K(W) = 0.$$

Dann ist aber $\mathrm{Ker}(f) = \{0\}$ und somit ist f injektiv. □

Korollar 4.23 (Invertierbare Matrizen)
Sind $A, B \in \mathrm{Mat}_n(K)$ *mit* $AB = \mathbb{1}_n$, *so gilt auch* $BA = \mathbb{1}_n$ *und* $A \in \mathrm{Gl}_n(K)$.

Beweis: Aus $AB = \mathbb{1}_n$ folgt

$$f_A \circ f_B = f_{AB} = f_{\mathbb{1}_n} = \mathrm{id}_{K^n},$$

so daß f_B injektiv mit Linksinverser f_A ist. Nach Korollar 4.22 ist f_B dann aber schon bijektiv, und die Linksinverse ist die Inverse von f_B. Damit folgt dann auch

$$f_{\mathbb{1}_n} = \mathrm{id}_{K^n} = f_B \circ f_A = f_{BA},$$

und damit $BA = \mathbb{1}_n$. □

Bemerkung 4.24 (Ringe und Moduln)
In den vorhergehenden Kapiteln haben wir stets angemerkt, welche Aussagen auch für Moduln und lineare Abbildungen über kommutativen Ringen mit Eins wahr bleiben. In diesem Kapitel gilt das im wesentlichen für keine Aussage. Die Beweise beruhen sämtlich auf dem Basisergänzungssatz und dem Austauschlemma, und beide Aussagen sind über beliebigen Ringen falsch, ihre Beweise benötigen Division. Allein Korollar 4.23 bleibt wahr, allerdings braucht man einen neuen Beweis.

Aufgaben

Aufgabe 4.25
Es sei V ein K-Vektorraum mit $\dim_K(V) = 5$, und U und U' Unterräume mit $\dim_K(U) = 3$ und $\dim_K(U') = 4$.

a. Welche Werte kann $\dim_K(U \cap U')$ annehmen?
b. Gib für jeden der Werte von $\dim_K(U \cap U')$ ein Beispiel (K, V, U, U') an.

Aufgabe 4.26
Finde einen K-Vektorraum V sowie zwei K-lineare Abbildungen $f, g : V \longrightarrow V$, so daß folgendes gilt:

a. f ist injektiv, aber nicht surjektiv.
b. g ist surjektiv, aber nicht injektiv.

Aufgabe 4.27
Es sei V ein K-Vektorraum mit $1 \leq \dim_K(V) = n < \infty$ und $g \in \mathrm{End}_K(V)$.
Zeige, es gibt eine Zahl $0 \leq k \leq n$ mit

$$\mathrm{Ker}(g^0) \subsetneqq \mathrm{Ker}(g^1) \subsetneqq \ldots \subsetneqq \mathrm{Ker}(g^k) = \mathrm{Ker}(g^{k+i})$$

für alle $i \geq 1$.

Aufgabe 4.28
Es sei $B := \big((3,5,2)^t, (1,1,-1)^t, (2,4,1)^t\big)$.

a. Zeige, B ist eine Basis von $\mathbb{R}^3$.
b. Ersetze *mit Hilfe des Austauschsatzes von Steinitz* zwei Vektoren in B durch die Vektoren $(1,3,2)^t$ und $(-2,1,2)^t$.

Aufgabe 4.29
Sei V ein K-Vektorraum und $F = (v_1, \ldots, v_5)$ eine linear unabhängige Familie in V. Welchen der Vektoren $v_1, \ldots, v_5$ kann man durch $v := v_2 - v_3 + v_4 - v_5$ ersetzen, so dass die daraus resultierende Familie wieder linear unabhängig ist? Begründe Deine Aussage.

Aufgabe 4.30

Sei K ein Körper.

a. Begründe, weshalb die Mengen $U := \{(a_1, \ldots, a_n)^t \in K^n \mid a_1 = \ldots = a_n\}$ und $U' := \left\{(a_1, \ldots, a_n)^t \in K^n \;\middle|\; a_1 + \ldots + a_n = 0\right\}$ Unterräume des K^n sind.

b. Bestimme $\dim_K(U)$, $\dim_K(U')$, $\dim_K(U \cap U')$ und $\dim_K(U + U')$.

§ 5 Lineare Abbildungen und Matrizen

In diesem Abschnitt sind V und W zwei endlich-dimensionale K-Vektorräume mit Basen $B = (b_1, \ldots, b_n)$ und $D = (d_1, \ldots, d_m)$, sofern nichts anderes gesagt wird.
Ferner bezeichnen wir mit $E = (e_1, \ldots, e_n)$ die kanonische Basis von K^n und mit $F = (f_1, \ldots, f_m)$ die kanonische Basis von K^m.

A) Matrixdarstellung linearer Abbildungen

Definition 5.1 (Matrixdarstellung einer linearen Abbildung)
Gegeben seien eine Basis $B = (b_1, \ldots, b_n)$ von V und eine Basis $D = (d_1, \ldots, d_m)$ von W sowie eine lineare Abbildung $f : V \longrightarrow W$.

a. Ist $x \in V$, so läßt sich x nach Proposition 3.12 auf eindeutige Weise darstellen als Linearkombination der Basis B

$$x = \lambda_1 \cdot b_1 + \ldots \lambda_n \cdot b_n.$$

Wir nennen den Vektor $M_B(x) = (\lambda_1, \ldots, \lambda_n)^t$ den *Koordinatenvektor* oder die *Koordinaten* von x bezüglich B.

b. Für jeden Basisvektor b_j in B läßt sich das Bild $f(b_j) \in W$ unter f nach Proposition 3.12 auf eindeutige Weise als Linearkombination der Basis D darstellen

$$f(b_j) = a_{1j} \cdot d_1 + \ldots + a_{mj} \cdot d_m.$$

Schreiben wir die Koeffizienten als Spalten in eine Matrix, so erhalten wir die Matrix

$$M_D^B(f) := \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \in \mathrm{Mat}(m \times n, K),$$

die sogenannte *Matrixdarstellung* von f bezüglich der Basen B und D.

Beispiel 5.2

a. Der Koordinatenvektor eines Vektors $x = (x_1, \ldots, x_n)^t \in K^n$ bezüglich der kanonischen Basis $E = (e_1, \ldots, e_n)$ ist der Vektor $M_E(x) = x$ selbst.

b. Um den Koordinatenvektor von $x = (4, 0)^t \in \mathbb{R}^2$ bezüglich der Basis $B = \big((1, 1)^t, (1, -1)^t\big)$ zu bestimmen, muß man ihn als Linearkombination bezüglich der Basis darstellen. Man sieht leicht, daß

$$x = 2 \cdot (1, 1)^t + 2 \cdot (1, -1)^t$$

gilt. Damit erhalten wir

$$M_B(x) = (2, 2)^t.$$

c. Sei $V = \mathbb{R}^2$ mit Basis $B = (b_1, b_2) = \big((1,2)^t, (1,1)^t\big)$ und $W = \mathbb{R}^3$ mit Basis $D = (d_1, d_2, d_3) = \big((1,1,0)^t, (0,1,1)^t, (0,0,1)^t\big)$, und sei $f : V \to W$ die lineare Abbildung, die definiert wird durch

$$\begin{aligned} b_1 &\mapsto 3d_1 - 4d_2 + 6d_3, \\ b_2 &\mapsto 3d_1 - 3d_2 + 4d_3. \end{aligned}$$

Dann gilt:

$$M_D^B(f) = \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix}.$$

d. Ist $A = (a_{ij}) \in \mathrm{Mat}(m \times n, K)$ eine Matrix und $f_A : K^n \longrightarrow K^m$ die zugehörige Abbildung, dann ist

$$f_A(e_j) = A \cdot e_j = j\text{-te Spalte von } A = a_{1j} \cdot f_1 + \ldots + a_{mj} \cdot f_m.$$

Die Matrixdarstellung f_A bezüglich der kanonischen Basen ist also

$$M_F^E(f_A) = A.$$

e. Ist $V = \mathrm{Lin}(\cos, \sin)$ der Unterraum des $\mathbb{R}^{\mathbb{R}}$ aus Beispiel 4.12 mit Basis $B = (\cos, \sin)$ und ist

$$D : V \longrightarrow V : f \mapsto f'$$

der Ableitungsoperator, dann gilt

$$D(\cos) = \cos' = -\sin = 0 \cdot \cos + (-1) \cdot \sin$$

und

$$D(\sin) = \sin' = \cos = 1 \cdot \cos + 0 \cdot \sin,$$

so daß wir die Matrixdarstellung

$$M_B^B(D) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

erhalten.

Bemerkung 5.3

a. Mit der Notation aus Satz 4.9 gilt

$$M_B(x) = \phi_B(x),$$

d. h. der Koordinatenvektor von x unter B ist das Bild unter der Karte ϕ_B. Die Zuordnung

$$\phi_B : V \longrightarrow K^n : x \mapsto M_B(x)$$

ist ein Isomorphismus von Vektorräumen, der die Vektoren in einem festgelegten Koordinatensystem darstellt.

b. Nach Definition gilt zudem:

$$\boxed{\text{Die } j\text{-te Spalte von } M_D^B(f) \text{ ist der Koordinatenvektor von } f(b_j) \text{ bez. } D.}$$

Proposition 5.4 (Rechnen in Koordinaten)
Ist $f : V \longrightarrow W$ *eine lineare Abbildung und* $x \in V$, *so gilt*

$$M_D\big(f(x)\big) = M_D^B(f) \circ M_B(x),$$

d. h. der Koordinatenvektor $M_D\big(f(x)\big)$ *von* $f(x)$ *bezüglich der Basis* D *ist das Matrixprodukt der Matrixdarstellung* $M_D^B(f)$ *von* f *bezüglich* B *und* D *mit dem Koordinatenvektor* $M_B(x)$ *von* x *bezüglich* B.

Beweis: Für einen Vektor $x = \sum_{j=1}^n \lambda_j b_j$ und eine lineare Abbildung mit Matrixdarstellung $M_D^B(f) = (a_{ij})$ gilt

$$f(x) = \sum_{j=1}^n \lambda_j \cdot f(b_j) = \sum_{j=1}^n \lambda_j \cdot \sum_{i=1}^m a_{ij} \cdot d_i = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} \cdot \lambda_j \right) \cdot d_i.$$

Daraus folgt dann

$$M_D\big(f(x)\big) = (a_{ij}) \circ (\lambda_1, \ldots, \lambda_n)^t = M_D^B(f) \circ M_B(x).$$

□

Beispiel 5.5
Betrachten wir die lineare Abbildung $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ aus Beispiel 5.2 und den Vektor $x = (0,1)^t$, so gilt

$$x = 1 \cdot (1,2)^t - 1 \cdot (1,1)^t = 1 \cdot b_1 - 1 \cdot b_2$$

und damit

$$M_B(x) = (1,-1)^t.$$

Daraus leiten wir

$$M_D(f(x)) = M_D^B(f) \cdot M_B(x) = \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix} \circ \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}$$

ab, woraus

$$f(x) = 0 \cdot d_1 - 1 \cdot d_2 + 2 \cdot d_3 = 0 \cdot (1,1,0)^t - 1 \cdot (0,1,1)^t + 2 \cdot (0,0,1)^t = (0,-1,1)^t$$

folgt. Wir können also die Bilder beliebiger Vektoren ausrechnen, obwohl wir die Abbildungsvorschrift nicht als geschlossene Formel in den Standardkoordinaten kennen. Das ist allerdings mühsam, und wir werden weiter unten sehen, wie man diese Abbildungsvorschrift aus der Matrixdarstellung gewinnen kann.

Satz 5.6 (Matrixdarstellung linearer Abbildungen)
Die Abbildung

$$M_D^B : \mathrm{Hom}_K(V,W) \longrightarrow \mathrm{Mat}(m \times n, K) : f \mapsto M_D^B(f)$$

ist ein Isomorphismus von K*-Vektorräumen.*

Beweis: Wir zeigen zunächst, daß die Abbildung linear ist. Sind $f, g \in \mathrm{Hom}_K(V, W)$ mit $M_D^B(f) = (a_{ij})$ und $M_D^B(g) = (b_{ij})$ und sind $\lambda, \mu \in K$, so gilt

$$(\lambda f+\mu g)(b_j) = \lambda\cdot f(b_j)+\mu\cdot g(b_j) = \lambda\cdot\sum_{i=1}^m a_{ij}\cdot d_i+\mu\cdot\sum_{i=1}^m b_{ij}\cdot d_i = \sum_{i=1}^m(\lambda\cdot a_{ij}+\mu\cdot b_{ij})\cdot d_i,$$

woraus wir die Matrixdarstellung

$$M_D^B(\lambda f + \mu g) = (\lambda \cdot a_{ij} + \mu \cdot b_{ij}) = \lambda \cdot (a_{ij}) + \mu \cdot (b_{ij}) = \lambda \cdot M_D^B(f) + \mu \cdot M_D^B(g)$$

erhalten. Die Abbildung M_D^B ist also K-linear.

Es bleibt, zu zeigen, daß M_D^B bijektiv ist. Sei dazu $A = (a_{ij}) \in \mathrm{Mat}(m \times n, K)$ eine beliebige $m \times n$-Matrix und setzen wir

$$y_j = \sum_{i=1}^m a_{ij} \cdot d_i \in W,$$

so gibt es nach dem Existenz- und Eindeutigkeitssatz für lineare Abbildungen 3.14 genau eine lineare Abbildung $f \in \mathrm{Hom}_K(V, W)$ mit

$$f(b_j) = y_j = \sum_{i=1}^m a_{ij} \cdot d_i$$

für $j = 1, \ldots, n$, d. h. es gibt genau eine lineare Abbildung f mit

$$M_D^B(f) = (a_{ij}).$$

Die Abbildung M_D^B ist also bijektiv. □

Bemerkung 5.7 (Matrixdarstellung bezüglich der kanonischen Basen)
Für $f \in \mathrm{Hom}_K\left(K^n, K^m\right)$ definieren wir eine Matrix $A_f \in \mathrm{Mat}(m \times n, K)$, deren j-te Spalte das Bild $f(e_j)$ des j-ten Einheitsvektors ist. Dann ist

$$M_F^E : \mathrm{Hom}_K\left(K^n, K^m\right) \longrightarrow \mathrm{Mat}(m \times n, K) : f \mapsto A_f$$

die Umkehrabbildung von

$$\mathrm{Mat}(m \times n, K) \longrightarrow \mathrm{Hom}_K\left(K^n, K^m\right) : A \mapsto f_A.$$

Diesen Spezialfall von Satz 5.6 hatten wir schon in Korollar 3.17 bewiesen, wobei die Linearität der Abbildung dabei unmittelbar aus Lemma 1.8 folgt.

Lemma 5.8 (Verträglichkeit von Matrixdarstellung und Komposition)
Sind $f \in \mathrm{Hom}_K(U, V)$ *und* $g \in \mathrm{Hom}_K(V, W)$ *und sind* B, C *bzw.* D *Basen von* U, V *bzw.* W*, dann gilt*

$$M_D^B(g \circ f) = M_D^C(g) \circ M_C^B(f).$$

Beweis: Ist $B = (b_1, \ldots, b_n)$, so gilt für den j-ten Einheitsvektor $e_j \in K^n$

$$M_B(b_j) = e_j$$

und mit Proposition 5.4 folgt dann

$$\begin{aligned} M^B_D(g \circ f) \circ e_j =& M^B_D(g \circ f) \circ M_B(b_j) = M_D\big((g \circ f)(b_j)\big) \\ =& M_D\big(g(f(b_j))\big) = M^C_D(g) \circ M_C\big(f(b_j)\big) \\ =& M^C_D(g) \circ \big(M^B_C(f) \circ M_B(b_j)\big) = \big(M^C_D(g) \circ M^B_C(f)\big) \circ e_j. \end{aligned}$$

Da die Multiplikation einer Matrix mit e_j die j-te Spalte dieser Matrix liefert, stimmen in $M^B_D(g \circ f)$ und in $M^C_D(g) \circ M^B_C(f)$ also die j-te Spalte überein und das für alle $j = 1, \ldots, n$. Die Matrizen sind also identisch. □

Bemerkung 5.9 (K-Algebren)
Ein K-Vektorraum $(B, +, \cdot)$, auf dem zusätzlich eine Multiplikation

$$\circ : B \times B \to B : (x, y) \mapsto x \circ y$$

definiert ist, so daß $(B, +, \circ)$ ein (nicht unbedingt kommutativer) Ring mit Eins 1_B ist, heißt eine K-*Algebra*, falls die Skalarmultiplikation mit der Ringmultiplikation verträglich ist, d. h. für $\lambda \in K$ und $x, y \in B$ gelten:

$$\lambda \cdot (x \circ y) = (\lambda \cdot x) \circ y = x \circ (\lambda \cdot y).$$

Ein K-*Algebrenisomorphismus* ist eine bijektive Abbildung $\varphi : A \to B$ zwischen zwei K-Algebren A und B, die mit allen drei Operationen verträglich ist und die 1_A auf 1_B abbildet.

Beispiele für K-Algebren, die für unsere Vorlesung von Bedeutung sind, sind $\big(\mathrm{End}_K(V), +, \cdot, \circ\big)$ und $\big(\mathrm{Mat}(n, K), +, \cdot, \circ\big)$, und das folgende Korollar besagt, daß diese isomorph zueinander sind.

Korollar 5.10 (M^B_B ist ein K-Algebrenisomorphismus)
Für zwei Endomorphismen $f\, g in\, \mathrm{End}_K(V)$ *gilt*

$$M^B_B(f \circ g) = M^B_B(f) \circ M^B_B(g).$$

Insbesondere, $M^B_B : \mathrm{End}_K(V) \longrightarrow \mathrm{Mat}_n(K)$ *ist ein* K-*Algebrenisomorphismus.*

Beweis: Die Aussage folgt unmittelbar aus Lemma 5.8. □

Unser Ziel ist es nun, in obigem Beispiel 5.5 aus $M^B_D(f)$ die Matrix $A_f = M^E_F(f)$ zu bestimmen. Dazu führen wir folgende allgemeine Begriffsbildung ein.

Definition 5.11 (Basiswechsel)
Sind $B = (b_1, \ldots, b_n)$ und $B' = \big(b'_1, \ldots, b'_n\big)$ zwei Basen von V, so besitzt jedes b_j eine eindeutige Darstellung als Linearkombination der Basis B'

$$b_j = a_{1j} \cdot b'_1 + \ldots + a_{nj} \cdot b'_n.$$

Schreiben wir die Koeffizienten als Spalten in eine Matrix, so erhalten wir die Matrix

$$\mathsf{T}^{\mathsf{B}}_{\mathsf{B}'} := \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \in \mathrm{Mat}_n(K),$$

den *Basiswechsel* oder die *Koordinatentransformationsmatrix* bezüglich (B, B').
Es gilt also:

$$\boxed{\text{Die } j\text{-te Spalte von } T^B_{B'} \text{ ist der Koordinatenvektor } M_{B'}(b_j) \text{ von } b_j \text{ bez. } B'.}$$

Bemerkung 5.12 (Basiswechsel als Matrixdarstellung)
Offensichtlich ist der Basiswechsel ein Spezialfall einer Matrixdarstellung

$$T^B_{B'} = M^B_{B'}(\mathrm{id}_V),$$

nämlich die Matrixdarstellung der Identität bezüglich der Basen B und B'.

Lemma 5.13 (Basiswechselmatrizen sind invertierbar)
Sind B und B' zwei Basen von V, so ist $T^B_{B'}$ invertierbar mit

$$\left(T^B_{B'}\right)^{-1} = T^{B'}_B.$$

Beweis: Mit Lemma 5.8

$$T^{B'}_B \circ T^B_{B'} = M^{B'}_B(\mathrm{id}_V) \circ M^B_{B'}(\mathrm{id}_V) = M^B_B(\mathrm{id}_V \circ \mathrm{id}_V) = M^B_B(\mathrm{id}_V) = \mathbb{1}_n.$$

□

Satz 5.14 (Basiswechsel bei Matrixdarstellungen)
*Seien B und B' Basen von V, D und D' Basen von W und $f \in \mathrm{Hom}_K(V, W)$.
Dann gilt:*

$$M^{B'}_{D'}(f) = T^D_{D'} \circ M^B_D(f) \circ T^{B'}_B.$$

Beweis: Die Aussage folgt unmittelbar aus Lemma 5.8

$$\begin{aligned} M^{B'}_{D'}(f) =& M^{B'}_{D'}(\mathrm{id}_W \circ f \circ \mathrm{id}_V) \\ =& M^D_{D'}(\mathrm{id}_W) \circ M^B_D(f) \circ M^{B'}_B(\mathrm{id}_V) \\ =& T^D_{D'} \circ M^B_D(f) \circ T^{B'}_B. \end{aligned}$$

□

Korollar 5.15 (Basiswechsel bei Endomorphismen)
Sind B und B' Basen von V, ist $T = T^{B'}_B$ und ist $f \in \mathrm{End}_K(V)$, so gilt

$$M^{B'}_{B'}(f) = T^{-1} \circ M^B_B(f) \circ T.$$

Beweis: Dies folgt aus Satz 5.14, weil nach Lemma 5.13 $\left(T^{B'}_B\right)^{-1} = T^B_{B'}$. □

Beispiel 5.16
Wir wollen nun für die Abbildung in Beispiel 5.5 die Matrixdarstellung $M^E_F(f)$ bezüglich der kanonischen Basen berechnen. Nach Satz 5.14 gilt:

$$M^E_F(f) = T^D_F \circ M^B_D(f) \circ T^E_B.$$

Um T^D_F auszurechnen, müssen wir d_1, d_2 und d_3 in der kanonischen Basis ausdrücken und die Koeffizienten als Spaltenvektoren in die Matrix T^D_F übertragen:

$$T^D_F = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Um T^E_B zu ermitteln, müßten wir die Einheitsvektoren e_1 und e_2 als Linearkombination der Basis B darstellen, was auf das Lösen zweier Gleichungssysteme hinaus liefe. Stattdessen können wir aber auch T^B_E bestimmen und anschließend invertieren, was sich im Falle einer (2×2)-Matrix anbietet, da das Invertieren sehr einfach ist (vgl. Aufgabe 1.14),

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

sofern die Matrix invertierbar ist.

Analog zum Fall von T^D_F erhalten wir

$$T^B_E = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix},$$

und somit

$$T^E_B = \left(T^B_E\right)^{-1} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix}.$$

Also gilt:

$$M^E_F(f) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \circ \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix} \circ \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

B) f-invariante Unterräume

Definition 5.17 (f-invarianter Unterraum)
Ist $f : V \longrightarrow V$ eine K-lineare Abbildung und $U \leq V$ ein Unterraum mit $f(U) \subseteq U$, so nennen wir U einen f-*invarianten Unterraum.*

Bemerkung 5.18 (f-invariante Unterräume)
Aus Aufgabe 2.46 wissen wir, daß jeder f-invariante Unterraum U zwei K-lineare Abbildungen

$$f_U : U \longrightarrow U : x \mapsto f(x)$$

und

$$f_{V/U} : V/U \longrightarrow V/U : \overline{x} \mapsto \overline{f(x)}$$

induziert. Mit Hilfe dieser Abbildungen erhalten wir eine vorteilhafte Blockgestalt bei der Matrixdarstellung.

Proposition 5.19 (Matrixdarstellung in Blockform)
Es sei $f : V \longrightarrow V$ *eine* K*-lineare Abbildung und* $U \leq V$ *ein* f*-invarianter Unterraum. Ferner sei* $B' = (x_1, \ldots, x_k)$ *eine Basis von* U *und* $B = (x_1, \ldots, x_n)$ *eine Ergänzung von* B' *zu einer Basis von* V.

Dann ist $B'' = \left(\overline{x_{k+1}}, \ldots, \overline{x_n}\right)$ *eine Basis von* V/U *und es gilt*

$$M_B^B(f) = \left(\begin{array}{c|c} M_{B'}^{B'}(f_U) & * \\ \hline 0 & M_{B''}^{B''}(f_{V/U}) \end{array}\right),$$

wobei $0 \in \mathrm{Mat}((n-k) \times k, K)$ *die Nullmatrix ist und* $* \in \mathrm{Mat}(k \times (n-k), K)$ *eine geeignete Matrix ist.*

Beweis: Daß B'' eine Basis von V/U ist, wissen wir bereits aus Bemerkung 4.19. Sei nun $M_B^B(f) = (a_{ij}) \in \mathrm{Mat}_n(K)$, $M_{B'}^{B'}(f_U) = (b_{ij}) \in \mathrm{Mat}_k(K)$ und $M_{B''}^{B''}(f_{V/U}) = (c_{ij}) \in \mathrm{Mat}_{n-k}(K)$.

Für $j = 1, \ldots, k$ gilt dann

$$\sum_{i=1}^{k} b_{ij} \cdot x_i + \sum_{i=k+1}^{n} 0 \cdot x_i = \sum_{i=1}^{k} b_{ij} \cdot x_i = f_U(x_j) = f(x_j) = \sum_{i=1}^{n} a_{ij} \cdot x_i.$$

Da die Darstellung als Linearkombination von Basisvektoren eindeutig ist, folgt somit

$$(a_{ij})_{i,j=1,\ldots,k} = (b_{ij})_{i,j=1,\ldots,k} = M_{B'}^{B'}(f_U)$$

und

$$(a_{ij})_{i=k+1,\ldots,n,\, j=1,\ldots,k} = 0 \in \mathrm{Mat}((n-k) \times k, K).$$

Für $j = k+1, \ldots, n$ erhalten wir analog

$$\sum_{i=k+1}^{n} c_{ij}\overline{x_i} = f_{V/U}\left(\overline{x_j}\right) = \overline{f(x_j)} = \overline{\sum_{i=1}^{n} a_{ij} \cdot x_i} = \sum_{i=1}^{n} a_{ij} \cdot \overline{x_i} = \sum_{i=k+1}^{n} a_{ij} \cdot \overline{x_i},$$

da $\overline{x_i} = \overline{0}$ für $i = 1, \ldots, k$. Aus der Eindeutigkeit der Darstellung als Linearkombination von Basisvektoren erhalten wir also wieder

$$(a_{ij})_{i,j=k+1,\ldots,n} = (c_{ij})_{i,j=k+1,\ldots,n} = M_{B''}^{B''}(f_{V/U}).$$

Insgesamt haben wir damit die Behauptung

$$M_B^B(f) = (a_{ij})_{i,j=1,\ldots,n} = \left(\begin{array}{c|c} M_{B'}^{B'}(f_U) & * \\ \hline 0 & M_{B''}^{B''}(f_{V/U}) \end{array}\right)$$

gezeigt. □

Proposition 5.20 (Matrixdarstellung in Blockdiagonalgestalt)
Sei $f : V \longrightarrow V$ *eine* K*-lineare Abbildung und* $V = U_1 \oplus \ldots \oplus U_k$ *die direkte Summe nicht-trivialer* f*-invarianter Unterräume* $U_1, \ldots, U_k$ *mit Basen* $B_1, \ldots, B_k$*.*

Dann ist $B = B_1 \cup \ldots \cup B_k$ *eine Basis von* V *und es gilt*

$$M_B^B(f) = \left(\begin{array}{c|c|ccc} M_{B_1}^{B_1}(f_{U_1}) & 0 & \cdots & & 0 \\ \hline 0 & \boxed{M_{B_2}^{B_2}(f_{U_2})} & & & \vdots \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & \cdots & & M_{B_k}^{B_k}(f_{U_k}) \end{array} \right).$$

Beweis: Daß B eine Basis von V ist, wissen wir aus Aufgabe 3.30.

Es sei $B = (x_1, \ldots, x_n)$ und $M_B^B(f) = (a_{ij}) \in \mathrm{Mat}_n(K)$. Halten wir ein $1 \leq l \leq k$ fest, so ist $B_l = (x_r, \ldots, x_s)$ für geeignete $1 \leq r < s \leq n$ und $M_{B_l}^{B_l}(f_{U_l}) = (b_{ij})_{i,j=r,\ldots,s} \in \mathrm{Mat}_{s-r+1}(K)$. Wie im Beweis von Proposition 5.19 sehen wir für $j = r, \ldots, s$ dann

$$\sum_{i=1}^{r-1} 0 \cdot x_i + \sum_{i=r}^{s} b_{ij} \cdot x_i + \sum_{i=s+1}^{n} 0 \cdot x_i = \sum_{i=1}^{k} b_{ij} \cdot x_i = f_{U_l}(x_j) = f(x_j) = \sum_{i=1}^{n} a_{ij} \cdot x_i,$$

woraus wieder

$$(a_{ij})_{i,j=r,\ldots,s} = (b_{ij})_{i,j=r,\ldots,s} = M_{B_l}^{B_l}(f_{U_l})$$

sowie $a_{ij} = 0$ für alle $i < r$ und $i > s$ folgt. Damit ist die Behauptung gezeigt. □

C) Äquivalenz von Matrizen und der Rang

Die Koordinatentransformationen in Vektorräumen mit Basen führen auf folgende Äquivalenzbegriffe für Matrizen.

Definition 5.21 (Äquivalenz von Matrizen)
Eine Matrix $A' \in \mathrm{Mat}(m \times n, K)$ heißt *äquivalent* zu $A \in \mathrm{Mat}(m \times n, K)$, falls es invertierbare Matrizen $S \in \mathrm{Gl}_m(K)$ und $T \in \mathrm{Gl}_n(K)$ gibt mit

$$A' = S \circ A \circ T.$$

Bemerkung 5.22 (Äquivalenz von Matrizen als Äquivalenzrelation)
Die Äquivalenz von Matrizen eine Äquivalenzrelation auf $\mathrm{Mat}(m \times n, K)$ ist.

Denn für $A, B, C \in \mathrm{Mat}(m \times n, K)$ gelten:

- $A = \mathbb{1}_m \circ A \circ \mathbb{1}_n$ und mithin ist A äquivalent zu A;
- wenn B äquivalent zu A ist, gibt es $S \in \mathrm{Gl}_m(K)$ und $T \in \mathrm{Gl}_n(K)$ mit $B = S \circ A \circ T$ und mithin $S^{-1} \circ B \circ T^{-1} = A$, so daß auch A äquivalent zu B;
- wenn B äquivalent zu A und C äquivelent zu B ist, so gibt es Matrizen $S, U \in \mathrm{Gl}_m(K)$ und $T, V \in \mathrm{Gl}_n(K)$ mit $B = S \circ A \circ T$ und $C = U \circ B \circ V$ und mithin gilt auch $C = (U \circ S) \circ A \circ (T \circ V)$ mit $U \circ S \in \mathrm{Gl}_m(K)$ und $T \circ V \in \mathrm{Gl}_n(K)$, so daß C auch äquivalent zu A ist.

Beispiel 5.23
Die Matrizen

$$A = \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix} \quad \text{und} \quad A' = \begin{pmatrix} 3 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}$$

sind äquivalent, da wir in Beispiel 5.16

$$A' = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \circ \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix} \circ \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix}$$

gezeigt haben.

Definition 5.24 (Rang)

a. Ist $f \in \mathrm{Hom}_K(V, W)$, so definieren wir den *Rang* von f als

$$\mathrm{rang}(f) := \dim_K\big(\mathrm{Im}(f)\big).$$

b. Ferner definieren wir für eine Matrix $A \in \mathrm{Mat}(m \times n, K)$ den *Rang* von A durch:

$$\mathrm{rang}(A) := \mathrm{rang}(f_A).$$

Bemerkung 5.25 (Rangabschätzung)

a. Für $f \in \mathrm{Hom}_K(V, W)$ gilt wegen der Dimensionsformel für lineare Abbildungen

$$\mathrm{rang}(f) = \dim_K(V) - \dim_K\big(\mathrm{Ker}(f)\big) \leq \dim_K(V)$$

und da $\mathrm{Im}(f)$ ein Unterraum von W ist, gilt auch $\mathrm{rang}(f) \leq \dim_K(W)$.

b. Man beachte, daß das Bild von f_A von den Spalten von A erzeugt wird, so daß der Rang von A die Anzahl linear unabhängiger Spalten von A ist.
Zudem folgt aus a. für $A \in \mathrm{Mat}(m \times n, K)$

$$\mathrm{rang}(A) \leq \min\{m, n\}.$$

Beispiel 5.26
Da der Rang einer Matrix die Anzahl linear unabhängiger Spalten ist, gelten

$$\mathrm{rang}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0, \quad \mathrm{rang}\begin{pmatrix} 1 & 0 & 2 \\ 1 & 0 & 2 \end{pmatrix} = 1, \quad \mathrm{rang}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix} = 2.$$

Satz 5.27 (Invertierbare Matrizen haben vollen Rang.)
Eine Matrix $A \in \mathrm{Mat}_n(K)$ *ist genau dann invertierbar, wenn* $\mathrm{rang}(A) = n$.
Insbesondere sind die Spalten einer invertierbaren $n \times n$*-Matrix eine Basis des* K^n.

Beweis: Nach Aufgabe 3.33 ist A genau dann invertierbar, wenn $f_A : K^n \longrightarrow K^n$ bijektiv ist. Wegen Korollar 4.22 ist dies genau dann der Fall, wenn f_A surjektiv ist, d.h. wenn $\mathrm{Im}(f_A) = K^n$. Wegen $\mathrm{Im}(f_A) \subseteq K^n$ und Lemma 4.13 ist dies wiederum genau dann der Fall, wenn

$$n = \dim_K(\mathrm{Im}(f_A)) = \mathrm{rang}(A).$$

Also ist A genau dann invertierbar, wenn $\text{rang}(A) = n$ gilt. In diesem Fall sind die Spalten von A linear unabhängig in K^n und bilden mithin eine Basis des K^n. □

Bemerkung 5.28
Sind die Matrizen $A, A' \in \text{Mat}(m \times n, K)$ äquivalent, so gibt es Basen B von K^n und D von K^m, so daß

$$A' = M_D^B(f_A),$$

d.h. A und A' sind Matrixdarstellungen derselben linearen Abbildung f_A bezüglich verschiedener Basen!

Dazu betrachten wir einfach die Matrizen $S \in \text{Gl}_m(K)$ und $T \in \text{Gl}_n(K)$ mit $A' = S \circ A \circ T$. Die Spalten von S^{-1} sind linear unabhängig und bilden eine Basis D von K^m nach Satz 5.27. Ebenso bilden die Spalten von T eine Basis B von K^n. Für diese Basen gilt aber nach Konstruktion

$$T_F^D = S^{-1} \quad \text{und} \quad T_E^B = T.$$

Insgesamt erhalten wir also

$$M_D^B(f_A) = T_D^F \circ M_F^E(f_A) \circ T_E^B = S \circ A \circ T = A'.$$

Wir werden nun zeigen, daß der Rang der Matrixdarstellung einer linearen Abbildung *nicht* von der Wahl der Basen abhängt, bezüglich derer man die Matrixdarstellung bildet.

Proposition 5.29 (Rang einer Matrixdarstellung)
Ist $f \in \text{Hom}_K(V, W)$*, so gilt*

$$\text{rang}(f) = \text{rang}\left(M_D^B(f)\right).$$

Insbesondere haben äquivalente Matrizen den gleichen Rang.

Beweis: Wir betrachten die Karten $\phi_B : V \longrightarrow K^n$ und $\phi_D : W \longrightarrow K^m$. Aus Proposition 5.4 ergibt sich

$$\begin{aligned}\phi_D \circ f \circ \phi_B^{-1}(e_j) =& \phi_D\big(f(b_j)\big) = M_D\big(f(b_j)\big)\\ =& M_D^B(f) \circ M_B(b_j) = M_D^B(f) \circ e_j = f_{M_D^B(f)}(e_j).\end{aligned}$$

Da die linearen Abbildungen $\phi_D \circ f \circ \phi_B^{-1}$ und $f_{M_D^B(f)}$ auf der Basis B übereinstimmen, sind sie identisch. Beachtet man nun noch, daß ϕ_D ein Isomorphismus ist und somit die Dimension eines Vektorraumes erhält, dann folg

$$\begin{aligned}\text{rang}\left(M_D^B(f)\right) &= \text{rang}\left(f_{M_D^B(f)}\right) = \text{rang}\left(\phi_D \circ f \circ \phi_B^{-1}\right)\\ &= \dim_K\left(\phi_D\left(f\left(\phi_B^{-1}(K^n)\right)\right)\right) = \dim_K\left(\phi_D\left(f(V)\right)\right)\\ &= \dim_K\left(f(V)\right) = \text{rang}(f)\end{aligned}$$

Sind A und A' äquivalent, so sind sie nach Bemerkung 5.28 Matrixdarstellungen der gleichen Abbildung f_A bezüglich verschiedener Basen und wir haben gerade gesehen, daß der Rang der Matrixdarstellung nicht von der Wahl der Basen abhängt. □

Beispiel 5.30
Die Abbildung f in Beispiel 5.2 hat den Rang zwei, wie man an ihrer Matrixdarstellung sieht:

$$M_D^B(f) = \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix}.$$

Satz 5.31 (Normalform einer Matrixdarstellung bezüglich Äquivalenz)
Es sei $f \in \mathrm{Hom}_K(V, W)$ *mit* $\mathrm{rang}(f) = r$. *Dann gibt es Basen* B *von* V *und* D *von* W *mit*

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

wobei hier 0 *jeweils die Nullmatrix der entsprechenden Größe meint.*[1] *Wir bezeichnen die rechte Seite der obigen Gleichung auch als die* Normalform *von* f *bezüglich Äquivalenz.*

Beweis: Wähle vermöge Lemma 3.23 ein Komplement U von $\mathrm{Ker}(f)$. Nach Satz 2.34 und Lemma 2.37 ist die folgende Abbildung ein Isomorphismus

$$f_{|U} : U \longrightarrow \mathrm{Im}(f) : x \mapsto f(x).$$

Wähle eine Basis $(d_1, \ldots, d_r)$ von $\mathrm{Im}(f)$. Dann ist $(b_1, \ldots, b_r)$ mit $b_i := (f_{|U})^{-1}(d_i)$ eine Basis von U, nach Aufgabe 3.29. Wähle nun eine Basis $(b_{r+1}, \ldots, b_n)$ von $\mathrm{Ker}(f)$, dann ist $B = (b_1, \ldots, b_n)$ eine Basis von $V = U \oplus \mathrm{Ker}(f)$. Ergänze ferner $(d_1, \ldots, d_r)$ zu einer Basis $D = (d_1, \ldots, d_m)$ von W vermöge Satz 3.21. Dann:

$$f(b_i) = \begin{cases} d_i, & i = 1, \ldots, r, \\ 0, & i = r+1, \ldots, n. \end{cases}$$

Also hat $M_D^B(f)$ die gewünschte Gestalt. □

Korollar 5.32 (Normalform einer Matrix bezüglich Äquivalenz)
Zu $A \in \mathrm{Mat}(m \times n, K)$ *mit* $r = \mathrm{rang}(A)$ *existieren Matrizen* $S \in \mathrm{Gl}_m(K)$ *und* $T \in \mathrm{Gl}_n(K)$ *mit*

$$S \circ A \circ T = \left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right). \tag{7}$$

Die rechte Seite heißt die Normalform *von* A *bezüglich Äquivalenz.*

Beweis: Anwendung des Satzes 5.31 auf $f_A : K^n \longrightarrow K^m$ liefert B und D von K^n bzw. K^m mit

$$\left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right) = M_D^B(f_A) = T_D^F \circ M_F^E(f_A) \circ T_E^B = T_D^F \circ A \circ T_E^B.$$

[1] Man bezeichnet die vier Matrizen $\mathbb{1}_r \in \mathrm{Mat}_r(K)$, $0 \in \mathrm{Mat}\big(r \times (n-r), K\big)$, $0 \in \mathrm{Mat}\big((m-r) \times r, K\big)$ und $0 \in \mathrm{Mat}((m-r) \times (n-r), K)$ auch als *Blöcke* von $M_D^B(f)$ und die Matrix $M_D^B(f)$ als eine *Blockmatrix*.

Die Behauptung folgt also, da $S := T^F_D$ und $T := T^B_E$ invertierbar sind. □

Beispiel 5.33

Die folgende Matrix $A \in \mathrm{Mat}(3 \times 4, \mathbb{R})$ hat Rang 2 und hat somit die Matrix B als Normalform:

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

Bemerkung 5.34 (Normalform als guter Repräsentant einer Äquivalenzklasse)

Aus Korollar 5.32 folgt, daß zwei Matrizen genau dann äquivalent sind, wenn sie den gleichen Rang haben. $\mathrm{Mat}(m \times n, K)$ zerfällt also in $\min\{m, n\} + 1$ Äquivalenzklassen und jede Äquivalenzklasse ist durch den Rang einer ihrer Matrizen eindeutig bestimmt. Darüber hinaus besitzt jede Äquivalenzklasse $\overline{A}$, $A \in \mathrm{Mat}(m \times n, K)$, einen besonders schönen Repräsentanten, nämlich

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

Diesen Repräsentanten der Äquivalenzklasse von A nennt man die *Normalform von* A *bezüglich Äquivalenz*.

Korollar 5.35 (Zeilen- und Spaltenrang)

a. $A \in \mathrm{Mat}_n(K)$ *ist genau dann invertierbar, wenn* A^t *invertierbar ist. In dem Fall gilt*

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

b. *Für eine Matrix* $A \in \mathrm{Mat}(m \times n, K)$ *gilt*

$$\mathrm{rang}(A) = \mathrm{rang}\left(A^t\right).$$

Insbesondere ist die Anzahl linear unabhängiger Spalten in A *gleich der Anzahl linear unabhängiger Zeilen!*

Beweis: a. Es sei A invertierbar und $B = A^{-1}$. Dann gilt

$$B^t A^t = (AB)^t = \mathbb{1}_n^t = \mathbb{1}_n,$$

so daß A^t nach Korollar 4.23 invertierbar ist mit Inverser $\left(A^t\right)^{-1} = \left(A^{-1}\right)^t$. Ist umgekehrt A^t invertierbar, so ist nach dem eben gezeigten auch $A = (A^t)^t$ invertierbar.

b. Nach Korollar 5.32 finden wir invertierbare Matrizen $S \in \mathrm{Gl}_m(K)$ und $T \in \mathrm{Gl}_n(K)$, so daß

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

in Normalform mit $r = \mathrm{rang}(A)$ gegeben ist. Dann ist aber

$$T^t \circ A^t \circ S^t = \left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right)^t \in \mathrm{Mat}(n \times m, K)$$

ebenfalls eine Matrix in Normalform. Es gilt also

$$\operatorname{rang}\left(T^t \circ A^t \circ S^t\right) = r$$

und wegen Teil a. ist die Matrix $T^t \circ A^t \circ S^t$ äquivalent zu A^t, so daß sie nach Proposition 5.29 den gleichen Rang hat wie A^t.

□

Beispiel 5.36
Die Matrix

$$A = \begin{pmatrix} 1 & 1 & 2 & 3 & 5 \\ 0 & 1 & 0 & 2 & 2 \\ 0 & 0 & 1 & 3 & 1 \end{pmatrix}$$

hat offenbar den Rang 3, da die ersten drei Spalten schon linear unabhängig sind. Mithin hat auch die transponierte Matrix

$$A^t = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \\ 3 & 2 & 3 \\ 5 & 2 & 1 \end{pmatrix}$$

den Rang 3, d.h. die drei Spalten sind linear unabhängig.

Bemerkung 5.37 (Ringe und Moduln)
Die Identifikation von linearen Abbildungen und Matrizen funktioniert auch für Moduln über Ringen, wenn sie endliche Basen besitzen. Die Beweise ändern sich nicht. Man erhält also die Aussagen der Sätze und Bemerkungen 5.6, 5.7, 5.14 und 5.15 ohne Änderung für Moduln, die endliche Basen besitzen — die zugehörigen Definitionen kann man ebenfalls ohne Änderung übernehmen. Die weiteren Aussagen des Abschnitts zur Äquivalenz von Matrizen und zu deren Rang gelten in dieser Form nicht allgemein für lineare Abbildungen von Moduln. Selbst wenn zwei Moduln V und W Basen besitzen, muß das Bild einer linearen Abbildung von V nach W keine Basis haben, so daß man den Rang der Abbildung dann gar nicht definieren kann.

Aufgaben

Aufgabe 5.38 (Zyklische Unterräume)
Es sei $f \in \operatorname{End}_K(V)$, $0 \neq x \in V$ und $m \in \mathbb{N}$ mit $f^{m-1}(x) \neq 0$ und $f^m(x) = 0$.

a. Zeige, $B = \left(f^{m-1}(x), f^{m-2}(x), \ldots, f(x), x\right)$ ist eine Basis von $U = \operatorname{Lin}(B)$.
b. Zeige, U ist f-invariant.
c. Bestimme $M_B^B(f_U)$.

Wir nennen U einen *zyklischen Unterraum* von V.

Aufgabe 5.39
Für Matrizen $A \in \mathrm{Mat}(n \times p, K)$ und $B \in \mathrm{Mat}(m \times n, K)$ gilt:

$$\mathrm{rang}(B \circ A) \leq \min\big\{\mathrm{rang}(A), \mathrm{rang}(B)\big\}.$$

Aufgabe 5.40
Für eine Matrix $A \in \mathrm{Mat}(m \times n, K)$ bezeichne $\mathrm{ZR}(A)$ die lineare Hülle der Zeilen von A und $\mathrm{SR}(A)$ die lineare Hülle der Spalten von A.
Zeige für $A \in \mathrm{Mat}(m \times n, K)$, $S \in \mathrm{Gl}_m(K)$ und $T \in \mathrm{Gl}_n(K)$

$$\mathrm{ZR}(A) = \mathrm{ZR}(SA) \quad \text{und} \quad \mathrm{SR}(A) = \mathrm{SR}(AT).$$

Aufgabe 5.41
Betrachte den Vektorraum P_n der Polynome vom Grad höchstens n (siehe Beispiel 2.6) mit Basis $B = (t^0, t^1, \ldots, t^n)$ und die formale Ableitung

$$d : P_n \longrightarrow P_n : \sum_{k=0}^{n} a_k \cdot t^k \mapsto \sum_{k=1}^{n} k \cdot a_k \cdot t^{k-1},$$

von der wir aus Beispiel 2.21 bereits wissen, daß sie K-linear ist.

a. Berechne die Matrixdarstellung $M_B^B(d)$ und den Rang von d.

b. Zeige, daß im Fall $n = 3$ auch $D = (t^0, t^0 + t^1, t^1 + t^2, t^2 + t^3)$ eine Basis von P_3 ist und berechne die Basiswechsel T_B^D und T_D^B sowie die Matrixdarstellung $M_D^D(d)$.

§ 6 Der Gauß-Algorithmus

In diesem Abschnitt wollen wir zeigen, daß man jede Matrix durch elementare Zeilenoperationen in Zeilen-Stufen-Form transformieren kann, und einen Algorithmus angeben, der dies tut, den *Gauß-Algorithmus*.

Definition 6.1 (Zeilen-Stufen-Form)
Es sei $A = (a_{ij}) \in \mathrm{Mat}(m \times n, K)$.

a. A ist in *Zeilen-Stufen-Form*, kurz ZSF, falls es ein r, mit $0 \le r \le m$ und Indizes $j_1, \ldots, j_r$ mit $1 \le j_1 < j_2 < \ldots < j_r \le m$ gibt, so daß folgendes gilt:
   (i) $a_{ij} = 0$ für $1 \le i \le r$ und $1 \le j < j_i$,
   (ii) $a_{ij} = 0$ für $r < i \le m$ und $j = 1, \ldots, n$, und
   (iii) $a_{ij_i} \ne 0$ für $i = 1, \ldots, r$.
   Die Körperelemente a_{ij_i} heißen die *Pivots* der Zeilen-Stufen-Form. Man beachte, daß A genau r linear unabhängige Zeilen hat und daß somit $r = \mathrm{rang}(A)$!

b. Eine Zeilen-Stufen-Form von A heißt *reduziert*, falls zusätzlich gilt:
   (iv) $a_{ij_i} = 1$ für $i = 1, \ldots, r$, und
   (v) $a_{kj_i} = 0$ für $k < i$ und $i = 1, \ldots, r$.

Bemerkung 6.2
Eine Matrix A in Zeilen-Stufen-Form ist also von der folgenden Gestalt:

$$A = \begin{pmatrix} 0 & \ldots & 0 & a_{1j_1} & * & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & * \\ 0 & \ldots & \ldots & 0 & \ldots & 0 & a_{2j_2} & * & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & * \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & \ldots & 0 & a_{3j_3} & * & \ldots & \ldots & \ldots & * \\ \vdots & & & & & & & & & & \ddots & & & & \vdots \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & a_{rj_r} & * & \ldots & * \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 0 \\ \vdots & & & & & & & & & & & & & & \vdots \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 0 \end{pmatrix}$$

Hat A reduzierte Zeilen-Stufen-Form, so sind die Pivots alle Eins und die Einträge in der Spalte oberhalb der Pivots sind alle Null.

Beispiel 6.3
Betrachte die Matrizen $A, B, C \in \mathrm{Mat}(4 \times 5, K)$ mit

$$A = \begin{pmatrix} 0 & 1 & 0 & 3 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 2 & 3 & 4 & 5 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 4 & 1 \end{pmatrix} \text{ und } C = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 3 & 4 & 0 \\ 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Die Matrix A ist in reduzierter ZSF mit $\mathrm{rang}(A) = r = 3$, $j_1 = 2$, $j_2 = 3$ und $j_3 = 5$.
Die Matrix B ist in ZSF mit $\mathrm{rang}(B) = r = 4$ und $j_1 = 1$, $j_2 = 2$, $j_3 = 3$ und $j_4 = 4$.

Die ZSF ist aber nicht reduziert.
Die Matrix C ist nicht in ZSF. Aber durch Vertauschen der beiden ersten Zeilen entsteht eine Matrix, die ZSF hat.

Für die folgende Definition erinnern wir uns (siehe Beispiel 3.11) an die Matrizen

$$E_i^j = (e_{lk})_{l,k=1,\ldots,n} = (\delta_{il} \cdot \delta_{jk})_{l,k=1,\ldots,n} \in \mathrm{Mat}_n(K),$$

die an der Stelle (i,j) den Eintrag 1 und sonst nur Nullen als Einträge haben.

Definition 6.4 (Elementarmatrizen)
Es seien $0 \neq \lambda \in K$, $n > 0$ und $1 \leq i,j \leq n$ mit $i \neq j$. Wir definieren die folgenden quadratischen Matrizen in $\mathrm{Mat}(n,K)$, die auch als *Elementarmatrizen* bezeichnet werden:

a. $S_i(\lambda) := \mathbb{1}_n + (\lambda - 1) \cdot E_i^i$,
b. $Q_i^j(\lambda) := \mathbb{1}_n + \lambda \cdot E_i^j$, und
c. $P_i^j := \mathbb{1}_n - E_i^i - E_j^j + E_i^j + E_j^i$.

Die Matrizen P_i^j heißen zudem *Permutationsmatrizen.*

Bemerkung 6.5 (Elementare Zeilen- und Spaltenoperationen)
Es seien $0 \neq \lambda \in K$, $1 \leq i,j \leq n$ mit $i \neq j$ und $A \in \mathrm{Mat}(n \times m, K)$.

I $S_i(\lambda) \circ A$ geht aus A hervor, indem man die i-te Zeile von A mit λ multipliziert.

II $Q_i^j(\lambda) \circ A$ geht aus A hervor, indem man zur i-ten Zeile das λ-fache der j-ten Zeile addiert.

III $P_i^j \circ A$ geht aus A hervor, indem man die i-te und j-te Zeile vertauscht.

Man nennt die Multiplikation von links mit diesen Matrizen auch *elementare Zeilenoperationen.* Analog erhält man *elementare Spaltenoperationen*, indem man mit den Matrizen von rechts multipliziert.

I' $A \circ S_j(\lambda)$ geht aus A hervor, indem man die j-te Spalte von A mit λ multipliziert.

II' $A \circ Q_i^j(\lambda)$ geht aus A hervor, indem man zur j-ten Spalte das λ-fache der i-ten Spalte addiert.

III' $A \circ P_i^j$ geht aus A hervor, indem man die i-te und j-te Spalte vertauscht.

Proposition 6.6 (Elementarmatrizen sind invertierbar.)
Es seien $0 \neq \lambda \in K$, $1 \leq i,j \leq n$ *mit* $i \neq j$ *und* $A \in \mathrm{Mat}(n \times m, K)$. *Dann gelten:*

a. $S_i\left(\lambda^{-1}\right) \circ S_i(\lambda) = \mathbb{1}_n$,
b. $Q_i^j(-\lambda) \circ Q_i^j(\lambda) = \mathbb{1}_n$, *und*
c. $P_i^j \circ P_i^j = \mathbb{1}_n$.

Insbesondere sind die Elementarmatrizen invertierbar und die Inversen sind wiederum Elementarmatrizen vom gleichen Typ.

Beweis: Wir führen den Beweis für b. vor. Die übrigen Teile lassen sich dann analog zeigen. Für $0 \neq \lambda \in K$ gilt, vermittels der Distributivität der Matrixmultiplikation:

$$Q_i^j(-\lambda) \circ Q_i^j(\lambda) = \left(\mathbb{1}_n - \lambda \cdot E_i^j\right) \circ \left(\mathbb{1}_n + \lambda \cdot E_i^j\right) = \mathbb{1}_n - \lambda^2 \cdot E_i^j \circ E_i^j = \mathbb{1}_n,$$

da $E_i^j \circ E_i^j = 0$ wegen $i \neq j$. Beachte dazu, daß für $E_i^j \circ E_i^j = (c_{lk})$ gilt:

$$c_{lk} = \sum_{p=1}^{n} \delta_{il}\delta_{jp}\delta_{ip}\delta_{jk},$$

und daß für $i \neq j$ und p beliebig gilt $\delta_{jp}\delta_{ip} = 0$. □

Satz 6.7 (Existenz der reduzierten Zeilen-Stufen-Form)
Jede Matrix $A \in \mathrm{Mat}(m \times n, K)$ *läßt sich mittels endlich vieler elementarer Zeilenoperationen in reduzierte Zeilen-Stufen-Form* rZSF(A) *überführen, d.h. es gibt Elementarmatrizen* $T_1, \ldots, T_k$*, so daß*

$$\mathrm{rZSF}(A) = T_1 \circ \ldots \circ T_k \circ A.$$

Beweis: Sei $A \in \mathrm{Mat}(m \times n, K)$. Ist $A = 0$, so hat A bereits ZSF mit $r = 0$ und wir sind fertig.

Ist $A \neq 0$, so führe folgende Schritte durch:

1. Schritt: Durchlaufe die Spalten von oben nach unten, mit der ersten Spalte beginnend, bis der erste Eintrag $a_{i_1 j_1} \neq 0$ gefunden ist:

$$\begin{pmatrix} 0 & \ldots & 0 & 0 & * & \ldots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & 0 & * & \ldots & * \\ \vdots & & \vdots & a_{i_1 j_1} & * & \ldots & * \\ \vdots & & \vdots & * & * & \ldots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \ldots & 0 & * & * & \ldots & * \end{pmatrix}$$

2. Schritt: Steht $a_{i_1 j_1}$ nicht in der ersten Zeile, d. h. $i_1 \neq 1$, dann vertausche die Zeilen a_1 und a_{i_1} - Zeilenoperation vom Typ III. Die so entstandene Matrix heiße $\widetilde{A}_1 = (\tilde{a}_{ij})$. Dann ist $\tilde{a}_{1j_1}$ unser erstes Pivot.

3. Schritt: Erzeuge in der Spalte $\tilde{a}^{j_1}$ von $\widetilde{A}_1$ unterhalb von $\tilde{a}_{1j_1}$ Nullen durch elementare Operationen vom Typ II, d. h. addiere für $k = 2, \ldots, m$ zur k-ten Zeile das $-\frac{\tilde{a}_{kj_1}}{\tilde{a}_{1j_1}}$-fache der ersten Zeile. Die Spalten mit Index kleiner als j_1 werden dadurch nicht geändert. Das Ergebnis ist dann eine Matrix von der Form:

$$A^{(1)} := \left(\begin{array}{cccc|ccc} 0 & \ldots & 0 & a_{1j_1}^{(1)} & * & \ldots & * \\ \hline 0 & \ldots & 0 & 0 & & & \\ \vdots & & \vdots & \vdots & & A_2 & \\ 0 & \ldots & 0 & 0 & & & \end{array}\right),$$

wobei A_2 eine $(m-1)\times(n-j_1)$-Matrix ist, sofern $j_1 \neq n$.

Ist $n - j_1 = 0$ oder $m-1=0$ oder $A^{(2)} = 0$, so sind wir fertig.

Andernfalls ist $A_2 \neq 0$, und wir führen Schritt 1-3 mit A_2 durch. Dabei kann man alle Zeilenoperationen auf die Matrix $A^{(1)}$ ausdehnen, ohne daß sich in den ersten j_1 Spalten etwas ändert, da dort nur Nullen stehen. Ist A_2 umgeformt, so erhält man eine Matrix $A^{(2)}$ der Form:

$$A^{(2)} = \left(\begin{array}{ccccccc|ccc}
0 & \ldots & 0 & a^{(2)}_{1j_1} & * & \ldots & \ldots & \ldots & \ldots & * \\
0 & \ldots & 0 & 0 & \ldots & 0 & a^{(2)}_{2j_2} & * & \ldots & * \\
\hline
0 & \ldots & 0 & 0 & \ldots & 0 & 0 & & & \\
\vdots & & \vdots & \vdots & & \vdots & \vdots & & A_3 & \\
0 & \ldots & 0 & 0 & \ldots & 0 & 0 & & &
\end{array}\right)$$

mit einem Pivot $a^{(2)}_{2j_2}$ und, sofern nicht $m-2=0$ oder $n-j_2=0$, einer Matrix A_3, die eine Zeile und mindestens eine Spalte weniger als A_2 hat.

Ist $A_3 = 0$, so sind wir fertig. ansonsten fahren wir fort wie bisher und erhalten Matrizen $A^{(3)}$, A_4, $A^{(4)}$, etc.. Das Verfahren stoppt, falls nach r-maligem Durchlaufen der Schritte 1-3 entweder $r=m$ oder $r=n$ oder $A_{r+1}=0$. In jedem der drei Fälle ist die Matrix $A^{(r)}$ in ZSF.

Um die Matrix $A^{(r)} = \left(a^{(r)}_{ij}\right)$ in reduzierte ZSF zu bringen, multiplizieren wir zunächst die Zeilen $a^{(r)}_i$, für $i=1,\ldots,r$, mit $\frac{1}{a^{(r)}_{ij_i}}$, was einer elementaren Zeilenoperation vom Typ I entspricht. Die so entstehende Matrix heiße $A' = (a'_{ij})$. Sodann addiert man für $i=1,\ldots,r$ und $k=1,\ldots,i-1$ zur k-ten Zeile das $-a'_{kj_i}$-fache der i-ten Zeile – elementare Operationen vom Typ II – und nennt in jedem Schritt i die neue Matrix wieder A'. Man sieht unmittelbar, daß die entstehende Matrix $A'' = \left(a''_{ij}\right)$ reduzierte ZSF hat, da in Spalte j_i die Elemente a'_{kj_i} in $a''_{kj_i} = 0$, für $k<i$, übergegangen sind. $\square$

Bemerkung 6.8 (Eindeutigkeit der reduzierten Zeilen-Stufen-Form)

a. Der Beweis des Satzes ist konstruktiv, das heißt, aus dem Beweis läßt sich ein Algorithmus zur Berechnung einer ZSF von A herleiten, der sogenannte *Gauß-Algorithmus*.

b. Die reduzierte Zeilenstufenform einer Matrix A ist eindeutig bestimmt, was die Bezeichnung rZSF(A) rechtfertigt.

Beweis der Eindeutigkeit der Zeilenstufenform: Es sei also $A \in \mathrm{Mat}(m \times n, K)$ eine $m \times n$-Matrix.

Da elementare Zeilenoperationen durch Multiplikation mit Elementarmatrizen von links realisiert werden, gilt für eine ZSF B von A, daß es eine invertierbare Matrix $S \in \mathrm{Gl}_m(K)$ gibt mit $B = S \circ A$ (vgl. auch Satz 6.7). Mit Aufgabe 5.40 folgt dann

$\mathrm{ZR}(A) = \mathrm{ZR}(B)$, insbesondere gilt mit Korollar 4.7 also, daß die Nicht-Null-Zeilen von B eine Basis von $\mathrm{ZR}(A)$ bilden, da

$$r := \dim_K \big(\mathrm{ZR}(A)\big) = \mathrm{rang}(A) = \mathrm{rang}(B). \tag{8}$$

Seien nun $B = (b_{ij})$ und $B' = \big(b'_{ij}\big)$ zwei reduzierte ZSF von A mit Zeilenvektoren $b_1, \ldots, b_m$ bzw. $b'_1, \ldots, b'_m$ und Pivotspalten $\{j_1, \ldots, j_r\}$ bzw. $\{j'_1, \ldots, j'_r\}$ - beachte, daß die Anzahl $r = \mathrm{rang}(A)$ nach (8) für beide gleich ist. Wir zeigen nun per Induktion, daß die Zeilen der Matrizen B und B' übereinstimmen.

Induktionsbehauptung: Für $i \in \mathbb{N}$ gilt entweder $i \geq r$ oder $b_{r-i} = b'_{r-i}$, insbesondere also $j_{r-i} = j'_{r-i}$.

Induktionsanfang: $i = 0$. O. E. gelte $j_r \geq j'_r$. Da $b_r \in \mathrm{ZR}(A) = \mathrm{Lin}\,(b'_1, \ldots, b'_r)$, gibt es $\lambda_1, \ldots, \lambda_r \in K$ mit

$$b_r = \sum_{i=1}^{r} \lambda_i b'_i.$$

Insbesondere gilt für $i = 1, \ldots, r-1$

$$0 = b_{rj'_i} = \lambda_i \quad \text{und} \quad b_{rj'_r} = \lambda_r,$$

nach (iv) und (v) in Definition 6.1 angewandt auf die reduzierte ZSF B' mit Pivotspalten $j'_1, \ldots, j'_r$ sowie (i) angewandt auf die ZSF B. Also folgt $b_r = \lambda_r \cdot b'_r$. Da $b_r \neq 0$, muß $\lambda_r \neq 0$ gelten und somit $j'_r = j_r$ wegen (i) in 6.1. Aber dann gilt nach (iv) in 6.1 $1 = b_{rj_r} = \lambda_r$ und somit $b_r = b'_r$.

Induktionsschritt: $0 < i < r-1$ und die Behauptung gelte schon für $0, \ldots, i-1$. O. E. gelte $j_{r-i} \geq j'_{r-i}$. Nach Induktionsvoraussetzung gilt nun $b_{r-i} \in \mathrm{ZR}(A) = \mathrm{Lin}\,\big(b'_1, \ldots, b'_{r-i}, b_{r-i+1}, \ldots, b_r\big)$ also gibt es $\lambda_1, \ldots, \lambda_r \in K$ mit

$$b_{r-i} = \sum_{k=1}^{r-i} \lambda_k b'_k + \sum_{k=r-i+1}^{r} \lambda_k b_k.$$

Insbesondere gilt nach (v) in Definition 6.1, angewandt auf die reduzierte ZSF B, für $k = r-i+1, \ldots, r$

$$0 = b_{r-i\, j_k} = \lambda_k,$$

da $r-i < k$, und (i) angewandt auf B sowie (v) auf B' liefert für $k = 1, \ldots, r-i-1$

$$0 = b_{r-i\, j'_k} = \lambda_k,$$

da $j'_k < j'_{r-i} \leq j_{r-i}$. Insgesamt erhalten wir also wieder

$$b_{r-i} = \lambda_{r-i} b'_{r-i}. \tag{9}$$

Wäre $j_{r-i} > j'_{r-i}$, dann wäre wieder mit (i) $0 = b_{r-i\, j'_{r-i}} = \lambda_{r-i}$ im Widerspruch zu (9) und $b_{r-i} \neq 0$. Also ist $j_{r-i} = j'_{r-i}$ und dann folgt mit (iv) aus 6.1, daß $\lambda_{r-i} = b_{r-i\, j_{r-i}} = 1$, und damit aus (9)) $b_{r-i} = b'_{r-i}$.

Also haben wir mit Induktion gezeigt, daß die Zeilen von B und B' übereinstimmen, d. h. daß die reduzierte Zeilenstufenform von A eindeutig bestimmt ist. $\square$

Beispiel 6.9
Wir überführen nun die folgende Matrix in reduzierte ZSF.

$$\begin{pmatrix} 0 & 0 & -1 & 2 & 3 \\ \underline{-1} & 1 & -3 & 0 & 2 \\ 1 & -1 & 1 & 4 & 3 \end{pmatrix} \xmapsto{\mathrm{I}\leftrightarrow\mathrm{II}} \begin{pmatrix} \underline{-1} & 1 & -3 & 0 & 2 \\ 0 & 0 & -1 & 2 & 3 \\ 1 & -1 & 1 & 4 & 3 \end{pmatrix} \xmapsto{\mathrm{III}\mapsto\mathrm{III}+\mathrm{I}}$$

$$\left(\begin{array}{cc|ccc} -1 & 1 & -3 & 0 & 2 \\ \hline 0 & 0 & \underline{-1} & 2 & 3 \\ 0 & 0 & -2 & 4 & 5 \end{array}\right) \xmapsto{\mathrm{III}\mapsto\mathrm{III}-2\cdot\mathrm{II}} \begin{pmatrix} \underline{-1} & 1 & -3 & 0 & 2 \\ 0 & 0 & \underline{-1} & 2 & 3 \\ 0 & 0 & 0 & 0 & \underline{-1} \end{pmatrix} \xmapsto[\mathrm{III}\mapsto-\mathrm{III}]{\mathrm{I}\mapsto-\mathrm{I},\mathrm{II}\mapsto-\mathrm{II}}$$

$$\begin{pmatrix} 1 & -1 & 3 & 0 & -2 \\ 0 & 0 & 1 & -2 & -3 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \xmapsto[\mathrm{II}\mapsto\mathrm{II}+3\cdot\mathrm{III}]{\mathrm{I}\mapsto\mathrm{I}+2\cdot\mathrm{III}} \begin{pmatrix} 1 & -1 & 3 & 0 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \xmapsto{\mathrm{I}\mapsto\mathrm{I}-3\cdot\mathrm{II}}$$

$$\begin{pmatrix} 1 & -1 & 0 & 6 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Die vierte Matrix besitzt bereits ZSF mit unterstrichenen Pivots, die letzte ist in reduzierter ZSF.

Wir bemerken, daß wir auch auf anderem Weg zum Ziel gekommen wären, und zwar durch andere Wahl der Pivots.

$$\begin{pmatrix} 0 & 0 & -1 & 2 & 3 \\ -1 & 1 & -3 & 0 & 2 \\ \underline{1} & -1 & 1 & 4 & 3 \end{pmatrix} \xmapsto{\mathrm{I}\leftrightarrow\mathrm{III}} \begin{pmatrix} 1 & -1 & 1 & 4 & 3 \\ -1 & 1 & -3 & 0 & 2 \\ 0 & 0 & -1 & 2 & 3 \end{pmatrix} \xmapsto{\mathrm{II}\mapsto\mathrm{II}+\mathrm{I}}$$

$$\left(\begin{array}{cc|ccc} 1 & -1 & 1 & 4 & 3 \\ \hline 0 & 0 & \underline{-2} & 4 & 5 \\ 0 & 0 & -1 & 2 & 3 \end{array}\right) \xmapsto{\mathrm{III}\mapsto\mathrm{III}-\frac{1}{2}\cdot\mathrm{II}} \begin{pmatrix} \underline{1} & -1 & 1 & 4 & 3 \\ 0 & 0 & \underline{-2} & 4 & 5 \\ 0 & 0 & 0 & 0 & \underline{\frac{1}{2}} \end{pmatrix} \xmapsto[\mathrm{III}\mapsto 2\cdot\mathrm{III}]{\mathrm{II}\mapsto-\frac{1}{2}\cdot\mathrm{II}}$$

$$\begin{pmatrix} 1 & -1 & 1 & 4 & 3 \\ 0 & 0 & 1 & -2 & -\frac{5}{2} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \xmapsto[\mathrm{II}\mapsto\mathrm{II}+\frac{5}{2}\cdot\mathrm{III}]{\mathrm{I}\mapsto\mathrm{I}-3\cdot\mathrm{III}} \begin{pmatrix} 1 & -1 & 1 & 4 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \xmapsto{\mathrm{I}\mapsto\mathrm{I}-\mathrm{II}}$$

$$\begin{pmatrix} 1 & -1 & 0 & 6 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

In der Praxis sind 1000×1000-Matrizen keine Seltenheit. Dort wird mit einer festen Stellenzahl gerechnet und deshalb treten bei großen Matrizen unter Umständen erhebliche Rundungsfehler auf. Es kommt der Wahl der richtigen Pivots eine große Bedeutung zu. Ist das gewählte Pivot zu klein, so kann bei Division durch dieses Pivot im dritten Schritt der Rundungsfehler riesig werden - für den Computer bedeutet dies in etwa, als ob man durch Null zu dividieren versuche. Deshalb wählt man in der Praxis das betragsmäßig größte Element als Pivot.

Rechnet man allerdings in Computeralgebrasystemen mit exakter Arithmetik, so spielt die Auslöschung durch Rundungsfehler keine Rolle. Dort muß man eher dafür sorgen, daß die Zahlen, d. h. die Zähler und Nenner, nicht zu groß werden, da dies zu erheblichen Geschwindigkeitsverlusten führen würde.

Wir wollen abschließend den Gauß-Algorithmus in leicht abgewandelter Form als rekursiven Algorithmus zur Bestimmung der reduzierten ZSF einer Matrix formulieren.

Algorithmus 6.10 (Gauß-Algorithmus)
INPUT: $A \in \mathrm{Mat}(m \times n, K)$.

OUTPUT: rZSF(A), die reduzierte Zeilen-Stufen-Form von A.

- **0. Schritt:** Falls $A = 0$, gehe zu Schritt 8.
- **1. Schritt:** Falls $m = 1$, gehe zu Schritt 7.
- **2. Schritt:** Durchlaufe die erste Spalte von oben nach unten, bis ein Element ungleich Null a_{i1} gefunden wurde oder das Ende der Spalte erreicht ist.
- **3. Schritt:** Wurde kein $a_{i1} \neq 0$ gefunden, bilde eine Untermatrix B von A durch Streichen der ersten Spalte von A und gehe zu Schritt 6. Andernfalls, vertausche die Zeilen a_1 und a_i.
- **4. Schritt:** Für $k = 2, \ldots, m$ addiere zur k-ten Zeile das $-\frac{a_{k1}}{a_{11}}$-fache der ersten.
- **5. Schritt:** Falls $n = 1$, gehe zu Schritt 7. Andernfalls bilde eine Untermatrix B von A, durch Streichen der ersten Zeile und der ersten Spalte von A.
- **6. Schritt:** Wende den Algorithmus auf die Untermatrix B an.[2]
- **7. Schritt:** Die Matrix A ist nun in ZSF. Für $i = m$ bis $i = 1$, d. h. rückwärts zählend, durchlaufe die Zeile a_i, beginnend mit der ersten Spalte, bis ein Element $a_{ij} \neq 0$ gefunden wurde oder das Ende der Zeile erreicht ist.
  In letzterem Fall tue nichts, in ersterem multipliziere die Zeile a_i mit $\frac{1}{a_{ij}}$ und addiere für $k = 1, \ldots, i-1$ zur k-ten Zeile das $-a_{kj}$-fache der i-ten Zeile.
- **8. Schritt:** Gib die (veränderte) Matrix A zurück.

A) Algorithmus zur Berechnung des Rangs einer Matrix

Lemma 6.11
Elementare Zeilen- oder Spaltenoperationen ändern den Rang einer Matrix nicht.

[2]Dies ist der Rekursionsschritt, indem der Algorithmus mit einer kleineren Untermatrix aufgerufen wird. Das Ergebnis, das man dabei zurück erhält, wird wieder in die Matrix A eingefügt.

Beweis: Multipliziert man eine Matrix A mit einer invertierbaren Matrix, so erhält man eine äquivalente Matrix. Wegen Proposition 5.29 ändert dies den Rang der Matrix nicht. Da elementare Zeilen- und Spaltenoperationen durch Multiplikation mit invertierbaren Matrizen entstehen, ändern auch diese den Rang der Matrix nicht. □

Algorithmus 6.12 (zur Bestimmung des Rangs)
INPUT: $A \in \mathrm{Mat}(m \times n, K)$.

OUTPUT: $\mathrm{rang}(A)$

1. **Schritt:** Überführe A in ZSF.
2. **Schritt:** Zähle die Anzahl r der Nicht-Nullzeilen in der ZSF.
3. **Schritt:** Gib r zurück.

Beispiel 6.13
In Beispiel 6.9 haben wir eine ZSF berechnet:

$$A := \begin{pmatrix} 0 & 0 & -1 & 2 & 3 \\ -1 & 1 & -3 & 0 & 2 \\ 1 & -1 & 1 & 4 & 3 \end{pmatrix} \longmapsto \quad \ldots \quad \longmapsto \begin{pmatrix} 1 & -1 & 0 & 6 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Die Matrix A hat also Rang 3.

B) Algorithmus zur Berechnung der Inversen einer Matrix

Satz 6.14 (Kriterium für die Invertierbarkeit einer Matrix)
Es sei $A \in \mathrm{Mat}(n, K)$. *Dann sind gleichwertig:*

a. A *ist invertierbar.*

b. $\mathrm{rZSF}(A) = \mathbb{1}_n$.

c. *Es gibt Elementarmatrizen* $T_1, \ldots, T_k \in \mathrm{Mat}(n, K)$ *mit:*

$$T_k \circ \ldots \circ T_1 \circ A = \mathbb{1}_n.$$

d. *Es gibt Elementarmatrizen* $T_1', \ldots, T_k' \in \mathrm{Mat}(n, K)$ *mit:*

$$A = T_1' \circ \ldots \circ T_k'.$$

Insbesondere wird die Gruppe $\mathrm{Gl}_n(K)$ *also von den Elementarmatrizen erzeugt.*

Beweis: Nach Korollar 5.27 gilt, daß A genau dann invertierbar ist, wenn $\mathrm{rang}(A) = n$. Also folgt die Äquivalenz von a.-d. aus Satz 6.7 unter Berücksichtigung von Proposition 6.6. □

Aus Satz 6.14 leitet sich folgendes Verfahren zur Bestimmung der Inversen einer invertierbaren Matrix ab. Hierzu beachte man, daß für Elementarmatrizen $T_1, \ldots, T_k$, für die gilt, daß $T_k \circ \ldots \circ T_1 \circ A = \mathbb{1}_n$, auch gilt, daß

$$T_k \circ \ldots \circ T_1 \circ (A, \mathbb{1}_n) = (\mathbb{1}_n, T_k \circ \ldots \circ T_1) = \big(\mathbb{1}_n, A^{-1}\big).$$

Algorithmus 6.15 (zur Bestimmung der Inversen)

INPUT: $A \in \mathrm{Mat}(n, K)$.

OUTPUT: Inverse von A, falls sie existiert, eine Fehlermeldung sonst.

1. **Schritt:** Erweitere die Matrix A um $\mathbb{1}_n$ zu $C = (A, \mathbb{1}_n) \in \mathrm{Mat}(n \times 2n, K)$.
2. **Schritt:** Überführe C in reduzierte ZSF $C' = \big(A', B\big)$.
3. **Schritt:** Falls rang $\big(A'\big) = n$, dann gib B zurück, sonst eine Fehlermeldung.

Beispiel 6.16

Wir betrachten die 3×3-Matrix

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \in \mathrm{Mat}(3 \times 3, K)$$

und versuchen die Inverse mittels des Algorithmus 6.15 zu bestimmen.

$$\begin{array}{rrr|rrrl}
 & A & & & \mathbb{1}_n & & \\
\hline
1 & 1 & 1 & 1 & 0 & 0 & \\
0 & 1 & 1 & 0 & 1 & 0 & \\
1 & 0 & 1 & 0 & 0 & 1 & \quad \mathrm{III} \mapsto \mathrm{III} - \mathrm{I} \\
\hline
1 & 1 & 1 & 1 & 0 & 0 & \\
0 & 1 & 1 & 0 & 1 & 0 & \\
0 & -1 & 0 & -1 & 0 & 1 & \quad \mathrm{III} \mapsto \mathrm{III} + \mathrm{II} \\
\hline
1 & 1 & 1 & 1 & 0 & 0 & \quad \mathrm{I} \mapsto \mathrm{I} - \mathrm{III} \\
0 & 1 & 1 & 0 & 1 & 0 & \quad \mathrm{II} \mapsto \mathrm{II} - \mathrm{III} \\
0 & 0 & 1 & -1 & 1 & 1 & \\
\hline
1 & 1 & 0 & 2 & -1 & -1 & \quad \mathrm{I} \mapsto \mathrm{I} - \mathrm{II} \\
0 & 1 & 0 & 1 & 0 & -1 & \\
0 & 0 & 1 & -1 & 1 & 1 & \\
\hline
1 & 0 & 0 & 1 & -1 & 0 & \\
0 & 1 & 0 & 1 & 0 & -1 & \\
0 & 0 & 1 & -1 & 1 & 1 &
\end{array}$$

Hieraus ergibt sich gemäß obigem Algorithmus zunächst, daß A invertierbar ist, und ferner, daß

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

C) Algorithmus zur Berechnung der Normalform einer Matrix

Korollar 6.17 (Normalform einer Matrix)
Sei $A \in \mathrm{Mat}(m \times n, K)$ *mit* $r = \mathrm{rang}(A)$*, so läßt sich* A *durch endlich viele elementare Zeilen- und Spaltenoperationen auf die folgende Form bringen:*

$$\left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right). \tag{10}$$

Beweis: Die Aussage folgt aus Korollar 5.32 und Korollar 6.14, da elementare Operationen nach Bemerkung 6.5 durch Multiplikation mit Elementarmatrizen realisierbar sind. □

Wir wollen nun noch an einem Beispiel zeigen, wie man eine Matrix mittels des gaußschen Verfahrens auf Normalform (10) bringt.

Algorithmus 6.18 (Normalform-Algorithmus)
INPUT: $A \in \mathrm{Mat}(m \times n, K)$.

OUTPUT: Normalform $\mathrm{NF}(A)$ von A bezüglich Äquivalenz sowie die zugehörigen Transformationsmatrizen $S \in \mathrm{Gl}_m(K)$ und $T \in \mathrm{Gl}_n(K)$

1. **Schritt:** Überführe A durch elementare Zeilenoperationen in (reduzierte) ZSF und überführe $\mathbb{1}_m$ durch die selben Zeilenoperationen in eine Matrix S.
2. **Schritt:** Überführe A durch elementare Spaltenoperationen in Normalform und überführe $\mathbb{1}_n$ durch die selben Spaltenoperationen in eine Matrix T.
3. **Schritt:** Gib die Normalform von A sowie die Matrizen S und T zurück.

Beispiel 6.19
Durch elementare Zeilen und Spaltenoperationen überführt man A_λ, $\lambda \in K$, in Normalform:

$\mathbb{1}_m$			A_λ			$\mathbb{1}_n$			
1	0	0	1	0	λ	1	0	0	
0	1	0	0	1	0	0	1	0	$\mathrm{ZIII} \mapsto \mathrm{ZIII} - \lambda \cdot \mathrm{ZI}$
0	0	1	λ	0	1	0	0	1	
1	0	0	1	0	λ	1	0	0	
0	1	0	0	1	0	0	1	0	$\mathrm{SIII} \mapsto \mathrm{SIII} - \lambda \cdot \mathrm{SI}$
$-\lambda$	0	1	0	0	$1-\lambda^2$	0	0	1	
1	0	0	1	0	0	1	0	$-\lambda$	falls $\lambda = \pm 1$ fertig,
0	1	0	0	1	0	0	1	0	sonst $\mathrm{SIII} \mapsto \frac{1}{1-\lambda^2} \cdot \mathrm{SIII}$
$-\lambda$	0	1	0	0	$1-\lambda^2$	0	0	1	
1	0	0	1	0	0	1	0	$-\frac{\lambda}{1-\lambda^2}$	
0	1	0	0	1	0	0	1	0	
$-\lambda$	0	1	0	0	1	0	0	$\frac{1}{1-\lambda^2}$	
S			$\mathrm{NF}(A_\lambda)$			T			

Für die Normalform $\mathrm{NF}(A) = SAT$ erhalten wir also

$$\mathrm{NF}(A_\lambda) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\lambda & 0 & 1 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 & -\frac{\lambda}{1-\lambda^2} \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{1-\lambda^2} \end{pmatrix},$$

falls $\lambda \neq \pm 1$, und sonst

$$\mathrm{NF}(A_\lambda) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\lambda & 0 & 1 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 & -\lambda \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Insbesondere gilt, $\mathrm{rang}(A) = 3$ für $\lambda^2 \neq 1$ und $\mathrm{rang}(A) = 2$ sonst.

D) Algorithmus zur Berechnung einer Basis

Der folgende Algorithmus zur Bestimmung einer Basis aus gegebenem Erzeugendensystem beruht auf der Tatsache, daß elementare Zeilenoperationen den Zeilenraum nicht verändern - vgl. Aufgabe 5.40.

Algorithmus 6.20 (Basisberechnung)
INPUT: Ein Erzeugendensystem F des Unterraums $U \subseteq K^n$.

OUTPUT: Eine Basis von U.

1. **Schritt:** Schreibe die Vektoren von F als Zeilen in eine Matrix A und überführe A in Zeilen-Stufen-Form.
2. **Schritt:** Gib die ersten $\mathrm{rang}(A)$ Zeilen als Vektoren zurück.

Beispiel 6.21
Betrachte $U = \mathrm{Lin}\left((1,0,-1,2,3)^t, (1,-1,1,4,3)^t, (0,2,-4,-4,0)^t\right) \leq \mathbb{R}^5$:

$$\begin{pmatrix} 1 & 0 & -1 & 2 & 3 \\ 1 & -1 & 1 & 4 & 3 \\ 0 & 2 & -4 & -4 & 0 \end{pmatrix} \longmapsto \quad \ldots \quad \longmapsto \begin{pmatrix} 1 & 0 & -1 & 2 & 3 \\ 0 & -1 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Also ist $B = \big((1,0,-1,2,3)^t, (0,-1,2,2,0)^t\big)$ eine Basis von U.

E) Algorithmus zum Test auf Injektivität / Surjektivität / Bijektivität

Bemerkung 6.22
Sei $A \in \mathrm{Mat}(m \times n, K)$. Aus der Dimensionsformel für lineare Abbildungen 4.20

$$\dim_K \big(\mathrm{Ker}(f_A)\big) = n - \mathrm{rang}(A)$$

folgt unmittelbar:

- f_A ist injektiv $\iff$ $\mathrm{rang}(A) = n$.
- f_A ist surjektiv $\iff$ $\mathrm{rang}(A) = m$.
- f_A ist bijektiv $\iff$ $\mathrm{rang}(A) = n = m$.

Algorithmus 6.23 (Test auf Injektivität / Surjektivität / Bijektivität)
INPUT: $A \in \mathrm{Mat}(m \times n, K)$.

OUTPUT: Meldung, ob f_A injektiv, surjektiv oder bijektiv ist.

1. **Schritt:** Bestimme den Rang r von A.
2. **Schritt:** Ist $r = m = n$, gib "f_A ist bijektiv" zurück. Ist $r = m < n$, gib "f_A ist surjektiv" zurück. Ist $r = n < m$, gib "f_A ist injektiv" zurück.

Beispiel 6.24
Die zur folgenden Matrix $A \in \mathrm{Mat}(3\times 5, \mathbb{R})$ gehörende Abbildung $f_A : \mathbb{R}^5 \longrightarrow \mathbb{R}^3$ ist weder injektiv noch surjektiv, da $\mathrm{rang}(A) = 2 < 3 = m$ und $\mathrm{rang}(A) = 2 < 5 = n$:

$$A = \begin{pmatrix} 1 & 0 & -1 & 2 & 3 \\ 1 & -1 & 1 & 4 & 3 \\ 0 & 2 & -4 & -4 & 0 \end{pmatrix} \longmapsto \quad \ldots \quad \longmapsto \begin{pmatrix} 1 & 0 & -1 & 2 & 3 \\ 0 & -1 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

F) Algorithmus zur Berechnung der Summe zweier Unterräume

Die Berechnung der Summe zweier Unterräume, die durch Erzeuger gegeben sind, ist einfach, da man nur die Erzeuger der beiden Unterräume vereinigen muß.

Algorithmus 6.25 (Summe zweier Unterräume)
INPUT: Erzeugendensysteme F und G von zwei Unterräumen U und U' des K^n.

OUTPUT: Eine Basis von $U + U'$.

1. **Schritt:** Bilde aus F und G ein Erzeugendensystem und berechne mittels 6.20 eine Basis von $U + U' = \langle F \cup G \rangle$.
2. **Schritt:** Gib diese Basis zurück.

G) Algorithmus zum Testen auf lineare Unabhängigkeit

Da eine endliche Familie von Vektoren genau dann linear unabhängig ist, wenn sie eine Basis ihres Erzeugnisses ist, und da die Dimension des Erzeugnisses einer solchen Familie gerade der Rang der Matrix ist, deren Spalten die Erzeuger sind, liefert Korollar 4.7 den folgenden Algorithmus.

Algorithmus 6.26 (Test auf lineare Unabhängigkeit)
INPUT: Eine Familie F von m Vektoren in K^n.

OUTPUT: Eins, falls F linear unabhängig ist, Null sonst.

1. **Schritt:** Ist F leer, gib Eins zurück, sonst schreibe die Vektoren in F als Spalten in eine Matrix A.
2. **Schritt:** Ist $\mathrm{rang}(A) = m$, so gib Eins zurück, sonst Null.

Ist $f = f_A$ für eine $m \times n$-Matrix A, dann wird das Bild von f von den Spalten von A erzeugt. Wir können eine Basis des Bildes also wie folgt bestimmen.

H) Algorithmus zur Berechnung des Bildes einer linearen Abbildung

Algorithmus 6.27 (Bild von f_A)
INPUT: $A \in \text{Mat}(m \times n, K)$.

OUTPUT: Eine Basis von $\text{Im}(f_A)$.

1. **Schritt:** Transponiere A und überführe die Transponierte in ZSF.
2. **Schritt:** Transponiere das Ergebnis wieder und gib die ersten $\text{rang}(A)$ Spaltenvektoren zurück.

Bemerkung 6.28 (Ringe und Moduln)
Der Gauß-Algorithmus zur Berechnung einer reduzierten ZSF funktioniert über beliebigen kommutativen Ringen mit Eins im allgemeinen nicht mehr, da man dazu teilen muß. Man kann eine Matrix aber auch durch elementare Zeilenoperationen in nicht-reduzierte ZSF überführen, ohne zu teilen. Für manche Fragen ist eine solche ZSF hinreichend, die dann über Ringen berechnet werden kann. Im Übrigen funktioniert auch der Algorithmus zur Berechnung der reduzierten Zeilenstufenform, wenn man zwischendurch nur durch Elemente teilen muß, die ein Inverses im Ring besitzen. Damit kann man bei invertierbaren Matrizen dann z.B. die Inverse berechnen.

Aufgaben

Aufgabe 6.29
Es seien $0 \neq \lambda \in K$, $1 \leq i, j \leq n$ mit $i \neq j$. Dann gelten:

$$Q_i^j(\lambda) = S_j\left(\lambda^{-1}\right) \circ Q_i^j(1) \circ S_j(\lambda),$$

und

$$P_i^j = Q_j^i(1) \circ Q_i^j(-1) \circ Q_j^i(1) \circ S_j(-1).$$

Aufgabe 6.30
Berechne den Rang der folgenden Matrix in Abhängigkeit von a und b:

$$\begin{pmatrix} 0 & b & b & b \\ a & 0 & b & b \\ a & a & 0 & b \end{pmatrix} \in \text{Mat}(3 \times 4, \mathbb{R}).$$

Aufgabe 6.31
Bestimme die Inverse der Matrix

$$\begin{pmatrix} 1 & 3 & -1 & 4 \\ 2 & 5 & -1 & 3 \\ 0 & 4 & -3 & 1 \\ -3 & 1 & -5 & -2 \end{pmatrix} \in \text{Mat}_4(\mathbb{R}).$$

Aufgabe 6.32

Transformiere die folgende Matrix A in Normalform bezüglich Äquivalenz und gib auch die Transformationsmatrizen S und T an:

$$A = \begin{pmatrix} 1 & -2 & 3 & 0 \\ 2 & -7 & 10 & -1 \\ -2 & 4 & -7 & 2 \\ 3 & -5 & 7 & 1 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{R}).$$

Aufgabe 6.33

Überprüfe die folgende Abbildung auf Injektivität und Surjektivität:

$$g : \mathbb{R}^4 \to \mathbb{R}^4 \;:\; \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto \begin{pmatrix} 2x_1 + x_2 + x_4 \\ 3x_1 + x_2 + x_3 + x_4 \\ 2x_1 + x_3 + x_4 \\ 2x_1 + x_2 + x_3 \end{pmatrix}.$$

Aufgabe 6.34

Es sei $U = \{(a_1, \ldots, a_5)^t \in \mathbb{R}^5 \mid a_1 - 2a_2 = 0 = 2a_4 + a_5\} \leq \mathbb{R}^5$. Bestimme die Dimension von U sowie eine Basis von U, die den Vektor $(2, 1, 1, -1, 2)^t$ enthält.

Aufgabe 6.35

Seien $U = \langle (1, 0, 1, 1)^t, (-1, 1, 0, 0)^t \rangle \leq \mathbb{R}^4$ und $U' = \langle (1, 0, 1, 0)^t, (1, 1, 1, 1)^t \rangle \leq \mathbb{R}^4$. Zeige, $\mathbb{R}^4 = U \oplus U'$.

§ 7 Lineare Gleichungssysteme

Definition 7.1 (Lineare Gleichungssysteme)

a. Ein *lineares Gleichungssystem* über K

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

besteht aus m Gleichungen in n *Unbestimmten* oder *Variablen* $x_1, \ldots, x_n$ mit $a_{ij}, b_i \in K$ für $1 \leq i \leq m$ und $1 \leq j \leq n$.
Da sich (LGS) mit $A = (a_{ij})$, $b = (b_1, \ldots, b_m)^t$ und $x = (x_1, \ldots, x_n)^t$ vermittels Matrixmultiplikation auch kurz schreiben läßt als

$$Ax = b,$$

sprechen wir meist von dem *linearen Gleichungssystem* $Ax = b$.

b. Die Matrix

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \in \mathrm{Mat}(m \times n, K)$$

heißt *Koeffizientenmatrix* und der Vektor $b = (b_1, \ldots, b_m)^t \in K^m$ die *Inhomogenität* des Gleichungssystems (LGS). Ferner heißt die Matrix

$$(A \mid b) := \left(\begin{array}{ccc|c} a_{11} & \ldots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \ldots & a_{mn} & b_m \end{array}\right) \in \mathrm{Mat}\big(m \times (n+1), K\big)$$

die *erweiterte Koeffizientenmatrix* von (LGS).

c. Das lineare Gleichungssystem (LGS) heißt *homogen*, falls $b = 0_{K^m}$ der Nullvektor in K^m ist. Ansonsten heißt das System *inhomogen*.

d. Ist ein lineares Gleichungssystem $Ax = b$ gegeben, so heißt das Gleichungssystem $Ax = 0$ (mit $0 = 0_{K^m}$) das *zugehörige homogene Gleichungssystem*.

e. Ein Vektor $c = (c_1, \ldots, c_n)^t \in K^n$ heißt *Lösung* von (LGS), wenn die Gleichung $Ac = b$ erfüllt ist. Die Menge aller Lösungen von (LGS) wird mit

$$\mathrm{Lös}(A, b) := \big\{c \in K^n \mid Ac = b\big\}.$$

bezeichnet.

Bei einem linearen Gleichungssystem sind also Körperelemente a_{ij} und b_i fest vorgegeben, während für die Unbestimmten x_j Körperelemente c_j gesucht werden, die das Gleichungssystem lösen.

Falls $K = \mathbb{R}$, so kann ein lineares Gleichungssystem entweder gar keine, genau eine oder unendlich viele Lösungen haben. Wir werden im Folgenden mehrere Verfahren zur Lösung kennenlernen und uns, im Fall von mehr als einer Lösung, mit der Struktur der Lösungsmenge $\text{Lös}(A, b)$ beschäftigen. Eine wichtige Rolle spielt dabei die lineare Abbildung $f_A : K^n \to K^m$.

Bemerkung 7.2 (Struktur des Lösungraums)
Es sei $A \in \text{Mat}(m \times n, K)$ und $b \in K^m$.

a. Aus den Definitionen folgt unmittelbar

$$\text{Lös}(A, 0) = \{c \in K^n \mid Ac = 0\} = \text{Ker}(f_A),$$

so daß $\text{Lös}(A, 0)$ ein Unterraum des K^n ist mit Dimension

$$\dim_K \big(\text{Lös}(A, 0)\big) = \dim_K \big(\text{Ker}(f_A)\big) = n - \text{rang}(A).$$

Insbesondere ist $Ax = 0$ genau dann eindeutig lösbar, wenn $\text{rang}(A) = n$.

b. Ebenfalls anhand der Definitionen sieht man, daß das lineare Gleichungssystem $Ax = b$ genau dann eine Lösung besitzt, wenn $b \in \text{Im}(f_A) = \{Ac \mid c \in K^n\}$.

Beispiel 7.3
Das lineare Gleichungssystem

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 1 \\ 2x_1 + 3x_2 &= 1 \\ x_2 - x_3 &= 0 \end{aligned}$$

ist inhomogen, hat als Koeffizientenmatrix

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 0 \\ 0 & 1 & -1 \end{pmatrix} \in \text{Mat}(3 \times 3, \mathbb{R}),$$

und als erweiterte Koeffizientenmatrix

$$(A \mid b) = \left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 2 & 3 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{array}\right) \in \text{Mat}(3 \times 4, \mathbb{R}).$$

Die Lösung ist in diesem Fall ganz einfach. Wir erhalten $x_3 = x_2$ aus der 3. Gleichung, $3x_2 = 1 - 2x_1$ aus der 2. und, wenn wir das in die erste Gleichung einsetzen, $x_1 + (1 - 2x_1) = 1$, also $x_1 = 0$. Einsetzen von $x_1 = 0$ in die 2. und 3. Gleichung liefert, daß $\left(0, \frac{1}{3}, \frac{1}{3}\right)^t$ die einzige Lösung ist.

Wir geben zunächst ein Kriterium für die Lösbarkeit eines Gleichungssystems.

Satz 7.4 (Kriterium für die Lösbarkeit eines LGS)
Ein Gleichungssystem $Ax = b$ *ist genau dann lösbar, wenn* $\text{rang}(A) = \text{rang}(A \mid b)$.

Beweis: Wir beachten, daß $\mathrm{Im}(f_A)$ von den Spaltenvektoren $a^1, \ldots, a^n$ von A erzeugt wird, und erhalten deshalb:

$$\begin{aligned}
Ax = b \text{ lösbar } &\Longleftrightarrow b \in \mathrm{Im}(f_A) = \mathrm{Lin}\left(a^1, \ldots, a^n\right) \\
&\Longleftrightarrow b \text{ ist Linearkombination von } a^1, \ldots, a^n \\
&\Longleftrightarrow \mathrm{Im}(f_A) = \mathrm{Lin}\left(a^1, \ldots, a^n\right) = \mathrm{Lin}\left(a^1, \ldots, a^n, b\right) = \mathrm{Im}(f_{(A|b)}) \\
&\Longleftrightarrow \mathrm{rang}(A) = \dim_K\left(\mathrm{Im}(f_A)\right) = \dim_K\left(\mathrm{Im}(f_{(A|b)})\right) = \mathrm{rang}(A \mid b),
\end{aligned}$$

wobei wir für die letzte Äquivalenz berücksichtigen, daß $\mathrm{Im}(f_A) \subseteq \mathrm{Im}(f_{(A|b)})$ gilt. □

Der folgende Satz gibt Auskunft über die Struktur der Lösungsmenge eines linearen Gleichungssystems. Wir haben bereits gesehen, daß diese ein Unterraum ist, wenn das Gleichungssystem homogen ist, und wir werden nun zeigen, daß sie ein affiner Unterraum ist, wenn das Gleichungssystem inhomogen ist.

Satz 7.5 (Struktur von $\mathrm{Lös}(A, b)$ als affiner Raum)
Seien $A \in \mathrm{Mat}(m \times n, K)$, $b \in K^m$ *und sei* $c \in K^n$ *eine Lösung des linearen Gleichungssystems* $Ax = b$. *Dann gilt:*

$$\mathrm{Lös}(A, b) = c + \mathrm{Lös}(A, 0).$$

Beweis: Sei zunächst $y \in \mathrm{Lös}(A, 0)$. Dann gilt:

$$A\big(c + y\big) = Ac + Ay = b + 0 = b,$$

also ist $c + y \in \mathrm{Lös}(A, b)$.

Ist umgekehrt $x \in \mathrm{Lös}(A, b)$. Dann gilt für $y := x - c$

$$Ay = A\big(x - c\big) = Ax - Ac = b - b = 0,$$

also ist $y \in \mathrm{Lös}(A, 0)$. Aber damit ist $x = c + y \in c + \mathrm{Lös}(A, 0)$. □

Wir wollen nun einen Algorithmus kennenlernen, der es uns erlaubt, die Lösung eines linearen Gleichungssystems $Ax = b$ in parametrisierter Form zu bestimmen, d. h. eine spezielle Lösung und eine Basis des Lösungsraumes des zugehörigen homogenen Gleichungssystems zu berechnen. Der wichtigste Schritt ist hierbei die Überführung der erweiterten Koeffizientenmatrix $(A \mid b)$ in reduzierte Zeilen-Stufen-Form.

Lemma 7.6 (Elementare Zeilenoperationen ändern den Lösungsraum nicht.)
Sind $A, A' \in \mathrm{Mat}(m \times n, K)$ *und* $b, b' \in K^m$ *und entsteht die Matrix* $(A' \mid b')$ *aus* $(A \mid b)$ *durch elementare Zeilenoperationen, so gilt*

$$\mathrm{Lös}(A, b) = \mathrm{Lös}\left(A', b'\right).$$

Beweis: Daß $(A' \mid b')$ aus $(A \mid b)$ durch elementare Zeilenoperationen hervorgeht, bedeutet, daß es eine invertierbare Matrix $S \in \mathrm{Gl}_m(K)$ gibt mit $A' = SA$ und $b' = Sb$.

Ist nun $c \in$ Lös(A, b), dann gilt $Ac = b$ und damit

$$b' = Sb = SAc = A'c.$$

Also ist $c \in$ Lös $\left(A', b'\right)$.

Ist andererseits $c \in$ Lös $\left(A', b'\right)$, dann gilt $A'c = b'$ und damit

$$b = S^{-1}b' = S^{-1}A'c = Ac.$$

Also ist $c \in$ Lös(A, b). □

Bemerkung 7.7
Aus Lemma 7.6 und Satz 6.7 folgt, daß wir die erweiterte Koeffizientenmatrix eines Gleichungssystems $Ax = b$ mit $A \in \mathrm{Mat}(m \times n, K)$ und $b \in K^m$ mittels Gauß-Algorithmus in (reduzierte) ZSF überführen können, ohne daß sich die Lösungsmenge ändert.

Wir betrachten deshalb den Fall, daß die Matrix A in ZSF gegeben ist, näher.

Satz 7.8 (Lösbarkeitskriterium für ein LGS mittels Gauß-Algorithmus)
Sei $A \in \mathrm{Mat}(m \times n, K)$ *eine Matrix in Zeilen-Stufen-Form und* $b \in K^m$. *Die erweiterte Koeffizientenmatrix habe die Gestalt*

$$(A \mid b) = \left(\begin{array}{cccccccccccccc|c} 0 & \ldots & 0 & a_{1j_1} & * & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & * & b_1 \\ 0 & \ldots & \ldots & 0 & \ldots & 0 & a_{2j_2} & * & \ldots & \ldots & \ldots & \ldots & * & b_2 \\ \vdots & & & & & & & \ddots & & & & & \vdots & \vdots \\ 0 & \ldots & \ldots & \ldots & \ldots\ldots & \ldots & \ldots & 0 & a_{rj_r} & * & \ldots & \ldots & * & b_r \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & b_{r+1} \\ \vdots & & & & & & & & & & & & \vdots & \vdots \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & b_m \end{array}\right) \tag{11}$$

mit Pivots $a_{ij_i} \neq 0$ *für* $i = 1, \ldots, r$. *Dann gilt:*

a. $Ax = b$ *ist genau dann lösbar, wenn* $b_{r+1} = \ldots = b_m = 0$.

b. *Sind* $b_{r+1} = \ldots = b_m = 0$ *und gilt* $r = n$, *so besitzt* $Ax = b$ genau eine *Lösung.*

c. *Sind* $b_{r+1} = \ldots = b_m = 0$ *und ist* $r < n$, *so hat* $Ax = b$ *mehr als eine Lösung. Genauer* Lös$(A, b) = c +$ Lös$(A, 0)$, *wobei* c *eine spezielle Lösung ist und* Lös$(A, 0)$ *die Dimension* $n - r$ *hat.*

Beweis: Die Aussagen folgen aus Satz 7.4, Satz 7.5 und Bemerkung 7.2. □

Bemerkung 7.9 (Parametrisierung von Lös(A, b))
Wir wollen nun angeben, wie man im Fall c. aus Satz 7.8 aus der Zeilen-Stufen-Form (11) von A die *Parametrisierung von* Lös(A, b) als sogenannte *affine* Abbildung

$$\phi : K^{n-r} \to \text{Lös}(A, b)$$

herleitet. Sei hierzu $A = \mathrm{rZSF}(A)$ in *reduzierter* ZSF gegeben.

Die Parameter $x_{j_1}, \ldots, x_{j_r}$ nennen wir die *gebundenen Parameter* und die x_j mit $j \in I := \{1, \ldots, n\} \setminus \{j_1, \ldots, j_r\}$ die *freien Parameter*. Dies rührt daher, daß sich aus (11) für eine Lösung x ergibt

$$x_{j_i} = b_i - \sum_{j \in I} a_{ij} x_j, \quad i = 1, \ldots, r. \tag{12}$$

D. h. die gebundenen Parameter hängen von den freien Parametern ab.

Identifizieren wir K^{n-r} nun mit K^I und schreiben somit $y = (y_j \mid j \in I)$ für einen Vektor $y \in K^{n-r}$, dann ergibt sich die Parametrisierung hieraus als

$$\phi : K^{n-r} \to \text{Lös}(A, b) : y \mapsto c + f(y), \tag{13}$$

wobei

$$c_j = \begin{cases} 0, & \text{falls } j \in I, \\ b_i, & \text{falls } j = j_i, \end{cases} \tag{14}$$

und

$$f : K^{n-r} \to K^n : y \mapsto (z_1, \ldots, z_n)^t, \tag{15}$$

mit

$$z_j = \begin{cases} y_j, & \text{falls } j \in I, \\ -\sum_{k \in I} a_{ik} y_k, & \text{falls } j = j_i. \end{cases} \tag{16}$$

Damit ist f eine lineare Abbildung und deshalb nennt man ϕ affin.

Man beachte, daß c in diesem Fall eine spezielle Lösung von $Ax = b$ ist, während $\text{Im}(f) = \text{Lös}(A, 0)$.

A) Der Gauß-Algorithmus zur Lösung eines (LGS)

Algorithmus 7.10 (Algorithmus zur Lösung eines LGS)

INPUT: Die erweiterte Koeffizientenmatrix $(A \mid b)$ eines LGS $Ax = b$.

OUTPUT: Eine spezielle Lösung c von $Ax = b$ und eine Basis B von $\text{Lös}(A, 0)$, sofern das Gleichungssystem lösbar ist.

1. **Schritt:** Berechne eine reduzierte Zeilen-Stufen-Form $(A' \mid b')$ von $(A \mid b)$ mit $r = \text{rang}(A')$.
2. **Schritt:** Ist $b'_{r+1} \neq 0$, dann ist das LGS nicht lösbar.
3. **Schritt:** Überführe $(A' \mid b')$ in eine $n \times (n+1)$-Matrix $(A'' \mid b'')$ durch Einfügen und Streichen von Nullzeilen, so daß die Pivotelemente anschließend auf der Diagonale der Matrix A'' stehen.
4. **Schritt:** Ersetze jede Null auf der Diagonale von A'' durch -1.
5. **Schritt:** Die spezielle Lösung ist $c := b''$ und die Spalten von A'', die eine -1 auf der Diagonale haben, sind eine Basis von $\text{Lös}(A, 0)$.

Beispiel 7.11
Wir betrachten das Gleichungssystem:

$$\begin{aligned} x_1 + x_2 + x_3 - 2x_4 &= 1 \\ x_1 + x_2 - x_3 &= -1 \\ 3x_1 + 3x_2 + x_3 - 4x_4 &= 1 \end{aligned} \tag{17}$$

In Matrixschreibweise lautet das Gleichungssystem:

$$\begin{pmatrix} 1 & 1 & 1 & -2 \\ 1 & 1 & -1 & 0 \\ 3 & 3 & 1 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

Durch den Gauß-Algorithmus überführen wir die erweiterte Koeffizientenmatrix in reduzierte Zeilen-Stufen-Form:

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

Wir sehen, daß $\operatorname{rang}(A) = \operatorname{rang}(A \mid b) = 2$, so daß das Gleichungssystem lösbar ist.

Um die Lösung zu berechnen, fügen wir als zweite Zeile eine Nullzeile ein, um eine 4×5-Matrix zu erzeugen und die Pivotelemente auf der Diagonalen zu haben, und ersetzen die Nullen auf der Diagonalen anschließend durch -1:

$$\left(\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \mapsto \left(\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \mapsto \left(\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & -1 & 0 \end{array}\right).$$

Damit erhalten wir die letzte Spalte

$$c = (0, 0, 1, 0)^t$$

als spezielle Lösung von (17) und die Spalten 2 und 4 als Basis

$$B = \big((1, -1, 0, 0)^t, (-1, 0, -1, -1)^t\big)$$

des Lösungsraums $\text{Lös}(A, 0)$ des homogenen Gleichungssystems $Ax = 0$. Insgesamt gilt damit

$$\text{Lös}(A, b) = c + \text{Lös}(A, 0) = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + s \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} + t \cdot \begin{pmatrix} -1 \\ 0 \\ -1 \\ -1 \end{pmatrix} \;\middle|\; s, t \in K \right\}.$$

Wollte man eine Parametrisierung wie in Bemerkung 7.9 angeben, so erhält man

$$\phi : K^2 \to \text{Lös}(A, b) \subset K^4 : \begin{pmatrix} x_2 \\ x_4 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} x_2 - x_4 \\ -x_2 \\ -x_4 \\ -x_4 \end{pmatrix}.$$

Wir wollen nun einige Algorithmen angeben, denen der Algorithmus zur Lösung eines linearen Gleichungssystems zugrunde liegt.

B) Algorithmus zur Berechnung des Kerns einer linearen Abbildung

Ist $f = f_A$ für eine $m \times n$-Matrix A, dann ist der Kern von f gerade die Lösungsmenge $\text{Lös}(A, 0)$ des homogenen Gleichungssystems $Ax = 0$.

Algorithmus 7.12 (Kern von f_A)
INPUT: $A \in \text{Mat}(m \times n, K)$.

OUTPUT: Eine Basis von $\text{Ker}(f_A)$.

1. **Schritt:** Bestimme eine Lösung (c, B) von $Ax = 0$ gemäß 7.10.
2. **Schritt:** Gib B als Basis zurück.

Beispiel 7.13
Wir wollen den Kern der Linearen Abbildung $f_A : K^4 \longrightarrow K^3$ berechnen, die durch die Koeffizientenmatrix

$$A = \begin{pmatrix} 1 & 1 & 1 & -2 \\ 1 & 1 & -1 & 0 \\ 3 & 3 & 1 & -4 \end{pmatrix}$$

in Beispiel 7.11 gegeben ist. Dann gehen wir wie in Beispiel 7.11 vor, wobei wir die Inhomogenität durch den Nullvektor ersetzen oder einfach gänzlich ignorieren können. Die Rechnungen ändern sich nicht und wir erhalten wie dort

$$B = \big((1, -1, 0, 0)^t, (-1, 0, -1, -1)^t\big)$$

als Basis von $\text{Ker}(f_A) = \text{Lös}(A, 0)$.

C) Algorithmus zur Berechnung einer Transformationsmatrix $T^B_{B'}$

Sind $B = (b_1, \ldots, b_n)$ und $B' = (b'_1, \ldots, b'_n)$ zwei Basen des K^n und wollen wir die Transformationsmatrix $T^B_{B'}$ bestimmen, so müssen wir die Basisvektoren in B als Linearkombination der Basisvektoren in B' darstellen und die so erhaltenen Koeffizienten liefern die Spalten von $T^B_{B'}$. Wir müssen also n Gleichungsysteme Lösen, bei denen die Koeffizientenmatrix stets $b'_1, \ldots, b'_n$ als Spaltenvektoren hat und bei denen die Inhomogenitäten durch die Vektoren $b_1, \ldots, b_n$ gegeben werden. Da die Koeffizientenmatrix sich nicht ändert, können wir die n Gleichungssysteme simultan lösen, indem wir der erweiterten Koeffizientenmatrix gleich alle Vektoren $b_1, \ldots, b_n$ als zusätzliche Spalten anhängen.

Algorithmus 7.14 (Transformationsmatrix $T^B_{B'}$)

INPUT: Zwei Basen $B = (b_1, \ldots, b_n)$ und $B' = (b'_1, \ldots, b'_n)$ im K^n.

OUTPUT: Die Transformationsmatrix $T^B_{B'}$.

1. **Schritt:** Schreibe die Vektoren $b'_1, \ldots, b'_n, b_1, \ldots, b_n$ in dieser Reihenfolge als Spalten in eine Matrix A.
2. **Schritt:** Bringe A auf reduzierte ZSF.
3. **Schritt:** Die letzten n Spalten von rZSF(A) sind $T^B_{B'}$.

Beispiel 7.15

Seien die zwei Basen $B = \big((1,1)^t, (1,-1)^t\big)$ und $B' = \big((1,2)^t, (-1,0)^t\big)$ des $\mathbb{R}^2$ gegeben.

B'		B	
1	-1	1	1
2	0	1	-1
1	-1	1	1
0	2	-1	-3
1	-1	1	1
0	1	$-\frac{1}{2}$	$-\frac{3}{2}$
1	0	$\frac{1}{2}$	$-\frac{1}{2}$
0	1	$-\frac{1}{2}$	$-\frac{3}{2}$
$\mathbb{1}_2$		$T^B_{B'}$	

D) Algorithmus zur Berechnung einer Matrixdarstellung $M^B_D(f)$

Wir wollen hier angeben, wie man die Matrixdarstellung einer linearen Abbildung $f : K^n \longrightarrow K^m$ bezüglich zweier Basen $B = (b_1, \ldots, b_n)$ von K^n und $D = (d_1, \ldots, d_m)$ von K^m berechnet. Die Grundidee ist ähnlich wie beim Algorithmus zur Berechnung der Transformationsmatrix.

Algorithmus 7.16 (Matrixdarstellung $M^B_D(f)$)

INPUT: Eine lineare Abbildung $f : K^n \longrightarrow K^m$, eine Basis $B = (b_1, \ldots, b_n)$ von K^n und eine Basis $D = (d_1, \ldots, d_m)$ im K^m.

OUTPUT: Die Matrixdarstellung $M^B_D(f)$.

1. **Schritt:** Schreibe die Vektoren $d_1, \ldots, d_m, f(b_1), \ldots, f(b_n)$ in dieser Reihenfolge als Spalten in eine Matrix A.
2. **Schritt:** Bringe A auf reduzierte ZSF.
3. **Schritt:** Die letzten n Spalten von rZSF(A) sind $M^B_D(f)$.

Beispiel 7.17

Für die Basen $B = \big((1,0,1)^t, (1,1,0)^t, (0,0,1)^t\big)$ des K^3 und $D = \big((1,1)^t, (1,-1)^t\big)$ des K^2 sowie die lineare Abbildung

$$f : K^3 \longrightarrow K^2 : (x,y,z)^t \mapsto (x+y+z, x-z)^t$$

wollen wir die Matrixdarstellung $M_D^B(f)$ berechnen.

$$\begin{array}{cc|ccc}
\multicolumn{2}{c|}{D} & \multicolumn{3}{c}{f(B)} \\ \hline
1 & 1 & 2 & 2 & 1 \\
1 & -1 & 0 & 1 & -1 \\ \hline
1 & 1 & 2 & 2 & 1 \\
0 & -2 & -2 & -1 & -2 \\ \hline
1 & 1 & 2 & 2 & 1 \\
0 & 1 & 1 & \frac{1}{2} & 1 \\ \hline
1 & 0 & 1 & \frac{3}{2} & 0 \\
0 & 1 & 1 & \frac{1}{2} & 1 \\ \hline
\multicolumn{2}{c|}{\mathbb{1}_2} & \multicolumn{3}{c}{M_D^B(f)}
\end{array}$$

Bemerkung 7.18
Natürlich könnte man auch zunächst die Matrixdarstellung $M_F^E(f)$ bezüglich der kanonischen Basen berechnen, da man dazu einfach die Vektoren $f(e_i)$ in die Spalten der Matrix schreiben muß. Analog erhält man T_E^B, indem man die Vektoren von B in die Spalten der Matrix schreibt. Dann muß man nur noch T_D^F mit Hilfe des Algorithmus' zur Berechnung einer Transformationsmatrix bestimmen und kann die Matrizen multiplizieren, um $M_D^B(f)$ zu erhalten.

E) Algorithmus zum Austauschverfahren von Steinitz

Beim Austauschsatz von Steinitz müssen wir die Vektoren in $F = (y_1, \ldots, y_r)$, die wir in die Basis $B = (x_1, \ldots, x_n)$ hineintauschen wollen, sukzessive als Linearkombination der Basisvektoren in (dem veränderten) B darstellen, d.h. wir müssen immer wieder lineare Gleichungssysteme lösen.

Algorithmus 7.19 (Austauschverfahren von Steinitz)
INPUT: Eine Basis $B = (x_1, \ldots, x_n)$ und eine linear unabhängige Familie $F = (y_1, \ldots, y_r)$ von Vektoren in $V = \mathrm{Lin}(B) \subseteq K^n$.
OUTPUT: Eine Basis B' von V, die F enthält.

1. **Schritt:** Für $i = 1, \ldots, r$ tue:
   - Schreibe die Vektoren in B als Spalten in eine Matrix A.
   - Bilde die erweiterte Matrix (A, y_i).
   - Überführe (A, y_i) in reduzierte Zeilen-Stufen-Form und suche in der letzten Spalte den ersten Eintrag ungleich Null.
   - Streiche den entsprechenden Vektor aus B und füge y_i als letzten Vektor in B ein.
2. **Schritt:** Gib B zurück.

Beispiel 7.20
Betrachte die linear unabhängige Familie $F = (y_1, y_2) = \big((1,2,1)^t, (1,2,2)^t\big)$ und

die Basis $B = (x_1, x_2, x_3) = \big((1,1,0)^t, (1,0,1)^t, (0,0,1)^t\big)$. Wir wollen nun F in B hineintauschen.

Wir bilden die erweiterte Matrix (A, y_1) und überführen sie in reduzierte ZSF:

$$(A, y_1) = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 1 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

Da der erste Eintrag in der letzten Spalte nicht Null ist, streichen wir aus B den Vektor x_1 und fügen y_1 als letzten Vektor ein. Wir erhalten die neue Basis

$$B = (x_2, x_3, y_1).$$

Dann bilden wir wieder die erweiterte Matrix (A, y_2) und überführen sie in reduzierte ZSF:

$$(A, y_2) = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 1 & 1 & 1 & 2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Der erste Eintrag der letzten Spalte, der nicht Null ist, ist der zweite, mithin müssen wir den zweiten Vektor in B streichen, das ist x_3, und fügen y_2 am Ende ein. Wir erhalten die Basis

$$B = (x_2, y_1, y_2).$$

Bemerkung 7.21 (Berechnung eines Komplementes oder einer Basis für K^n/U)
Will man ein Komplement eines Unterraums U in K^n berechnen, so berechnet man zunächst eine Basis von U und tauscht diese anschließend mit Steinitz in die kanonische Basis von K^n. Die verbleibenden Vektoren der kanonischen Basis sind dann eine Basis für ein Komplement und zugleich sind deren Restklassen eine Basis für den Faktorraum K^n/U (siehe Bemerkung 4.19). Der obige Algorithmus erlaubt also auch die Berechnung eine Komplementes und einer Basis eines Faktorraums.

F) Algorithmus zur Berechnung der Gleichungen eines Unterraumes

Wir haben gesehen, daß Unterräume des K^n als Lösungsmengen von homogenen linearen Gleichungssystemen auftauchen. Um etwa den Schnitt zweier Unterräume des K^n zu bestimmen, ist es nützlich, aus dem Erzeugendensystem eines Unterraumes ein Gleichungssystem bestimmen zu können, das den Unterraum beschreibt.

Algorithmus 7.22 (Gleichungen eines Unterraumes)
INPUT: Eine Familie $F = (x_1, \ldots, x_m)$ von Vektoren im K^n.

OUTPUT: Eine Matrix $A \in \mathrm{Mat}(k \times n, K)$ mit $\mathrm{Lös}(A, 0) = \mathrm{Lin}\,(F)$.

1. **Schritt:** Schreibe die Vektoren aus F als Zeilen in eine Matrix $B \in \mathrm{Mat}(m \times n, K)$ und bestimme eine Basis $(y_1, \ldots, y_k)$ von $\mathrm{Ker}(f_B) = \mathrm{Lös}(B, 0)$.
2. **Schritt:** Schreibe die $y_1, \ldots, y_k$ als Zeilenvektoren in eine Matrix A.
3. **Schritt:** Gib A zurück.

Beispiel 7.23
Finde ein lineares Gleichungssystem $Ax = 0$ mit Lösungsmenge

$$\text{Lös}(A, 0) = \text{Lin}\left((1, 2, 1)^t, (0, 1, 0)^t\right) \leq \mathbb{R}^3.$$

Dazu Bilden wir die 2×3-Matrix

$$B = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

und berechnen ihren Kern:

$$\text{rZSF}(B, 0) = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

Somit ist der Vektor in der dritten Spalte eine Basis von $\text{Lös}(B, 0)$ und wir erhalten

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

G) Algorithmus zur Berechnung des Durchschnitts zweier Unterräume

Abschließend sind wir nun in der Lage, einen Algorithmus anzugeben, mittels dessen sich eine Basis des Schnitts zweier Unterräume des K^n ermitteln läßt.

Algorithmus 7.24 (Durchschnitt zweier Unterräume)
INPUT: Zwei Familien F und G von Vektoren in K^n.

OUTPUT: Eine Basis des Schnitts von $\text{Lin}(F)$ und $\text{Lin}(G)$.

1. **Schritt:** Bestimme Matrizen A und A' gemäß 7.22, so daß $\text{Lin}(F) = \text{Lös}(A, 0)$ und $\text{Lin}(G) = \text{Lös}\left(A', 0\right)$.
2. **Schritt:** Bilde aus den Zeilen von A und A' eine gemeinsame Matrix A''.
3. **Schritt:** Bestimme eine Basis B von $\text{Ker}(f_{A''}) = \text{Lös}\left(A'', 0\right)$ gemäß 7.12 und gib B zurück.

Beispiel 7.25
Wir wollen den Durchschnitt der Unterräume

$$U = \text{Lin}\left((1, 2, 1)^t, (0, 1, 0)^t\right)$$

und

$$U' = \{(x, y, z)^t \in \mathbb{R}^3 \mid x + y + z = 0\}$$

berechnen. Der zweite Unterraum ist bereits als Lösungsmenge eines Gleichungssystems mit Koeffizientenmatrix $A' = (1\ 1\ 1)$ gegeben. Für den ersten Unterraum haben wir eine solche Darstellung $\text{Lös}(A, 0)$ bereits in Beispiel 7.23 berechnet. Wir bilden eine neue Matrix A''

$$A'' = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

aus A und A' und lösen das zugehörige homogene Gleichungssystem

$$(A'', 0) = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

Damit ist die dritte Spalte

$$B = \left((-1, 2, -1)^t\right)$$

eine Basis von $U \cap U'$.

H) Beispiele linearer Gleichungssysteme in Anwendung

Wir geben jetzt einige Beispiele von Gleichungssystemen, die zum Teil aus Anwendungen kommen. Wir werden diese nicht in der Vorlesung besprechen. Sie sollen dem interessierten Leser die große praktische Bedeutung linearer Gleichungssysteme illustrieren.

Beispiel 7.26 (Wie alt ist der Vater?)
Ein Vater hat einen Sohn und eine Tochter. Der Vater ist viermal so alt wie sein Sohn und der Sohn ist fünf Jahre älter als seine Schwester. In fünf Jahren sind Vater und Sohn zusammen sechsmal so alt wie die Tochter.

Wie alt sind Vater, Sohn und Tochter?

Das lineare Gleichungssystem mit v = Alter des Vaters, s = Alter des Sohnes, und t = Alter der Tochter lautet:

$$v = 4s, \;\; s = t + 5, \;\; (v + 5) + (s + 5) = 6(t + 5).$$

Das Gleichungssystem schreiben wir systematisch folgendermaßen auf:

$$\begin{aligned} v - 4s + 0 \cdot t &= 0, \\ 0 \cdot v + s - t &= 5, \\ v + s - 6t &= 20. \end{aligned}$$

Dies ist ein lineares Gleichungssystem in den Unbestimmten v, s, t.

Die Lösung mit Hilfe des Gaußschen Algorithmus geht wie folgt:

$$\left(\begin{array}{rrr|r} 1 & -4 & 0 & 0 \\ 0 & 1 & -1 & 5 \\ 1 & 1 & -6 & 20 \end{array}\right) \mapsto \left(\begin{array}{rrr|r} 1 & -4 & 0 & 0 \\ 0 & 1 & -1 & 5 \\ 0 & 5 & -6 & 20 \end{array}\right) \mapsto \left(\begin{array}{rrr|r} 1 & -4 & 0 & 0 \\ 0 & 1 & -1 & 5 \\ 0 & 0 & 1 & 5 \end{array}\right)$$

$$\mapsto \left(\begin{array}{rrr|r} 1 & -4 & 0 & 0 \\ 0 & 1 & 0 & 10 \\ 0 & 0 & 1 & 5 \end{array}\right) \mapsto \left(\begin{array}{rrr|r} 1 & 0 & 0 & 40 \\ 0 & 1 & 0 & 10 \\ 0 & 0 & 1 & 5 \end{array}\right)$$

Als Lösung erhalten wir also: $t = 5, \;\; s = 10, \;\; v = 40$, d. h. der Vater ist 40 Jahre alt, sein Sohn zehn und seine Tochter fünf.

Beispiel 7.27 (Schnitt zweier Ebenen)
Wir definieren eine *Ebene* im $\mathbb{R}^3$ als Lösungsmenge einer linearen Gleichung

$$E : a_1x_1 + a_2x_2 + a_3x_3 = b$$

mit $a_1, a_2, a_3, b \in \mathbb{R}$ und $a_i \neq 0$ für mindestens ein i.

Dies stimmt mit der Anschauung überein (sind alle a_i und b gleich 0, so erhalten wir als Lösungsmenge den ganzen $\mathbb{R}^3$, sind alle $a_i = 0$ und $b \neq 0$, so ist die Lösungsmenge leer).

Um den Schnitt der beiden Ebenen, die durch die Gleichungen $E_1 : x_1+x_2+2x_3 = 2$ und $E_2 : x_1+x_3 = 4$ gegeben sind, zu bestimmen, müssen wir also das Gleichungssystem aus diesen beiden Gleichungen lösen, wobei wir wie in Abschnitt A) beschrieben vorgehen:

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 2 \\ 1 & 0 & 1 & 4 \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & 1 & -2 \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & -1 & 0 \end{array}\right).$$

Wir erhalten als Lösungsmenge

$$E_1 \cap E_2 = \begin{pmatrix} 4 \\ -2 \\ 0 \end{pmatrix} + \mathbb{R} \cdot \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

Dies ist offensichtlich die Parameterdarstellung einer Geraden im $\mathbb{R}^3$ durch die Punkte $(4, -2, 0)^t$ und $(5, -1, -1)^t$.

Beispiel 7.28 (Schnitt zweier Ebenen)
Im allgemeinen werden sich zwei Ebenen, E_1, E_2, im $\mathbb{R}^3$ in einer Geraden schneiden, in Spezialfällen können die Ebenen aber parallel sein ($E_1 \cap E_2 = \emptyset$) oder übereinstimmen ($E_1 = E_2$).

Sei E_1 die Ebene

$$E_1 : x_1 + x_2 + 2x_2 = 3$$

und E_2 eine beliebige Ebene

$$E_2 : a_1x_1 + a_2x_2 + a_3x_3 = b.$$

Wir wollen feststellen für welche a_1, a_2, a_3, b entweder $E_1 \cap E_2$ eine Gerade, leer oder E_1 ist:

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ a_1 & a_2 & a_3 & b \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & a_2 - a_1 & a_3 - 2a_1 & b - 3a_1 \end{array}\right).$$

Die letzte Gleichung lautet

$$(a_2 - a_1)x_2 + (a_3 - 2a_1)x_3 = b - 3a_1.$$

Ein wenig Überlegung liefert (da die Lösungsmenge der ersten Gleichung E_1 ist, und da die Lösungsmenge der zweiten Gleichung unabhängig von x_1 ist):

$$E_1 \cap E_2 = \emptyset \quad \Leftrightarrow \quad a_2 - a_1 = a_3 - 2a_1 = 0, (b - 3a_1) \neq 0, \tag{18}$$

$$E_1 = E_2 \quad \Leftrightarrow \quad a_2 - a_1 = a_3 - 2a_1 = b - 3a_1 = 0. \tag{19}$$

In allen anderen Fällen ist $E_1 \cap E_2$ eine Gerade.

Im Fall $E_1 = E_2$ haben wir wieder ein Gleichungssystem (19) mit drei Gleichungen in den vier Unbestimmten a_1, a_2, a_3, b zu lösen:

$$\left(\begin{array}{cccc|c} -1 & 1 & 0 & 0 & 0 \\ -2 & 0 & 1 & 0 & 0 \\ -3 & 0 & 0 & 1 & 0 \end{array}\right) \mapsto \left(\begin{array}{cccc|c} -1 & 1 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 \\ 0 & -3 & 0 & 1 & 0 \end{array}\right) \mapsto \left(\begin{array}{cccc|c} -1 & 1 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 \\ 0 & 0 & -\frac{3}{2} & 1 & 0 \end{array}\right)$$

$$\mapsto \left(\begin{array}{cccc|c} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & -\frac{2}{3} & 0 \end{array}\right) \mapsto \left(\begin{array}{cccc|c} 1 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 1 & 0 & -\frac{1}{3} & 0 \\ 0 & 0 & 1 & -\frac{2}{3} & 0 \end{array}\right)$$

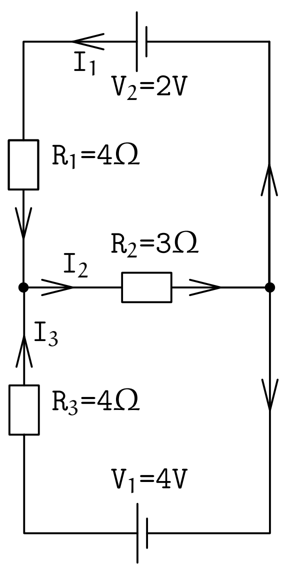
Als Lösung ergibt sich $a_1 = -\frac{b}{3}$, $a_2 = \frac{b}{3}$ und $a_3 = \frac{2b}{3}$, oder kurz

$$(a_1, a_2, a_3, b) = t \cdot (-1, 1, 2, 3)$$

mit $t \in \mathbb{R}$ beliebig. Daraus können wir aber alle drei Fälle ablesen:

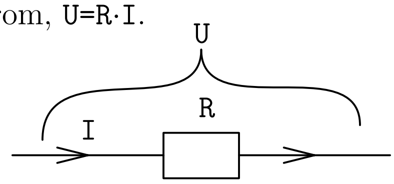
$E_1 = E_2$ genau dann, wenn die Gleichung von E_2 ein Vielfaches $\neq 0$ der Gleichung von E_1 ist; $E_1 \cap E_2 = \emptyset$ genau dann, wenn der Koeffizientenvektor (a_1, a_2, a_3) ein Vielfaches $\neq 0$ des Koeffizientenvektors von E_1 ist, aber die rechte Seite b von E_2 nicht das gleiche Vielfache der rechten Seite von E_1 ist; und $E_1 \cap E_2$ ist eine Gerade in allen anderen Fällen.

Beispiel 7.29 (Elektrische Netzwerke)
In einem einfachen elektrischen Netzwerk, wie z. B.

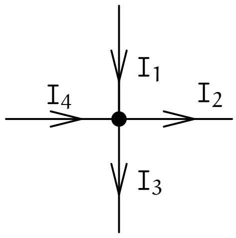


bezeichnet man mit U die Spannung, mit I den Strom und mit R den Widerstand, gemessen in Volt (V), Ampere (A) und Ohm (Ω) respektive. Dabei gelten folgende Gesetze:

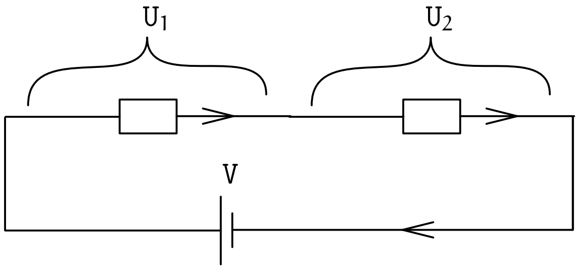
- *Ohmsches Gesetz*: Der Spannungsabfall über einen Widerstand ist das Produkt von Widerstand und Strom, U=R·I.



- *1. Kirchhoffsches Gesetz (Knotengleichung):* Die Summe der in einen Knoten hineinfließenden Ströme ist gleich der Summe der hinausfließenden Ströme. Beispiel: $I_1+I_4=I_2+I_3$



- *2. Kirchhoffsches Gesetz (Maschengleichung):* Die Summe der Spannungsverluste in einem geschlossenen Kreis ist gleich der Gesamtspannung in einem Kreis. Beispiel: $V=U_1+U_2$



Im obigen Beispiel stellt man mit Hilfe der drei Gesetze das folgende lineare Gleichungssystem auf:

$$
\begin{array}{rcll}
I_1 + I_3 & = & I_2, & \text{(Knotengleichung)} \\
4I_1 + 3I_2 & = & 2, & \text{(1. Maschengleichung)} \\
4I_3 + 3I_2 & = & 4. & \text{(2. Maschengleichung)}
\end{array}
$$

Wir erhalten das folgende Gleichungssystem:

$$
\left(\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 4 & 3 & 0 & 2 \\ 0 & 3 & 4 & 4 \end{array}\right) \mapsto \left(\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 0 & 7 & -4 & 2 \\ 0 & 3 & 4 & 4 \end{array}\right) \mapsto \left(\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 0 & 7 & -4 & 2 \\ 0 & 0 & 40 & 22 \end{array}\right)
$$

$$
\mapsto \left(\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 0 & 1 & -\frac{4}{7} & \frac{2}{7} \\ 0 & 0 & 1 & \frac{11}{20} \end{array}\right) \mapsto \left(\begin{array}{rrr|r} 1 & 0 & 0 & \frac{1}{20} \\ 0 & 1 & 0 & \frac{3}{5} \\ 0 & 0 & 1 & \frac{11}{20} \end{array}\right)
$$

woraus sich die folgende Lösung ergibt:

$$I_3 = \tfrac{11}{20}, I_2 = \tfrac{3}{5} \text{ und } I_1 = \tfrac{1}{20}.$$

Beispiel 7.30 (Kubische Splines)
Im "Computer aided geometric design" (CAGD) werden zum Design von Flächen und Kurven (z. B. im Automobil- oder Flugzeugbau) Flächen- und Kurvenstücke verwendet (meist durch sogenannte kubische Splines realisiert), die dann an den Endpunkten oder Randkurven glatt zusammenpassen müssen. Am bekanntesten sind die Bézier-Kubiken, die von dem französischen Auto-Designer bei Renault, P. Bézier, eingeführt wurden (diese werden heute z. B. auch in der Text-Beschreibungssprache PostScript verwendet).

Ein typisches Problem ist z.B. die Bestimmung einer kubischen Parabel

$$f(x) = ax^3 + bx^2 + cx + d$$

durch zwei Punkte (x_1, y_1), (x_2, y_2) in der Ebene mit vorgegebener Steigung m_1 in (x_1, y_1) und m_2 in (x_2, y_2).

Für $(x_1, y_1) = (0, 2)$, $(x_2, y_2) = (4, 0)$, $m_1 = -3$, $m_2 = -3$ ergibt sich aus

$$f'(x) = 3ax^2 + 2bx + c$$

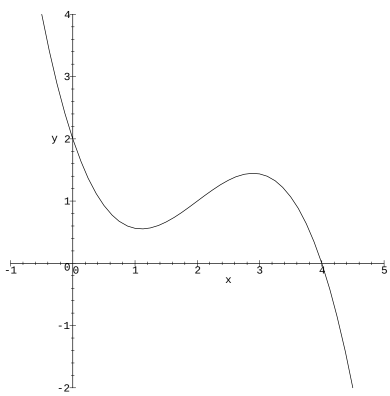
und

$$f(0) = 2, f(4) = 0, f'(0) = -3 \text{ und } f'(4) = -3$$

das lineare Gleichungssystem

$$\begin{aligned} d &= 2, \\ 64a + 16b + 4c + d &= 0, \\ c &= -3, \\ 48a + 8b + c &= -3, \end{aligned}$$

also $d = 2$, $c = -3$, $6a + b = 0$, $32a + 8b = 5$, und damit $a = -\frac{5}{16}$ und $b = \frac{15}{8}$. Die Kurve $y = -\frac{5}{16}x^3 + \frac{15}{8}x^2 - 3x + 2$ hat etwa die folgende Gestalt



Die Aufgabe ist, wie leicht zu sehen ist, stets lösbar und daher können kubische Splines stückweise definiert und glatt aneinander gesetzt werden.

Beispiel 7.31 (Leontieff-Modell)
Die folgende Planungsaufgabe zeigt, daß durchaus Gleichungen mit vielen Veränderlichen in der Praxis auftauchen.

Ein Konzern besitzt n Fabriken $F_1, \ldots, F_n$, in der Fabrik F_i wird das Produkt P_i hergestellt.

Zur Produktion einer Einheit von P_k werden a_{jk} Einheiten von P_j benötigt; wir nehmen an $a_{ii} = 0$.

Am Ende eines Produktionszyklus sind x_k Einheiten von P_k hergestellt, $k = 1, \ldots, n$; wir haben also einen Produktionsvektor $x = (x_1, \ldots, x_n)$.

Zur Herstellung von $x = (x_1, \ldots, x_n)$ werden

$$\sum_{k=1}^{n} a_{jk} x_k = a_{j1} x_1 + \cdots + a_{jn} x_n$$

Einheiten von P_j verbraucht.

Für den Markt verbleiben damit

$$y_j = x_j - \sum_{k=1}^{n} a_{jk} x_k$$

Einheiten von P_j.

Die Planungsaufgabe lautet nun:

Der Mehrbedarf $y = (y_1, \ldots, y_n)$ ist vorgegeben. Gesucht ist ein Produktionsvektor $x = (x_1, \ldots, x_n)$ mit

$$\begin{array}{rcl} x_1 - (a_{11} x_1 + \cdots + a_{1n} x_n) & = & y_1 \\ \vdots & & \vdots \\ x_n - (a_{n1} x_1 + \cdots + a_{nn} x_n) & = & y_n. \end{array}$$

Also ist ein lineares Gleichungssystem zu lösen. Allerdings, und das macht das Problem schwerer, ist zu beachten, daß alle $x_i \geq 0$ sein müssen (natürlich sind auch die y_j und die $a_{j_k} \geq 0$).

(Das Modell heißt Leontieff-Modell und ist nach Vassili Leontieff benannt, der 1973 den Nobelpreis für Wirtschaftswissenschaften erhielt.)

Ein einfaches Beispiel mit zwei Fabriken, Verbrauchsmatrix

$$\begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{3} & 0 \end{pmatrix}$$

und zunächst unbestimmtem Mehrbedarf (y_1, y_2) liefert das Gleichungssystem

$$\left(\begin{array}{rr|r} 1 & -\frac{1}{2} & y_1 \\ -\frac{1}{3} & 1 & y_2 \end{array}\right) \mapsto \left(\begin{array}{rr|c} 1 & -\frac{1}{2} & y_1 \\ 0 & \frac{5}{6} & \frac{1}{3}y_1 + y_2 \end{array}\right) \mapsto \left(\begin{array}{rr|c} 1 & -\frac{1}{2} & y_1 \\ 0 & 1 & \frac{2}{5}y_1 + \frac{6}{5}y_2 \end{array}\right)$$
$$\mapsto \left(\begin{array}{rr|c} 1 & 0 & \frac{6}{5}y_1 + \frac{3}{5}y_2 \\ 0 & 1 & \frac{2}{5}y_1 + \frac{6}{5}y_2 \end{array}\right)$$

und damit $x_1 = \frac{6}{5}y_1 + \frac{3}{5}y_2$, $x_2 = \frac{2}{5}y_1 + \frac{6}{5}y_2$.

Beispiel 7.32 (Finde eine Gleichungssystem zu gegebener Lösung.)
Ein Gleichungssystem besitze die spezielle Lösung $(1, 0, 1)^t$ und das zugehörige homogene System besitze $(1, 1, 1)^t$ als Lösung und habe den Rang zwei. Finde ein Gleichungssystem, das diese Bedingungen erfüllt.

Da die Lösungen Vektoren im $\mathbb{R}^3$ sind, ist es ein System in drei Variablen.

Da der Rang zwei ist, hat die Zeilen-Stufen-Form zwei Zeilen. Da die Lösungsmenge nicht von der Form abhängt, können wir das System in Zeilen-Stufen-Form annehmen:

Problem: Finde eine Gerade im $\mathbb{R}^3$, die selbst durch $(1, 0, 1)^t$ geht und für die die in den Nullpunkt verschobene Gerade durch $(1, 1, 1)^t$ geht.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1, \\ a_{22}x_2 + a_{23}x_3 &= b_2. \end{aligned}$$

$(1, 0, 1)^t$ ist Lösung:

$$\begin{aligned} a_{11} + a_{13} &= b_1, && (1) \\ a_{23} &= b_2. && (2) \end{aligned}$$

$(1, 1, 1)^t$ ist Lösung des homogenen Systems:

$$\begin{aligned} a_{11} + a_{12} + a_{13} &= 0, && (3) \\ a_{22} + a_{23} &= 0. && (4) \end{aligned}$$

Das zugehörige lineare Gleichungssystem in $a_{11}, a_{12}, , a_{13}, a_{22}, a_{23}$, b_1, b_2 lautet:

$$\begin{array}{c} \\ (1) \\ (3) \\ (4) \\ (2) \end{array} \begin{array}{c} \begin{array}{ccccccc} a_{11} & a_{12} & a_{13} & a_{22} & a_{23} & b_1 & b_2 \end{array} \\ \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & -1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{pmatrix} \end{array} \mapsto \begin{array}{c} \begin{array}{ccccccc} a_{11} & a_{12} & a_{13} & a_{22} & a_{23} & b_1 & b_2 \end{array} \\ \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{pmatrix} \end{array}.$$

Das System hat unendlich viele Lösungen, und da der Rang 2 sein soll, muß $a_{22} \neq 0$ und damit auch $a_{23} = -a_{22} \neq 0$ sein.

Wir wählen

$$\begin{aligned} a_{22} &= 1 \Rightarrow a_{23} = b_2 = -1, \\ a_{12} &= 1 \Rightarrow b_1 = -1, \\ a_{11} &= 1 \Rightarrow a_{13} = -2. \end{aligned}$$

Also ist

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

ein geeignetes Gleichungssystem.

I) Ringe und Moduln

Bemerkung 7.33 (Ringe und Moduln)
Man kann lineare Gleichungssysteme in analoger Weise über kommutativen Ringen mit Eins betrachten. Es bleibt richtig, daß die Lösungsmenge eines homogenen LGS ein Modul $\text{Lös}(A, 0) = \text{Ker}(f_A)$, und daß $Ax = b$ genau dann lösbar ist, wenn $b \in \text{Im}(f_A)$. Auch die Strukturaussage $\text{Lös}(A, b) = c + \text{Lös}(A, 0)$ in Satz 7.5 bleibt wahr. Alle Aussagen, die den Rang einer Matrix verwenden, sind jedoch nicht mehr richtig. Außerdem kann sich die Lösungmenge eines Gleichungssystems ändern, wenn man Zeilen der Matrix mit einer Konstanten multipliziert, die kein Inverses im Ring besitzt. Es ist also Vorsicht geboten, wenn man den abgewandelten Gauß-Algorithmus, der ohne Division auskommt, verwenden will, um die erweiterte Matrix auf ZSF zu bringen. Ist der Ring ein sogenannter Integritätsbereich, wie etwa die ganzen Zahlen $\mathbb{Z}$, so entstehen dadurch keine wirklichen Probleme, da aus $\lambda \cdot x = 0$ mit $\lambda \neq 0$ immer noch $x = 0$ folgt.

Aufgaben

Aufgabe 7.34
Berechne die Lösungsmenge des folgenden Gleichungssystems, sofern es lösbar ist:

$$\begin{array}{rcrcrcr} -x & + & 6y & + & 2z & = & 4 \\ 2x & - & 2y & - & z & = & 2 \\ 3x & - & 4y & - & 2z & = & 1 \end{array}$$

Aufgabe 7.35
Für welche $a, b \in \mathbb{R}$ besitzt das lineare Gleichungssystem

$$\begin{array}{rcrcrcr} ax & + & & & z & = & ab \\ -2x & + & by & + & az & = & -b \\ & & by & + & (a+1)z & = & b \end{array}$$

außer $(b, 1, 0)$ noch weitere Lösungen. Bestimme sie.

Aufgabe 7.36

Bestimme die Lösungsmenge des folgenden linearen Gleichungssystems über $\mathbb{R}$ in Abhängigkeit vom Parameter $t \in \mathbb{R}$:

$$\begin{array}{ccccccc} x & + & y & + & z & = & 1 \\ & & ty & + & z & = & 1 \\ tx & + & ty & + & z & = & 1+t \end{array}$$

Aufgabe 7.37

Bestimme eine Basis des Kerns und des Bildes von f_A mit

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 1 & 1 \\ -1 & -1 & 0 & 1 & 1 \\ -1 & -1 & -1 & 0 & 1 \\ -1 & -1 & -1 & -1 & 0 \end{pmatrix} \in \mathrm{Mat}_5(\mathbb{R}).$$

Aufgabe 7.38

Es sei $U = \{(x+y, y, y-x)^t \mid x, y \in \mathbb{R}\}$ und $U' = \{(x, y, z)^t \in \mathbb{R}^3 \mid z = 2x + y\}$. Bestimme Basen von $U + U'$, $U \cap U'$, $\mathbb{R}^3/U$ und $\mathbb{R}^3/U'$.

Aufgabe 7.39

Bestimme eine Basis für $U \cap U'$ mit

$$U = \langle (2, -1, 1, -1)^t, (1, -2, 2, 1)^t, (3, -1, 0, 2)^t \rangle \leq \mathbb{R}^4$$

und

$$U' = \langle (3, -2, 3, 8)^t, (2, 1, -5, 3)^t \rangle \leq \mathbb{R}^4.$$

Aufgabe 7.40

Es sei $U = \langle (1, 2, 3, 4)^t, (1, 1, 1, 1)^t \rangle \leq \mathbb{R}^4$. Bestimme mit Hilfe des Austauschsatzes von Steinitz eine Basis von $\mathbb{R}^4/U$.

Aufgabe 7.41

Wir betrachten

$$B = \big((1, 1, 1, 1)^t, (-1, 0, 0, 1)^t, (0, -1, 0, 1)^t, (0, 0, -1, 1)^t \big)$$

und

$$D = \big((1, 1, 0)^t, (0, 1, 1)^t, (0, 0, 1)^t \big).$$

a. Zeige, dass B eine Basis des $\mathbb{R}^4$ und D eine Basis des $\mathbb{R}^3$ ist.

b. Bestimme $M_D^B(f)$ für $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^3 : (x_1, x_2, x_3, x_4)^t \mapsto (x_1 - x_2, x_3, x_2 + x_4)^t$.

c. Bestimme umgekehrt die Funktionsvorschrift für $g \in \mathrm{Hom}_{\mathbb{R}}\left(\mathbb{R}^4, \mathbb{R}^3\right)$ mit

$$M_D^B(g) = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 2 & 1 & 3 & 1 \\ 0 & -1 & 2 & 0 \end{pmatrix}.$$

Aufgabe 7.42
Es sei V ein $\mathbb{R}$-Vektorraum, $B = (x_1, x_2, x_3)$ eine Basis von V und $B' = (y_1, y_2, y_3)$ mit $y_1 = x_1 + x_3$, $y_2 = x_1 + x_2$ und $y_3 = x_1 + x_2 + x_3$.

a. Zeige, dass B' eine Basis von V ist.

b. Bestimme $M_{B'}^{B'}(f)$, wobei $f \in \mathrm{End}_{\mathbb{R}}(V)$ gegeben ist durch

$$M_B^B(f) = \begin{pmatrix} a & 0 & b \\ -b & a & a \\ a & b & b \end{pmatrix} \text{ mit } a, b \in \mathbb{R}.$$

Aufgabe 7.43
Seien $B = \big((1,1,1)^t, (1,1,0)^t, (1,0,-1)^t\big)$ und $B' = \big((2,1)^t, (1,1)^t\big)$. E bzw. E' seien die kanonischen Basen des $\mathbb{R}^3$ bzw. des $\mathbb{R}^2$. Ferner sei $f \in \mathrm{Hom}_K\big(\mathbb{R}^3, \mathbb{R}^2\big)$ gegeben durch $f\big((x,y,z)^t\big) = (x - y + z, 2x + y)^t$.

a. Zeige, dass B und B' Basen des $\mathbb{R}^3$ bzw. des $\mathbb{R}^2$ sind.

b. Bestimme $M_{E'}^E(f)$.

c. Bestimme $M_{B'}^B(f)$ sowie die Transformationsmatrizen T_E^B und $T_{B'}^{E'}$ mit $T_{B'}^{E'} \cdot M_{E'}^E(f) \cdot T_E^B = M_{B'}^B(f)$.

§ 8 Die Determinante

Wir werden jetzt eine ganz neue Möglichkeit kennenlernen, um quadratische lineare Gleichungssysteme zu lösen, nämlich mit Hilfe von Determinanten. Die Determinante ordnet einer quadratischen Matrix über einem Körper ein Element des Körpers zu, das genau dann ungleich Null ist, wenn die Matrix invertierbar ist. Die Determinante liefert aber nicht nur ein nützliches Kriterium für die Invertierbarkeit, sie ist vor allem aus theoretischen Gründen von unschätzbarem Wert. Z. B. liefert die Cramersche Regel mit Hilfe der Determinante eine geschlossene Formel für die Lösung eines linearen Gleichungssystems. Aus dieser Formel lassen sich Eigenschaften der Lösungen als Funktionen der Koeffizienten der Matrix bestimmen.

Die Determinante einer Matrix ist eine *polynomiale Funktion* in den Einträgen der Matrix. Sind diese Einträge etwa reelle oder komplexe Zahlen, so hängt die Determinante stetig von den Einträgen ab. Daraus folgt z. B. die wichtige Tatsache, daß eine invertierbare Matrix bei kleiner Störung der Einträge invertierbar bleibt. Damit wird eine Verbindung zur Analysis hergestellt. Eine weitere wichtige Bedeutung in der Analysis hat die Determinante für die Volumenberechnung (siehe auch Bemerkung 12.33.

Wir werden die Eigenschaften der Determinante soweit entwickeln, wie sie in der linearen Algebra wichtig sind. Allerdings führt uns die Determinante auch hier schon auf eine höhere Stufe: die Determinante ist nicht nur linear, sie ist *multilinear*, wie wir gleich sehen werden.

A) Die Leibnitz-Formel für die Determinante

Definition 8.1 (Determinante)
Wir definieren für $A \in \mathrm{Mat}_n(K)$ die *Determinante* von A durch die *Leibniz-Formel*

$$\det(A) := |A| := \sum_{\sigma \in \mathbb{S}_n} \mathrm{sgn}(\sigma) \cdot a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}. \tag{20}$$

Beispiel 8.2 (Determinanten für $n = 1, 2, 3$)

a. Ist $n = 1$, dann ist $A = (a) \in \mathrm{Mat}(1, K)$ und $\det(A) = a$.

b. Ist $n = 2$, dann ist $\mathbb{S}_2 = \{\mathrm{id}, (1\ 2)\}$ und damit folgt:

$$\det(A) = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

d. h. $\det(A)$ ist das Produkt der Elemente der Hauptdiagonalen minus dem Produkt der Elemente der Gegendiagonalen. Z.B.

$$\det \begin{pmatrix} 5 & 2 \\ 6 & 3 \end{pmatrix} = 5 \cdot 3 - 6 \cdot 2 = 3.$$

c. Für $n = 3$ hat $\mathbb{S}_n$ bereits sechs Elemente. Man berechnet in diesem Fall die Determinante mit der *Regel von Sarrus*:

$$\begin{array}{ccccc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}$$

Die Produkte der Elemente längs der gestrichelten Linien tauchen bei der Berechnung der Determinante als positive Summanden auf, die Produkte der Elemente längs der gepunkteten Linien als negative Summanden. D. h., wir erhalten:

$$\begin{aligned} \det(A) &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad -a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}. \end{aligned}$$

Wenden wir das obige Schema auf die Matrix

$$A = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 5 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

an, so erhalten wir

$$\begin{array}{ccccc} 2 & 0 & 1 & 2 & 0 \\ 0 & 5 & 0 & 0 & 5 \\ -1 & 0 & 1 & -1 & 0 \end{array}$$

und damit

$$\det(A) = 2 \cdot 5 \cdot 1 - (-1) \cdot 5 \cdot 1 = 15.$$

d. Für $n = 4$ ergeben sich schon $4! = 24$ Summanden und für $n = 10$ gar $10! = 3628800$. In numerischen Anwendungen sind 1000×1000-Matrizen keine Seltenheit, so daß es sich von selbst versteht, daß dabei nicht die Definition, bei der dann für die Determinante über 10^{2567} Produkte berechnet werden müßten, zur Berechnung verwendet werden kann. In der Tat wird zur Berechnung von Determinanten über Körpern wieder der Gauß-Algorithmus eine wichtige Rolle spielen.

Proposition 8.3 (Determinanten von Dreiecksmatrizen)
Ist $A = (a_{ij}) \in \mathrm{Mat}_n(K)$ *eine obere (bzw. untere)* Dreiecksmatrix, *d. h.* $a_{ij} = 0$ *für* $i > j$ *(bzw.* $i < j$*), dann ist*

$$\det(A) = a_{11} \cdots a_{nn}$$

das Produkt der Diagonalelemente.

Beweis: Ist $\mathrm{id} \neq \sigma \in \mathbb{S}_n$, so gilt $i > \sigma(i)$ (bzw. $i < \sigma(i)$) für mindestens ein i. Wegen der Voraussetzung $a_{i\sigma(i)} = 0$ für $i > \sigma(i)$ (bzw. $i < \sigma(i)$) bleibt von den Summanden in (20) also nur der für id übrig. □

Beispiel 8.4

$$\det\begin{pmatrix} 2 & 3 & 12 & -3 \\ 0 & -1 & 5 & -111 \\ 0 & 0 & 3 & 5 \\ 0 & 0 & 0 & 1 \end{pmatrix} = 2\cdot(-1)\cdot 3\cdot 1 = -6.$$

Lemma 8.5 (Alternative Leibnitz-Formel)
Für die Determinante von $A \in \mathrm{Mat}_n(K)$ *gilt*

$$\det(A) = \sum_{\sigma\in\mathbb{S}_n} \mathrm{sgn}(\sigma)\cdot a_{\sigma(1)1}\cdot\ldots\cdot a_{\sigma(n)n}. \tag{21}$$

Beweis: Man beachte, daß für $\sigma \in \mathbb{S}_n$ auch σ^{-1} eine Permutation der Zahlen $1,\ldots,n$ ist, d. h. $\{1,\ldots,n\} = \left\{\sigma^{-1}(1),\ldots,\sigma^{-1}(n)\right\}$. Zudem wissen wir aus Satz B2.7 e., daß $\mathrm{sgn}(\sigma) = \mathrm{sgn}\left(\sigma^{-1}\right)$, und es ist gleich, ob wir über $\sigma\in\mathbb{S}_n$ summieren oder über $\sigma^{-1}\in\mathbb{S}_n$, da auf beide Weisen alle Elemente von $\mathbb{S}_n$ je einmal erreicht werden. Aus diesen Vorbetrachtungen ergibt sich:

$$\begin{aligned}
\det(A) &= \sum\nolimits_{\sigma\in\mathbb{S}_n} \mathrm{sgn}(\sigma)\cdot a_{\sigma^{-1}(1)\sigma(\sigma^{-1}(1))}\cdot\ldots\cdot a_{\sigma^{-1}(n)\sigma(\sigma^{-1}(n))} \\
&= \sum\nolimits_{\sigma\in\mathbb{S}_n} \mathrm{sgn}(\sigma)\cdot a_{\sigma^{-1}(1)1}\cdot\ldots\cdot a_{\sigma^{-1}(n)n} \\
&\overset{\text{B2.7e.}}{=} \sum\nolimits_{\sigma\in\mathbb{S}_n} \mathrm{sgn}\left(\sigma^{-1}\right)\cdot a_{\sigma^{-1}(1)1}\cdot\ldots\cdot a_{\sigma^{-1}(n)n} \\
&= \sum\nolimits_{\sigma^{-1}\in\mathbb{S}_n} \mathrm{sgn}\left(\sigma^{-1}\right)\cdot a_{\sigma^{-1}(1)1}\cdot\ldots\cdot a_{\sigma^{-1}(n)n} \\
&= \sum\nolimits_{\pi\in\mathbb{S}_n} \mathrm{sgn}(\pi)\cdot a_{\pi(1)1}\cdot\ldots\cdot a_{\pi(n)n}.
\end{aligned}$$

□

Proposition 8.6 (Die Determinante der Transponierten)
Für $A\in\mathrm{Mat}_n(K)$ *gilt:*

$$\det(A) = \det\left(A^t\right).$$

Beweis: Sei $A = (a_{ij})$ und $A^t = (a'_{ij})$, dann gilt $a'_{ij} = a_{ji}$. Mithin erhalten wir mit Hilfe von Lemma 8.5

$$\begin{aligned}
\det(A) &= \sum_{\sigma\in\mathbb{S}_n} \mathrm{sgn}(\sigma)\cdot a_{1\sigma(1)}\cdot\ldots\cdot a_{n\sigma(n)} \\
&= \sum_{\sigma\in\mathbb{S}_n} \mathrm{sgn}(\sigma)\cdot a'_{\sigma(1)1}\cdot\ldots\cdot a'_{\sigma(n)n} \quad = \quad \det\left(A^t\right).
\end{aligned}$$

□

Beispiel 8.7
Beispiel 8.2 b. aufgreifend gilt

$$\det\begin{pmatrix} 5 & 6 \\ 2 & 3 \end{pmatrix} = \det\begin{pmatrix} 5 & 2 \\ 6 & 3 \end{pmatrix} = 3.$$

B) Die Determinante als Volumenform

Definition 8.8 (Multilineare Abbildungen)
Es seien V und W zwei K-Vektorräume.

a. Eine Abbildung
$$f : V^n = V \times \overset{n}{\cdots} \times V \to W$$
heißt *multilinear*, falls f in jedem Argument linear ist, d. h. es gelten
$$f(x_1, \ldots, x_i + y_i, \ldots, x_n) = f(x_1, \ldots, x_i, \ldots, x_n) + f(x_1, \ldots, y_i, \ldots, x_n)$$
und
$$f(x_1, \ldots, \lambda x_i, \ldots, x_n) = \lambda \cdot f(x_1, \ldots, x_i, \ldots, x_n)$$
für jedes $i \in \{1, \ldots, n\}$ und für alle $x_1, \ldots, x_n, y_i \in V$ und $\lambda \in K$.

b. Eine multilineare Abbildung $f : V^n \to W$ heißt *alternierend*, falls für $(x_1, \ldots, x_n) \in V^n$ mit $x_i = x_j$ für ein $i \neq j$, gilt:
$$f(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) = 0.$$

Lemma 8.9
Ist $f : V^n \to W$ eine alternierende multilineare Abbildung, dann gilt für $\sigma \in \mathbb{S}_n$
$$f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = \operatorname{sgn}(\sigma) \cdot f(x_1, \ldots, x_n).$$
Insbesondere gilt $f(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) = -f(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n)$.

Beweis: Wir betrachten zunächst den Fall, daß $\sigma = (i\ j)$ eine Transposition ist. Da f alternierend und multilinear ist, folgt die Behauptung für σ aus
$$\begin{aligned}
0 &= f(x_1, \ldots, x_i + x_j, \ldots, x_i + x_j, \ldots, x_n) \\
&= f(x_1, \ldots, x_i, \ldots, x_i, \ldots, x_n) + f(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) \\
&\quad + f(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n) + f(x_1, \ldots, x_j, \ldots, x_j, \ldots, x_n) \\
&= f(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) + f(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n).
\end{aligned}$$
Ist $\sigma \in \mathbb{S}_n$ beliebig, so können wir $\sigma = \tau_1 \circ \ldots \circ \tau_k$ als Produkt von Transpositionen schreiben und die Behauptung folgt mittels Induktion nach der Anzahl k der Transpositionen. Den Induktionsanfang $k = 1$ haben wir bereits gezeigt. Ist $k \geq 2$ und setzen wir $\pi = \tau_2 \circ \ldots \tau_k$, so folgt mit der Vorüberlegung
$$\begin{aligned}
f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) &= f(x_{\tau_1(\pi(1))}, \ldots, x_{\tau_1(\pi(n))}) = -f(x_{\pi(1)}, \ldots, x_{\pi(n)}) \\
&\overset{\text{Ind.}}{=} -\operatorname{sgn}(\pi) \cdot f(x_1, \ldots, x_n) \overset{\text{Satz B2.7}}{=} \operatorname{sgn}(\tau_1 \circ \pi) \cdot f(x_1, \ldots, x_n) \\
&= \operatorname{sgn}(\sigma) \cdot f(x_1, \ldots, x_n).
\end{aligned}$$
□

Bemerkung 8.10
Wir können $\mathrm{Mat}_n(K)$ auf recht natürliche Weise mit $K^n \times \overset{n}{\cdots} \times K^n$ identifizieren, indem wir eine Matrix $A = (a_{ij})$ mit dem n-Tupel ihrer Spaltenvektoren $(a^1, \ldots, a^n)$ gleichsetzen. Das wollen wir im folgenden tun.

Satz 8.11 (Die Determinante als Volumenform)

a. *Die Determinante*
$$\det : \mathrm{Mat}_n(K) \to K : A \mapsto \det(A)$$
ist eine alternierende multilineare Abbildung mit $\det(\mathbb{1}_n) = 1$.

b. *Ist* $f : \mathrm{Mat}_n(K) \longrightarrow K$ *eine alternierende multilineare Abbildung und* $A \in \mathrm{Mat}_n(K)$*, so gilt*
$$f(A) = f(\mathbb{1}_n) \cdot \det(A).$$

Beweis:

a. Wir werden im Beweis die Formel (21) aus Lemma 8.5 zur Berechnung der Determinante verwenden, da sie auf die Bedürfnisse der Determinante als multilineare Abbildung bezüglich der Spalten zugeschnitten ist.
Es seien $a^j = (a_{1j}, \ldots, a_{nj})^t$, $j = 1, \ldots, n$, und $b^i = (b_{1i}, \ldots, b_{ni})^t$. Wir setzen $A := (a^1 \ldots a^i \ldots a^n)$, $B := (a^1 \ldots b^i \ldots a^n)$ und $C := (a^1 \ldots \lambda a^i + \mu b^i \ldots a^n)$. Dann gilt
$$\begin{aligned} \det(C) &= \sum_{\sigma \in \mathbb{S}_n} \mathrm{sgn}(\sigma) \cdot a_{\sigma(1)1} \cdot \ldots \cdot (\lambda a_{\sigma(i)i} + \mu b_{\sigma(i)i}) \cdot \ldots \cdot a_{\sigma(n)n} \\ &= \lambda \cdot \sum_{\sigma \in \mathbb{S}_n} \mathrm{sgn}(\sigma) \cdot a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(i)i} \cdot \ldots \cdot a_{\sigma(n)n} \\ &\quad + \mu \cdot \sum_{\sigma \in \mathbb{S}_n} \mathrm{sgn}(\sigma) \cdot a_{\sigma(1)1} \cdot \ldots \cdot b_{\sigma(i)i} \cdot \ldots \cdot a_{\sigma(n)n} \\ &= \lambda \cdot \det(A) + \mu \cdot \det(B), \end{aligned}$$
so daß det multilinear ist.
Sei nun $a^i = a^j$, für ein $i \neq j$. Ist $\tau = (i\ j)$, die Transposition, die i und j vertauscht, dann besitzt $\mathbb{S}_n$ nach Satz B2.7 die Zerlegung $\mathbb{S}_n = \mathbb{A}_n \,\dot\cup\, \mathbb{A}_n\tau$. Ferner gilt für $\sigma \in \mathbb{A}_n$
$$\mathrm{sgn}(\sigma) = 1 \quad \text{und} \quad \mathrm{sgn}(\sigma\tau) = -1.$$
Wir erhalten also
$$\begin{aligned} \det(A) &= \sum_{\sigma \in \mathbb{A}_n} a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(i)i} \cdot \ldots \cdot a_{\sigma(j)j} \cdot \ldots \cdot a_{\sigma(n)n} \\ &\quad - \sum_{\sigma \in \mathbb{A}_n} a_{\sigma\tau(1)1} \cdot \ldots \cdot a_{\sigma\tau(i)i} \cdot \ldots \cdot a_{\sigma\tau(j)j} \cdot \ldots \cdot a_{\sigma\tau(n)n} \\ &= \sum_{\sigma \in \mathbb{A}_n} a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(i)i} \cdot \ldots \cdot a_{\sigma(j)j} \cdot \ldots \cdot a_{\sigma(n)n} \\ &\quad - \sum_{\sigma \in \mathbb{A}_n} a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(j)i} \cdot \ldots \cdot a_{\sigma(i)j} \cdot \ldots \cdot a_{\sigma(n)n} = 0, \end{aligned}$$
und somit ist det alternierend.
Außerdem folgt $\det(\mathbb{1}_n) = 1 \cdot \ldots \cdot 1 = 1$ aus Proposition 8.3.

b. Mit den Notationen von a. gilt $a^i = \sum_{j=1}^n a_{ji} e^j$, wenn e^j der j-te Einheitsvektor ist. Aus der Multilinearität von f folgt:
$$f(A) = \sum_{j_1=1}^n a_{j_1 1} f\big(e^{j_1}\ a^2 \ldots a^n\big) = \sum_{j_1=1}^n a_{j_1 1} \sum_{j_2=1}^n a_{j_2 2} f\big(e^{j_1}\ e^{j_2}\ a^3 \ldots a^n\big)$$

$$= \ldots = \sum_{j_1,\ldots,j_n=1}^{n} a_{j_1 1} \cdot \ldots \cdot a_{j_n n} f\big(e^{j_1} \;\ldots\; e^{j_n}\big).$$

Genau dann, wenn die $j_1, \ldots, j_n$ paarweise verschieden sind, existiert eine Permutation $\sigma \in \mathbb{S}_n$ mit $\big(e^{j_1} \ldots e^{j_n}\big) = \big(e^{\sigma(1)} \ldots e^{\sigma(n)}\big)$, und wegen Lemma 8.9 gilt dann

$$f(e^{\sigma(1)} \ldots e^{\sigma(n)}) = \operatorname{sgn}(\sigma) \cdot f(e^1 \ldots e^n) = \operatorname{sgn}(\sigma) \cdot f(\mathbb{1}_n).$$

Andernfalls stimmen zwei der j_i überein und $f\big(e^{j_1} \;\ldots\; e^{j_n}\big) = 0$, da f alternierend ist. Insgesamt haben wir damit gezeigt:

$$f(A) = \sum_{\sigma \in \mathbb{S}_n}^{n} a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n} \cdot \operatorname{sgn}(\sigma) \cdot f(\mathbb{1}_n) = \det(A) \cdot f(\mathbb{1}_n).$$

□

Bemerkung 8.12 (Das Volumen des Parallelotops)
Eine alternierende multilineare Abbildung $f : \operatorname{Mat}_n(K) \longrightarrow K$ wird auch eine *Volumenform* genannt. Aus Satz 8.11 b. folgt, daß die Determinante die einzige Volumenform f mit $f(\mathbb{1}_n) = 1$ ist, d.h. det ist durch die Eigenschaften in Satz 8.11 a. eindeutig bestimmt.

Die Determinante hat eine wichtige geometrische Interpretation, die den Begriff *Volumenform* rechtfertigt. Seien $x_1, \ldots, x_n \in \mathbb{R}^n$ und sei

$$P\big(x_1, \ldots, x_n\big) := \big\{\lambda_1 x_1 + \ldots + \lambda_n x_n \in \mathbb{R}^n \mid 0 \le \lambda_i \le 1, i = 1, \ldots, n\big\}$$

das von den Vektoren $x_1, \ldots, x_n$ aufgespannte *Parallelotop* (siehe Abbildung 3).

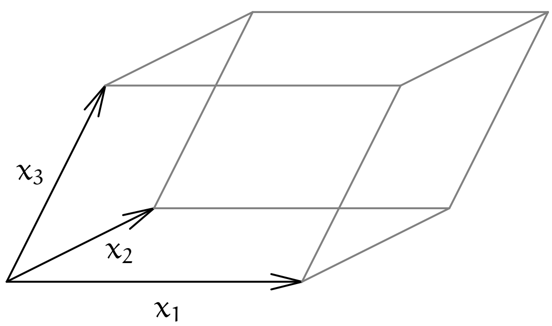


Abbildung 3. Das Parallelotop $P(x_1, x_2, x_3)$ im $\mathbb{R}^3$

Dann definiert man das n-dimensionale Volumen von $P(x_1, \ldots, x_n)$ mit Hilfe der Determinante als

$$\text{Volumen}\,\big(P(x_1, \ldots, x_n)\big) = |\det\big(x_1 \ldots x_n\big)|.$$

In Dimension $n = 1$ ist $|\det(x_1)| = |x_1|$ in der Tat die Länge der Strecke von 0 nach x_1, und diese ist gerade $P(x_1)$. Wir werden in Bemerkung 12.33 zeigen, daß auch in Dimension $n = 2$ und $n = 3$ das so definierte Volumen mit dem euklidischen Flächeninhalt bzw. mit dem euklidischen Volumen übereinstimmt, daß die Definition also sinnvoll ist. Sie wird im Rahmen der mehrdimensionalen Integrationstheorie und der Verallgemeinerung der Substitutionsregel eine wichtige Rolle spielen. □

C) Der Gauß-Algorithmus zur Berechnung der Determinante

Korollar 8.13 (Spaltenoperationen und die Determinante)
Es sei $A \in \mathrm{Mat}_n(K)$ *und* $\lambda \in K$.

a. *Bei Vertauschung zweier Spalten von* A *ändert sich das Vorzeichen von* $\det(A)$.
b. *Bei Multiplikation einer Spalte von* A *mit* λ *multipliziert sich* $\det(A)$ *mit* λ.
c. *Bei Addition des* λ*-fachen einer Spalte zu einer anderen Spalte ändert sich* $\det(A)$ *nicht.*
d. *Enthält* A *eine Nullspalte, so ist* $\det(A) = 0$.
e. *Sind zwei Spalten von* A *gleich, so ist* $\det(A) = 0$.

Beweis:

a. Das ist ein Spezialfall von Lemma 8.9, da det nach Satz 8.11 alternierend ist.
b. Dies folgt aus der Multilinearität von det, siehe Satz 8.11.
c. Für $A = \left(a^1 \ldots a^n\right)$ und $A' = \left(a^1 \ldots a^j + \lambda a^i \ldots a^n\right)$ folgt aus der Multilinearität und da det alternierend ist:
$$\det(A') = \det(A) + \lambda \cdot \det(a^1 \ldots a^i \ldots a^i \ldots a^n) = \det(A) + \lambda \cdot 0 = \det(A).$$
d. Ist eine Spalte von A Null, so folgt $\det(A) = 0$ aus b. mit $\lambda = 0$.
e. Das folgt, da det alternierend ist.

□

Da die Determinante einer Matrix gleich der Determinante der Transponierten ist, sind die Begriffe Spalte und Zeile austauschbar. Eine exaktere Formulierung bietet das folgende Korollar.

Korollar 8.14 (Zeilenoperationen und die Determinante)
Wir können $\det : \mathrm{Mat}_n(K) \to K$ *auch als multilineare Abbildung auf den Zeilen einer Matrix* A *auffassen. Entsprechend gilt Korollar 8.13 auch für Zeilen statt Spalten.*

Da sich die Determinante bei der Addition eines Vielfachen einer Zeile zu einer anderen nicht ändert, können wir den Gauß-Algorithmus zur Berechnung von Determinanten einsetzen.

Algorithmus 8.15 (Algorithmus zur Berechnung der Determinante über K)
INPUT: $A \in \mathrm{Mat}_n(K)$.

OUTPUT: $\det(A)$.

1. Schritt: Setze $d = 1$.

2. Schritt: Überführe A mittels Gauß-Algorithmus in nicht-reduzierte ZSF, d. h. führe im Gauß-Algorithmus 6.10 Schritt sieben nicht aus. Jedesmal, wenn dabei zwei Zeilen vertauscht werden, ersetze d durch $-d$. - Wird bei der Gaußreduktion ein Pivotelement zu Null, gib Null zurück und brich ab.

3. Schritt: Gib das Produkt von d mit den Diagonalelementen der ZSF zurück.

Beispiel 8.16

Wir wollen die Determinante der Matrix

$$A = \begin{pmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 0 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R})$$

berechnen. Dazu überführen wir sie mittels des Gauß-Algorithmus in ZSF und merken uns die Zeilenvertauschungen.

$$\begin{pmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 0 \end{pmatrix} \xmapsto[d=-1]{\mathrm{I}\leftrightarrow\mathrm{II}} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 0 \end{pmatrix} \xmapsto[\mathrm{III}\to\mathrm{III}-7\mathrm{I}]{\mathrm{II}\to\mathrm{II}-4\mathrm{I}} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -21 \end{pmatrix}$$

$$\xmapsto{\mathrm{III}\to\mathrm{III}-2\mathrm{II}} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & -9 \end{pmatrix}$$

Damit gilt dann

$$\det(A) = d \cdot 1 \cdot (-3) \cdot (-9) = -27.$$

Beispiel 8.17

Sei $A \in \mathrm{Mat}(n+1, K)$ gegeben durch

$$A = \begin{pmatrix} 0 & 1 & 2 & 3 & \dots & n \\ 1 & 0 & 1 & 2 & \dots & n-1 \\ 2 & 1 & 0 & 1 & \dots & n-2 \\ 3 & 2 & 1 & 0 & \dots & n-3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ n & n-1 & n-2 & n-3 & \dots & 0 \end{pmatrix}.$$

Ziehe für $i = 1, \dots, n$ von der i-ten Zeile die $(i+1)$-te Zeile ab. Wir erhalten:

$$A' = \begin{pmatrix} -1 & 1 & 1 & 1 & \dots & 1 & 1 \\ -1 & -1 & 1 & 1 & \dots & 1 & 1 \\ -1 & -1 & -1 & 1 & \dots & 1 & 1 \\ -1 & -1 & -1 & -1 & \dots & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -1 & -1 & -1 & -1 & \dots & -1 & 1 \\ n & n-1 & n-2 & n-3 & \dots & 1 & 0 \end{pmatrix}.$$

Addiere nun für $i = 2, \ldots, n+1$ die erste Spalte zur i-ten Spalte. Dann erhalten wir:

$$A'' = \begin{pmatrix} -1 & 0 & 0 & 0 & \ldots & 0 & 0 \\ * & -2 & 0 & 0 & \ldots & 0 & 0 \\ * & * & -2 & 0 & \ldots & 0 & 0 \\ * & * & * & -2 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ & * & * & * & \ldots & -2 & 0 \\ * & * & * & * & \ldots & * & n \end{pmatrix}.$$

Es folgt:

$$\det(A) = \det(A'') = (-1) \cdot (-2)^{n-1} \cdot n = -n \cdot (-2)^{n-1}.$$

Bemerkung 8.18

In Beispiel 8.17 haben wir durch ganz wenige Zeilen- und Spaltenoperationen die Matrix in Dreiecksgestalt überführt. Das lag aber an der speziellen Struktur der Matrix. Im allgemeinen Fall braucht der oben beschriebene Algorithmus zur Berechnung der Determinante mit Hilfe des Gauß-Algorithmus $\sim \frac{n^3}{3}$ Multiplikationen für eine $n \times n$-Matrix. In der Definition der Determinante tauchen dagegen $n!$ Summanden von je n Produkten auf, mit $n! \sim \left(\frac{n}{e}\right)^n$, wobei e die Eulersche Zahl ist. Man sagt, daß der Gauß-Algorithmus *polynomial*, die Definition aber *exponentiell* in der Größe der Matrix ist. Grundsätzlich gelten polynomiale Algorithmen als effizient, exponentielle dagegen als unakzeptabel ineffizient. Allerdings gibt es Fälle, wo keine polynomialen Algorithmen bekannt sind.

D) Der Determinantenmultiplikationssatz und der Kästchensatz

Satz 8.19 (Determinantenmultiplikationssatz)

Für Matrizen $A, B \in \mathrm{Mat}_n(K)$ *gilt*

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

Beweis: Wähle $A \in \mathrm{Mat}_n(K)$ fest und betrachte die Abbildung

$$f : \mathrm{Mat}_n(K) \longrightarrow K : B \mapsto \det(A \circ B).$$

f ist multilinear bezüglich der Spalten von B, da A auf jede Spalte von B linear wirkt. Außerdem ist f alternierend, da mit B auch $A \circ B$ zwei gleiche Spalten hat. Damit folgt aus Satz 8.11:

$$\det(A \circ B) = f(B) = f(\mathbb{1}_n) \cdot \det(B) = \det(A) \cdot \det(B).$$

□

Beispiel 8.20

In Beispiel 8.2 b. gilt

$$A = \begin{pmatrix} 5 & 2 \\ 6 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \circ \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R}),$$

so folgt aus dem Determinantenmultiplikationssatz 8.19 und weil die beiden Matrizen auf der rechten Seite Dreiecksmatrizen sind:

$$\det\begin{pmatrix} 5 & 2 \\ 6 & 3 \end{pmatrix} = \det\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \cdot \det\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} = 3 \cdot 1 = 3.$$

Das folgende Korollar ist eine Verallgemeinerung der Aussage in Aufgabe 1.14.

Korollar 8.21 (Determinante und Invertierbarkeit)
Genau dann ist $A \in \mathrm{Mat}_n(K)$ *invertierbar, wenn* $\det(A) \neq 0$. *In diesem Fall gilt*

$$\det\left(A^{-1}\right) = \frac{1}{\det(A)}.$$

Beweis: Ist A invertierbar, so gilt

$$1 = \det(\mathbb{1}_n) = \det\left(A \circ A^{-1}\right) = \det(A) \cdot \det\left(A^{-1}\right).$$

Dies zeigt, daß $\det(A)$ nicht Null sein kann, und zudem ist damit die obige Formel bewiesen.

Ist A nicht invertierbar, so sind die Spalten von A linear abhängig und durch mehrfache Addition von Vielfachen bestimmter Spalten zu einer anderen können wir eine Nullspalte erzeugen. Nach Korollar 8.13 c. ändert sich dabei der Wert der Determinante nicht, und nach Korollar 8.13 d. muß er somit 0 sein. □

Beispiel 8.22
Die Matrix A in Beispiel 8.20 ist invertierbar und ihre Inverse hat Determinante $\frac{1}{3}$. Dies wissen wir, ohne die Inverse auszurechnen. Diese können wir mit Hilfe von Aufgabe 1.14 berechnen. Für eine invertierbare 2×2-Matrix gilt

$$B = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Mat}_n(K)$$

gilt

$$B^{-1} = \frac{1}{ad - bc} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{\det(B)} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

so daß wir im Beispiel

$$A^{-1} = \frac{1}{3} \cdot \begin{pmatrix} 3 & -2 \\ -6 & 5 \end{pmatrix}$$

erhalten.

Bemerkung 8.23 (det ist ein Gruppenepimorphismus.)
In der Sprache der Algebra folgt aus Satz 8.19 und Korollar 8.21, daß

$$\det : \left(\mathrm{Gl}_n(K), \circ\right) \to \left(K^*, \cdot\right)$$

ein Gruppenepimorphismus ist. Dazu beachte man, daß det surjektiv ist wegen

$$\det\left(\begin{array}{c|ccc} \lambda & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & \mathbb{1}_{n-1} & \\ 0 & & & \end{array}\right) = \lambda.$$

Satz 8.24 (Kästchensatz)
Es sei $A \in \mathrm{Mat}_n(K)$ *eine* Blockmatrix *der Form*

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

mit $B \in \mathrm{Mat}_k(K)$, $C \in \mathrm{Mat}(k \times l, K)$, $D \in \mathrm{Mat}_l(K)$, $0 \in \mathrm{Mat}(l \times k, K)$ *und* $n = k+l$. *Dann gilt:*

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

Beweis: Man beachte, daß

$$A = \left(\begin{array}{c|c} B & C \\ \hline 0 & D \end{array}\right) = \left(\begin{array}{c|c} \mathbb{1}_k & 0 \\ \hline 0 & D \end{array}\right) \circ \left(\begin{array}{c|c} B & C \\ \hline 0 & \mathbb{1}_l \end{array}\right).$$

Wegen des Determinantenmultiplikationssatzes 8.19 reicht es mithin zu zeigen:

$$\det\left(\begin{array}{c|c} \mathbb{1}_k & 0 \\ \hline 0 & D \end{array}\right) = \det(D) \tag{22}$$

und

$$\det\left(\begin{array}{c|c} B & C \\ \hline 0 & \mathbb{1}_l \end{array}\right) = \det(B). \tag{23}$$

Die Abbildung

$$f : \mathrm{Mat}_l(K) \to K : D' \mapsto \det\left(\begin{array}{c|c} \mathbb{1}_k & 0 \\ \hline 0 & D' \end{array}\right)$$

ist offensichtlich multilinear und alternierend, und wegen Satz 8.11 b. gilt mithin

$$\det\left(\begin{array}{c|c} \mathbb{1}_k & 0 \\ \hline 0 & D \end{array}\right) = f(D) = f(\mathbb{1}_l) \cdot \det(D) = \det(\mathbb{1}_n) \cdot \det(D) = \det(D),$$

d. h. (22) ist erfüllt.

Analog ist die Abbildung

$$g : \mathrm{Mat}_k(K) \to K : B' \mapsto \det\left(\begin{array}{c|c} B' & C \\ \hline 0 & \mathbb{1}_l \end{array}\right)$$

alternierend und multilinear in den Spalten von B', also eine Volumenform. Wieder folgt aus Satz 8.11 mit Hilfe von Proposition 8.3, daß

$$\det\left(\begin{array}{c|c} B & C \\ \hline 0 & \mathbb{1}_l \end{array}\right) = g(B) = g(\mathbb{1}_k) \cdot \det(B) = \det\left(\begin{array}{c|c} \mathbb{1}_k & C \\ \hline 0 & \mathbb{1}_l \end{array}\right) \cdot \det(B) = \det(B),$$

womit auch (23) gezeigt ist. □

Beispiel 8.25 (Vandermonde-Determinante)
Wir wollen mit Hilfe des Gauß-Algorithmus zeigen, daß

$$\det\begin{pmatrix} 1 & a_0 & a_0^2 & a_0^3 & \ldots & a_0^n \\ 1 & a_1 & a_1^2 & a_1^3 & \ldots & a_1^n \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & a_n^3 & \ldots & a_n^n \end{pmatrix} = \prod_{0\le i<j\le n}(a_j - a_i)$$

für $a_0, \ldots, a_n \in K$ und $n \geq 1$ gilt. Die Determinate dieser Matrix ist als *Vandermonde-Determinante* bekannt.

Beweis: Wir beweisen die Aussage mit Hilfe von Induktion nach n. Für $n = 1$ ist

$$\det(A) = \begin{vmatrix} 1 & a_0 \\ 1 & a_1 \end{vmatrix} = a_1 - a_0$$

und die Aussage stimmt. Sei also $n > 1$ und die Aussage sei für Matrizen dieser Gestalt der Größe n (beachte, daß A die Größe $n+1$ hat) bereits gezeigt. Addieren wir für $j = n+1, \ldots, 2$ zur j-ten Spalte das $-a_0$-fache der $j-1$-ten Spalte, so ändert sich die Determinante nicht und wir erhalten

$$\det(A) = \begin{vmatrix} 1 & a_0 & a_0^2 & a_0^3 & \ldots & a_0^n \\ 1 & a_1 & a_1^2 & a_1^3 & \ldots & a_1^n \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & a_n^3 & \ldots & a_n^n \end{vmatrix}$$

$$= \left|\begin{array}{c|ccccc} 1 & 0 & 0 & 0 & \ldots & 0 \\ \hline 1 & a_1 - a_0 & a_1^2 - a_1a_0 & a_1^3 - a_1^2a_0 & \ldots & a_1^n - a_1^{n-1}a_0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n - a_0 & a_n^2 - a_na_0 & a_n^3 - a_n^2a_0 & \ldots & a_n^n - a_n^{n-1}a_0 \end{array}\right|$$

Aufgrund des Kästchensatzes und wegen $\det(1) = 1$ gilt dann

$$\det(A) = \begin{vmatrix} a_1 - a_0 & a_1^2 - a_1a_0 & a_1^3 - a_1^2a_0 & \ldots & a_1^n - a_1^{n-1}a_0 \\ \vdots & \vdots & \vdots & & \vdots \\ a_n - a_0 & a_n^2 - a_na_0 & a_n^3 - a_n^2a_0 & \ldots & a_n^n - a_n^{n-1}a_0 \end{vmatrix}$$

$$= \begin{vmatrix} a_1 - a_0 & (a_1 - a_0)\cdot a_1 & (a_1 - a_0)\cdot a_1^2 & \ldots & (a_1 - a_0)\cdot a_1^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ a_n - a_0 & (a_n - a_0)\cdot a_n & (a_n - a_0)\cdot a_n^2 & \ldots & (a_n - a_0)\cdot a_n^{n-1} \end{vmatrix}$$

Klammern wir nun in der i-ten Zeile $a_i - a_0$ aus, so erhalten wir

$$\det(A) = \prod_{i=1}^{n}(a_i - a_0)\cdot \begin{vmatrix} 1 & a_1 & a_1^2 & a_1^3 & \ldots & a_1^{n-1} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & a_n^3 & \ldots & a_n^{n-1} \end{vmatrix}.$$

Auf die letzte Determinante können wir Induktion anwenden und erhalten

$$\det(A) = \prod_{i=1}^{n}(a_i - a_0) \cdot \prod_{1\leq i<j\leq n}(a_j - a_i) = \prod_{0\leq i<j\leq n}(a_j - a_i).$$

□

E) Laplacescher Entwicklungssatz und Cramersche Regel

Wir kommen jetzt zu einer alternativen Berechnung der Determinante. Im Gegensatz zum Gaußalgorithmus kommt sie ohne Division aus und funktioniert deshalb über jedem kommutativen Ring mit Eins (siehe Bemerkung 8.37). Zu ihrer Herleitung führen wir zunächst verschiedene Hilfmatrizen ein.

Definition 8.26
Es sei $A = (a_{ij}) = (a^1 \ldots a^n) \in \mathrm{Mat}_n(K)$, $n \geq 2$, und $b = (b_1, \ldots, b_n)^t \in K^n$.

Wir definieren die *Ersetzungsmatrix*

$$A_i(b) := (a^1 \ldots a^{i-1}\; b\; a^{i+1} \ldots a^n),$$

in der die i-te Spalte von A durch b ersetzt wurde.

Ist $b = e_j$ der j-te Einheitsvektor, so gilt:

$$A_i(e_j) = \begin{pmatrix} a_{11} & \ldots & 0 & \ldots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{j1} & \ldots & 1 & \ldots & a_{jn} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \ldots & 0 & \ldots & a_{nn} \end{pmatrix}.$$

Ersetzen wir in $A_i(e_j)$ zusätzlich noch die j-te Zeile durch den i-ten Einheitsvektor, dann erhält man die Matrix

$$S_{ji}(A) = \begin{pmatrix} a_{11} & \ldots & 0 & \ldots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ 0 & \ldots & 1 & \ldots & 0 \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \ldots & 0 & \ldots & a_{nn} \end{pmatrix}.$$

Streicht man in der Matrix A die j-te Zeile und die i-te Spalte, so erhält man die *Streichungsmatrix*

$$A_{ji} = \left(\begin{array}{ccc|ccc} a_{11} & \ldots & a_{1\,i-1} & a_{1\,i+1} & \ldots & a_{1n} \\ \vdots & & \vdots & & \vdots & \\ a_{j-1\,1} & \ldots & a_{j-1\,i-1} & a_{j-1\,i+1} & \ldots & a_{j-1\,n} \\ \hline a_{j+1\,1} & \ldots & a_{j+1\,i-1} & a_{j+1\,i+1} & \ldots & a_{j+1\,n} \\ \vdots & & \vdots & & \vdots & \\ a_{n\,1} & \ldots & a_{n\,i-1} & a_{n\,i+1} & \ldots & a_{n\,n} \end{array}\right).$$

Lemma 8.27
Für $A \in \mathrm{Mat}_n(K)$, $n \geq 2$, $1 \leq i, j \leq n$, *gilt:*

$$\det\Big(A_i\big(e_j\big)\Big) = \det\big(S_{ji}(A)\big) = (-1)^{i+j}\det(A_{ji}).$$

Beweis: $S_{ji}(A)$ entsteht aus $A_i\big(e_j\big)$ durch Subtraktion des a_{jk}-fachen der i-ten Spalte von der k-ten Spalte, $k \in \{1, \ldots, n\} \setminus \{i\}$. Also gilt nach Korollar 8.13:

$$\det\Big(A_i\big(e_j\big)\Big) = \det\big(S_{ji}(A)\big).$$

Durch $i - 1$ Spaltenvertauschungen und $j - 1$ Zeilenvertauschungen entsteht aus $S_{ji}(A)$ die Matrix

$$\left(\begin{array}{c|ccc} 1 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & A_{ji} & \\ 0 & & & \end{array}\right).$$

Also folgt aus dem Kästchensatz 8.24 unter Beachtung der Korollare 8.13 und 8.14

$$\det\big(S_{ji}(A)\big) = (-1)^{i+j}\det(A_{ji}).$$

□

In der folgenden Definition beachte man die Vertauschung der Indizes!

Definition 8.28
Für $A \in \mathrm{Mat}_n(K)$, $n \geq 2$, $1 \leq i, j \leq n$ heißt

$$a_{ij}^{\#} := (-1)^{i+j}\det(A_{ji})$$

ein *Kofaktor* von A. Die Matrix der Kofaktoren

$$A^{\#} := \big(a_{ij}^{\#}\big) \in \mathrm{Mat}_n(K)$$

heißt die *Adjunkte* oder *Komplementärmatrix* von A.

Satz 8.29 (Satz über die Adjunkte)
Für $A \in \mathrm{Mat}_n(K)$, $n \geq 2$, *gilt:*

$$A^{\#} \circ A = A \circ A^{\#} = \det(A) \cdot \mathbb{1}_n.$$

Beweis: Sei $A^{\#} \circ A = (c_{ik})$. Dann gilt mit Lemma 8.27:

$$\begin{aligned} c_{ik} = \sum_{j=1}^{n} a_{ij}^{\#} \cdot a_{jk} &= \sum_{j=1}^{n} a_{jk} \cdot \det\big(a^1 \ldots a^{i-1}\ e_j\ a^{i+1} \ldots a^n\big) \\ &= \det\left(a^1 \ldots a^{i-1}\ \sum_{j=1}^{n} a_{jk}e_j\ a^{i+1} \ldots a^n\right) \\ &= \det\big(a^1 \ldots a^{i-1}\ a^k\ a^{i+1} \ldots a^n\big) = \delta_{ik} \cdot \det(A), \end{aligned}$$

wobei δ_{ik} das Kronecker-Symbol ist. Das dritte Gleichheitszeichen folgt aus der Multilinearität von det, das letzte, da det alternierend ist.

Der Beweis, daß $A \circ A^{\#} = \det(A) \cdot \mathbb{1}_n$ geht analog. □

Korollar 8.30
Es sei $A \in \mathrm{Mat}_n(K)$ *invertierbar, so ist*

$$A^{-1} = \frac{1}{\det(A)} \cdot A^{\#}.$$

Wir wollen an dieser Stelle einmal die vielen Aussagen, die zur Invertierbarkeit einer quadratischen Matrix über einem Körper äquivalent sind, sammeln.

Korollar 8.31
Für eine Matrix $A \in \mathrm{Mat}(n, K)$ *sind gleichwertig:*

a. A *ist invertierbar.*

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

c. $\det(A) \neq 0$.

d. f_A *ist bijektiv.*

e. f_A *ist injektiv.*

f. f_A *ist surjektiv.*

g. $\mathrm{rZSF}(A) = \mathbb{1}_n$.

h. A *ist das Produkt endlich vieler Elementarmatrizen.*

i. *Es gibt eine Matrix* $B \in \mathrm{Mat}(n, K)$ *mit* $A \circ B = \mathbb{1}_n$.

Beweis: Die unterschiedlichen Äquivalenzen sind in den Sätzen 3.33, 4.22, 4.23, 5.27, 6.14, 7.2 und 8.21 gezeigt worden. □

Beispiel 8.32
Für eine 2×2-Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

gilt $\det(A) = ad - bc$ und

$$A^{\#} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Ist also $ad - bc \neq 0$, so gilt:

$$A^{-1} = \frac{1}{ad - bc} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Damit ist Aufgabe 1.14 bewiesen.

Sei nun konkret $K = \mathbb{Q}$ und

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

Dann ist $\det(A) = 1$ und somit gilt

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

Der Satz über die Adjunkte führt zu einer rekursiven Berechnungsformel für die Determinante, die für theoretische Überlegungen sehr nützlich ist. Sie ist auch als rekursive Prozedur sehr einfach zu programmieren, aber nicht sehr effizient. Sie hat die gleiche Komplexität, wie die Leibnizsche Formel (20) zur Definition der Determinante.

Satz 8.33 (Laplacescher Entwicklungssatz)
Es sei $A \in \mathrm{Mat}_n(K)$.

a. *Wir nennen die folgende Formel, die* Entwicklung nach der i-ten Zeile*:*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det(A_{ij}). \tag{24}$$

b. *Entsprechend nennen wir die folgende Formel, die* Entwicklung nach der j-ten Spalte*:*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det(A_{ij}). \tag{25}$$

Beweis: Nach Satz 8.29 gilt für $A \circ A^{\#} = (c_{ik})$

$$\det(A) = c_{ii} = \sum_{j=1}^{n} a_{ij} \cdot a^{\#}_{ji} = \sum_{j=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det(A_{ij}).$$

Damit folgt (24), und (25) zeigt man analog durch die Betrachtung von $A^{\#} \circ A$. □

Bemerkung 8.34
Entwickelt man $A = (a_{ij})$ nach der ersten Zeile, so gilt:

$$\begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} = a_{11} \cdot \begin{vmatrix} a_{22} & \ldots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \ldots & a_{nn} \end{vmatrix} - a_{12} \cdot \begin{vmatrix} a_{21} & a_{23} & \ldots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n3} & \ldots & a_{nn} \end{vmatrix}$$
$$+ \ldots + (-1)^{n+1} a_{1n} \cdot \begin{vmatrix} a_{21} & \ldots & a_{2\,n-1} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{n\,n-1} \end{vmatrix}.$$

Benutzt man dieses Verfahren, so entwickelt man am Besten nach Zeilen bzw. Spalten, die möglichst viele Nullen enthalten. Die Vorzeichen merkt man sich am

Günstigsten mit der sogenannten *Schachbrettregel*:

$$\begin{vmatrix} + & - & + & - & \dots \\ - & + & - & + & \dots \\ + & - & + & - & \dots \\ - & + & - & + & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{vmatrix}$$

Für kleine Matrizen, insbesondere wenn die Matrix dünn besetzt ist, ist dieses Verfahren zur Berechnung der Determinante (und zur Berechnung der Inversen) durchaus anwendbar. Für größere Matrizen ist auf jeden Fall der Gaußsche Eliminationsalgorithmus vorzuziehen.

Wir berechnen nun die Determinante der Matrix

$$A = \begin{pmatrix} 0 & 2 & 0 \\ 1 & 3 & 4 \\ 2 & 5 & 3 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$$

mit Hilfe der Entwicklung nach der ersten Zeile. Dann gilt

$$\det(A) = 0\cdot\det(A_{11}) - 2\cdot\det(A_{12}) + 0\cdot\det(A_{13}) = -2\cdot\det\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} = (-2)\cdot(-5) = 10.$$

Algorithmus 8.35 (Laplace-Entwicklung)
INPUT: $A \in \mathrm{Mat}_n(K)$.

OUTPUT: $\det(A)$.

1. **Schritt:** Initialisiere det auf Null.
2. **Schritt:** Falls $n = 1$, setze $\det = a_{11}$ und gehe zu Schritt 3. Sonst tue für $i = 1, \ldots, n$:
   - Bilde eine Hilfsmatrix B durch Streichen der ersten Spalte und der i-ten Zeile von A.
   - Rufe den Algorithmus mit B auf und merke Dir das Ergebnis in einer Hilfsvariablen x.
   - Addiere zu det die Zahl $(-1)^{i+1} \cdot a_{i1} \cdot x$.
3. **Schritt:** Gib det zurück.

Der Satz über die Adjunkte liefert auch eine für theoretische Überlegungen sehr wichtige geschlossene Formel für die Lösungen eines linearen Gleichungssystems. Dies ist die berühmte *Cramersche Regel*. Wir werden sie in der mehrdimensionalen Analysis nutzen, um zu sehen, daß die Lösung eines eindeutig lösbaren linearen Gleichungssystems *stetig* von den Koeffizienten der erweiterten Koeffizientenmatrix (A, b) abhängt und somit kleine Störungen der Einträge nur zu kleinen Störungen in der Lösung führen.

Satz 8.36 (Cramersche Regel)
Es sei $A \in \mathrm{Mat}_n(K)$ *invertierbar und* $b \in K^n$.

Für die eindeutig bestimmte Lösung $x = (x_1, \ldots, x_n)^t \in K^n$ *von* $Ax = b$ *gilt dann*

$$\begin{aligned} x_i &= \tfrac{1}{\det(A)} \cdot \det\big(A_i(b)\big) \\ &= \tfrac{1}{\det(A)} \cdot \det \begin{pmatrix} a_{11} & \ldots & a_{1\,i-1} & b_1 & a_{1\,i+1} & \ldots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \ldots & a_{n\,i-1} & b_n & a_{n\,i+1} & \ldots & a_{nn} \end{pmatrix}. \end{aligned}$$

Beweis: Wegen Korollar 8.30 ist

$$x = A^{-1}b = \frac{1}{\det(A)} \cdot A^{\#}b$$

die eindeutig bestimmte Lösung des linearen Gleichungssystems. Also folgt mit Lemma 8.27 und der Multilinearität der Determinante

$$\begin{aligned} x_i &= \tfrac{1}{\det(A)} \cdot \sum_{j=1}^{n} a_{ij}^{\#} \cdot b_j = \tfrac{1}{\det(A)} \cdot \sum_{j=1}^{n} \det\big(A_i(e_j)\big) \cdot b_j \\ &= \tfrac{1}{\det(A)} \cdot \sum_{j=1}^{n} \det\big(a^1 \ldots a^{i-1}\; e_j\; a^{i+1} \ldots a^n\big) \cdot b_j \\ &= \tfrac{1}{\det(A)} \cdot \det\big(a^1 \ldots a^{i-1}\; b\; a^{i+1} \ldots a^n\big) \\ &= \tfrac{1}{\det(A)} \cdot \det\big(A_i(b)\big). \end{aligned}$$

□

Bemerkung 8.37 (Determinanten über kommutativen Ringen mit Eins)
Ist K nur ein kommutativer Ring mit Eins, so können wir die Determinante einer Matrix in $\mathrm{Mat}_n(K)$ ebenfalls durch die Leibniz-Formel definieren, und alle Aussagen dieses Abschnitts, die *ohne Division* auskommen, gelten mit dem gleichen Beweis.

Wir können den Gauß-Algorithmus 8.15 über beliebigen Ringen in der angegebenen Form *nicht* mehr anwenden, da dabei Divisionen nötig sind. Außerdem gilt Korollar 8.21 *nicht* mehr in der angegebenen Form, und ebenso gilt Korollar 8.31 in *nicht* in vollem Umfang.

Alle anderen Aussagen gelten jedoch ohne jede Änderung. Dies trifft insbesondere auf den Satz zur Adjunkten 8.29 zu, den wir später für Matrizen mit Koeffizienten in einem Polynomring anwenden wollen. Außerdem können wir den Laplaceschen Entwicklungssatz im Gegensatz zum Gaußschen Algorithmus über jedem kommutativen Ring mit Eins anwenden, um die Determinante auszurechnen. Es gibt aber auch hier geschicktere Verfahren, indem man den Gaußschen Algorithmus abwandelt zum sogenannten Bareiss Algorithmus (siehe [**Coh96**]).

Für die Aussage in Korollar 8.30 beachte man, daß aus dem Determinantenmultiplikationssatz 8.19 und dem Satz zur Adjunkten 8.29 unmittelbar folgt, daß eine quadratische Matrix über einem kommutativen Ring genau dann invertierbar ist, wenn $\det(A)$ invertierbar ist. In diesem Fall darf man dann auch in dem Ring durch $\det(A)$ teilen. Das trifft auf Korollar 8.30 ebenso zu wie auf die Cramersche Regel 8.36.

Betrachten wir konkret den Ring $\mathbb{Z}$ der ganzen Zahlen, dann sind nur 1 und -1 invertierbar. Mithin sind nur ganzzahlige Matrizen mit Determinante 1 oder -1 über $\mathbb{Z}$ invertierbar, d.h. nur für solche enthält die Inverse wieder nur ganze Zahlen. Ein Beispiel dafür haben wir in Beispiel 8.32 gesehen. Betrachten wir stattdessen die Matrix

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix},$$

so gilt $\det(A) = -2 \notin \{1, -1\}$ und die Einträge von

$$A^{-1} = -\frac{1}{2} \cdot \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$$

sind nicht mehr alle ganzzahlig, obwohl A nur ganzzahlige Einträge hatte. A ist als Matrix in $\mathrm{Mat}_2(\mathbb{Q})$ also invertierbar, als Matrix in $\mathrm{Mat}_2(\mathbb{Z})$ aber nicht.

Aufgaben

Aufgabe 8.38
Berechne die Determinanten der folgenden Matrizen:

a. $\begin{pmatrix} 1 & -4 \\ -3 & 8 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$.

b. $\begin{pmatrix} -1 & 0 & -1 \\ 0 & 5 & 0 \\ 3 & 0 & 3 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R})$.

c. $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R})$.

Aufgabe 8.39
Sei K ein Körper und $\lambda \in K$. Bestimme die Determinante der Matrix

$$\begin{pmatrix} 1 & \lambda & \lambda^2 & \ldots & \lambda^{n-1} \\ \lambda^{n-1} & 1 & \lambda & \ldots & \lambda^{n-2} \\ \lambda^{n-2} & \lambda^{n-1} & 1 & \ldots & \lambda^{n-3} \\ \vdots & \vdots & \vdots & & \vdots \\ \lambda & \lambda^2 & \lambda^3 & \ldots & 1 \end{pmatrix} \in \mathrm{Mat}(n \times n, K).$$

Aufgabe 8.40
Für $n \in \mathbb{N} \setminus \{0\}$ definieren wir

$$A_n = \begin{pmatrix} 1 & 1 & 0 & \dots & \dots & \dots & 0 \\ 1 & 1 & 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & 1 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 1 & 1 & 0 \\ 0 & \dots & \dots & 0 & 1 & 1 & 1 \\ 0 & \dots & \dots & \dots & 0 & 1 & 1 \end{pmatrix} \in \operatorname{Mat}_n(\mathbb{R})$$

als die Matrix, deren Einträge auf der Diagonalen sowie auf der oberen und unteren Nebendiagonalen alle eins sind, während alle anderen Einträge null sind. Ferner setzen wir $d_n = \det(A_n)$.

a. Zeige, für $n \geq 3$ gilt die Rekursionsformel $d_n = d_{n-1} - d_{n-2}$.

b. Zeige, für $k \in \mathbb{N}$ gilt

$$d_n = \begin{cases} 1, & \text{falls } n \equiv 1 (\text{mod } 6) \text{ oder } n \equiv 0 (\text{mod } 6), \\ 0, & \text{falls } n \equiv 2 (\text{mod } 6) \text{ oder } n \equiv 5 (\text{mod } 6), \\ -1, & \text{falls } n \equiv 3 (\text{mod } 6) \text{ oder } n \equiv 4 (\text{mod } 6). \end{cases}$$

Aufgabe 8.41
Sei V ein n-dimensionaler $\mathbb{C}$-Vektorraum und $f : V \longrightarrow V$ $\mathbb{C}$-linear. Mittels Einschränkung der Skalarmultiplikation können wir V als $\mathbb{R}$-Vektorraum und f als $\mathbb{R}$-lineare Abbildung auffassen. Des Weiteren bezeichnen wir mit $\det_{\mathbb{C}}(f)$ die Determinante von f als $\mathbb{C}$-lineare Abbildung und $\det_{\mathbb{R}}(f)$ die Determinante von f als $\mathbb{R}$-lineare Abbildung. Zeige:

$$\det_{\mathbb{R}}(f) = |\det_{\mathbb{C}}(f)|^2.$$

Hinweis: Für eine $\mathbb{C}$-Basis $(v_1, \dots, v_n)$ von V betrachte man die zugehörige $\mathbb{R}$-Basis $(v_1, \dots, v_n, iv_1, \dots, iv_n)$ sowie jeweils die zugehörige Matrixdarstellung von f. Wem der allgemeine Fall zu schwer ist, der beschränke sich auf die Abbildung $f : \mathbb{C} \longrightarrow \mathbb{C} : z \longmapsto (a + ib) \cdot z$ mit $a, b \in \mathbb{R}$ fest vorgegeben. Was ist eine Basis von $\mathbb{C}$ als $\mathbb{R}$-Vektorraum?

Aufgabe 8.42
Berechne die Determinante folgender Matrix mit Hilfe des Gauß-Algorithmus':

$$A = \begin{pmatrix} 1 & 3 & 0 & 2 & 1 \\ -2 & 0 & 3 & 1 & 3 \\ 0 & 2 & 5 & 0 & -1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 3 & 4 & 1 \end{pmatrix} \in \operatorname{Mat}_5(\mathbb{R}).$$

Aufgabe 8.43

Berechne die folgende Determinante mit Hilfe des Determinantenentwicklungssatzes:

$$B = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 3 & 1 & 3 \\ 2 & 5 & 0 & -1 \\ 1 & 0 & 1 & 1 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Q}).$$

Aufgabe 8.44

Bestimme für welche $s \in \mathbb{R}$ die Abbildung

$$f_s : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 : (x, y, z)^t \mapsto (x + z, x + 2y + z, sx + y - z)^t$$

invertierbar ist und berechne für diese die Inverse mit Hilfe der Adjunkten.

Aufgabe 8.45

Es sei $A \in \mathrm{Mat}_n(\mathbb{R})$ mit ungeradem n und $A^t = -A$. Zeige, A ist nicht invertierbar. Bleibt die Aussage wahr, wenn wir $\mathbb{R}$ durch einen anderen Körper ersetzen?

Aufgabe 8.46

Zeige, ist $A \in \mathrm{Mat}_n(K)$ eine invertierbare obere Dreiecksmatrix, so ist auch A^{-1} eine obere Dreiecksmatrix.

Aufgabe 8.47

Löse das folgende lineare Gleichungssystem mit Hilfe der Cramerschen Regel:

$$x + 2z = 3, \quad 3x + y = 5 \quad \text{und} \quad -x + y = 1.$$

KAPITEL II

Normalformen von Endomorphismen

§ 9 Endomorphismen und ihre Eigenwerte

In diesem Abschnitt sei V ein K-Vektorraum mit $1 \leq \dim_K(V) = n < \infty$.

A) Invarianten von Endomorphismen unter Konjugation

Bemerkung 9.1 (Endomorphismen)
Wir erinnern uns, daß K-lineare Abbildungen

$$f : V \longrightarrow V$$

auch *Endomorphismen* des K-Vektorraums V genannt werden (siehe Definition 2.19) und daß

$$\mathrm{End}_K(V) = \{f : V \longrightarrow V \mid f \text{ ist } K\text{-linear}\}$$

die K-*Algebra* der Endomorphismen von V ist (siehe Bemerkung 5.9).

Zudem wissen wir, wie sich die Matrixdarstellungen von Endomorphismen unter Basiswechsel verhalten. Sind B und D zwei Basen des Vektorraums V und ist $T = T^D_B$, so gilt (siehe Korollar 5.15)

$$M^D_D(f) = T^{-1} \circ M^B_B(f) \circ T.$$

Dabei ist es von großer Wichtigkeit, daß wir jeweils im Definitions- und Zielbereich von f *dieselbe* Basis verwenden, und das wollen wir von nun an stets tun, wenn wir Matrixdarstellungen von Endomorphismen betrachten!

Wir können deshalb Eigenschaften von Matrizen, die unter Transformationen der Form

$$A \mapsto T^{-1} \circ A \circ T$$

erhalten bleiben, auch für Endomorphismen definieren, indem wir dazu ihre Matrixdarstellungen bezüglich einer beliebigen Basis verwenden. In diesem Abschnitt wollen wir einige Beispiele hierfür kennen lernen.

Definition 9.2 (Konjugiert oder ähnlich)
Zwei quadratische Matrizen $A, B \in \mathrm{Mat}_n(K)$ heißen *konjugiert* oder *ähnlich*, wenn es eine invertierbare Matrix $T \in \mathrm{Gl}_n(K)$ gibt, so daß $B = T^{-1} \circ A \circ T$ ist.

Bemerkung 9.3 (Konjugation ist eine Äquivalenzrelation)
Konjugation von Matrizen ist ein Beispiel für eine Äquivalenzrelation auf der Menge $\mathrm{Mat}_n(K)$ der quadratischen $n \times n$-Matrizen über K. D.h.

- jede Matrix ist zu sich selbst konjugiert, denn $A = \mathbb{1}_n^{-1} \cdot A \cdot \mathbb{1}_n$;
- ist A zu B konjugiert, so ist auch B zu A konjugiert, da aus $B = T^{-1} \circ A \circ T$ auch $A = (T^{-1})^{-1} \circ B \circ T^{-1}$ folgt;
- ist A zu B und B zu C konjugiert, so ist auch A zu C konjugiert, da aus $B = T^{-1} \circ A \circ T$ und $C = S^{-1} \circ B \circ S$ auch $C = (T \circ S)^{-1} \circ A \circ (T \circ S)$ folgt.

Definition 9.4 (Das Charakteristische Polynom einer Matrix)
Es sei $A = (a_{ij}) \in \mathrm{Mat}_n(K)$ eine quadratische Matrix.

a. Die Summe der Diagonaleinträge von A heißt die *Spur* von A,

$$\mathrm{Spur}(A) = a_{11} + a_{22} + \ldots + a_{nn}.$$

b. Wir nennen

$$\chi_A := \det(t \cdot \mathbb{1}_n - A) \in K(t)$$

das *charakteristische Polynom* von A, wobei

$$t \cdot \mathbb{1}_n - A = \begin{pmatrix} t - a_{11} & -a_{12} & \ldots & -a_{1n} \\ -a_{21} & t - a_{22} & \ldots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \ldots & t - a_{nn} \end{pmatrix} \in \mathrm{Mat}_n\big(K(t)\big)$$

eine quadratische Matrix mit Polynomen als Einträgen ist, die wir als Elemente des Körper $K(t)$ der rationalen Funktionen auffassen (siehe Bemerkung B7.21). In Proposition 9.6 zeigen wir, daß χ_A in der Tat ein Polynom ist. Für Aussagen zu Polynomen verweisen wir auf Anhang B7.

Beispiel 9.5
Für die Matrix

$$A = \begin{pmatrix} 5 & 2 \\ 6 & 3 \end{pmatrix}$$

gilt

$$\chi_A = \det \begin{pmatrix} t-5 & -2 \\ -6 & t-3 \end{pmatrix} = (t-5) \cdot (t-3) - (-2) \cdot (-6) = t^2 - 8t + 3.$$

Man beachte, daß der konstante Term von χ_A gerade $\det(A) = 3$ und daß der Koeffizient von t gerade $-\mathrm{Spur}(A) = -8$ ist.

Proposition 9.6 (Charakteristisches Polynom)
Es sei $A \in \mathrm{Mat}_n(K)$ *eine quadratische Matrix. Dann ist*

$$\chi_A = t^n + \alpha_{n-1} \cdot t^{n-1} + \alpha_{n-2} \cdot t^{n-2} + \ldots + \alpha_1 \cdot t + \alpha_0 \in K[t]$$

ein normiertes Polynom vom Grad n *mit* $\alpha_{n-1} = -\mathrm{Spur}(A)$ *und* $\alpha_0 = (-1)^n \cdot \det(A)$.

Beweis: Ist $A = (a_{ij})$ und $t \cdot \mathbb{1}_n - A = \big(p_{ij}\big)$, dann folgt aus der Leibnitzschen Formel für die Determinante

$$\chi_A = \det(t \cdot \mathbb{1}_n - A) = (t - a_{11}) \cdots (t - a_{nn}) + \sum_{\mathrm{id} \neq \sigma \in \mathbb{S}_n} \operatorname{sgn}(\sigma) \cdot p_{1\sigma(1)} \cdots p_{n\sigma(n)}.$$

Da für $\sigma \neq \mathrm{id}$ mindestens zwei Faktoren in $p_{1\sigma(1)} \cdots p_{n\sigma(n)}$ konstante Polynome sind, ergibt $\sum_{\mathrm{id} \neq \sigma \in \mathbb{S}_n} \operatorname{sgn}(\sigma) \cdot p_{1\sigma(1)} \cdots p_{n\sigma(n)}$ ein Polynom vom Grad kleiner gleich $n - 2$. Damit lassen sich die Koeffizienten α_n und α_{n-1} von t^n und t^{n-1} in χ_A aus $(t - a_{11}) \cdots (t - a_{nn})$ herleiten und sind wie oben angegeben $\alpha_n = 1$ und

$$\alpha_{n-1} = -a_{11} - a_{22} - \ldots - a_{nn} = -\operatorname{Spur}(A).$$

Ferner ist

$$\alpha_0 = \chi_A(0) = \det(-A) = (-1)^n \cdot \det(A)$$

der konstante Term im charakteristischen Polynom. □

Bemerkung 9.7

Man beachte, daß es bei der Berechnung von $\chi_A(\lambda)$ für $\lambda \in K$ keinen Unterschied macht, ob wir zuerst t durch λ ersetzen und dann die Leibnitzformel zum Berechnen der Determinante anwenden oder ob wir zuerst die Determinante berechnen und dann t durch λ ersetzen. Das liegt daran, daß der Einsetzhomomorphismus mit der Multiplikation und Addition verträglich ist, vgl. Proposition B7.11. Diese Tatsache haben wir im obigen Beweis bei der Berechnung des konstanten Terms des charakteristischen Polynoms verwendet.

Proposition 9.8 (Konjugierte Matrizen haben dasselbe charakt. Polynom.)
Sind $A, B \in \operatorname{Mat}_n(K)$ *konjugiert, so gilt* $\chi_A = \chi_B$.

Insbesondere gilt auch $\det(A) = \det(B)$ *und* $\operatorname{Spur}(A) = \operatorname{Spur}(B)$.

Beweis: Sei $T \in \operatorname{Gl}_n(K)$ mit $B = T^{-1} \circ A \circ T$, dann gilt auch

$$T^{-1} \circ (t \cdot \mathbb{1}_n - A) \circ T = t \cdot T^{-1} \circ \mathbb{1}_n \circ T - T^{-1} \circ A \circ T = t \cdot \mathbb{1}_n - B.$$

Der Determinantenmultiplikationssatz 8.19 gilt auch für kommutative Ringe mit Eins (siehe Bemerkung 8.37), also insbesondere für Matrizen mit Einträgen im Polynomring, und somit folgt

$$\begin{aligned} \chi_B &= \det(t \cdot \mathbb{1}_n - B) = \det\left(T^{-1} \circ (t \cdot \mathbb{1}_n - A) \circ T\right) \\ &= \det(T^{-1}) \cdot \det(t \cdot \mathbb{1}_n - A) \cdot \det\left(T\right) \\ &= \frac{1}{\det(T)} \cdot \det(t \cdot \mathbb{1}_n - A) \cdot \det(T) = \chi_A. \end{aligned}$$

Die Aussage zur Determinante und zur Spur folgt unmittelbar aus Proposition 9.6. □

Damit können wir das charakteristische Polynom, die Determinante und die Spur eines Endomorphismus definieren.

Definition 9.9 (Das charakteristische Polynom eines Endomorphismus)
Sei V ein endlich-dimensionaler K-Vektorraum der Dimension mindestens Eins mit Basis B und $f \in \mathrm{End}_K(V)$. Wir definieren das *charakteristische Polynom* von f durch

$$\chi_f := \chi_{M_B^B(f)},$$

die *Determinante* von f durch

$$\det(f) := \det\left(M_B^B(f)\right)$$

und die *Spur* von f durch

$$\mathrm{Spur}(f) := \mathrm{Spur}\left(M_B^B(f)\right).$$

Da die Matrixdarstellungen eines Endomorphismus f zu verschiedenen Basen nach Korollar 5.15 konjugiert sind, sind diese Definitionen unter Berücksichtigung von Proposition 9.8 unabhängig von der Wahl der Basis B.

Beispiel 9.10
Wir betrachten den Endomorphismus

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 : (x, y)^t \mapsto (5x + 2y, 6x + 3y)^t.$$

Ist E die Standardbasis des $\mathbb{R}^2$, so gilt

$$M_E^E(f) = \begin{pmatrix} 5 & 2 \\ 6 & 3 \end{pmatrix},$$

und mithin gilt

$$\chi_f = \chi_{M_E^E(f)} = \det\begin{pmatrix} t-5 & -2 \\ -6 & t-3 \end{pmatrix} = t^2 - 8t + 3.$$

Alternativ könnte man die Basis $B = \left((1,1)^t, (0,1)^t\right)$ betrachten und erhält dann

$$M_B^B(f) = \begin{pmatrix} 7 & 2 \\ 2 & 1 \end{pmatrix},$$

und in der Tat gilt auch

$$\chi_f = \chi_{M_B^B(f)} = \det\begin{pmatrix} t-7 & -2 \\ -2 & t-1 \end{pmatrix} = (t-7)\cdot(t-1) - 4 = t^2 - 8t + 3.$$

Das charakteristische Polynom verträgt sich gut mit f-invarianten Unterräumen.

Proposition 9.11
Es sei $1 \leq \dim_K(V) < \infty$, $f \in \mathrm{End}_K(V)$.

a. *Ist* $U \subseteq V$ *ein* f*-invarianter Unterraum, dann gilt*

$$\chi_f = \chi_{f_U} \cdot \chi_{f_{V/U}}.$$

b. *Ist* $V = U_1 \oplus \ldots \oplus U_k$*, wobei die* U_i f*-invariant seien, dann gilt*

$$\chi_f = \chi_{f_{U_1}} \cdot \ldots \cdot \chi_{f_{U_k}}.$$

Beweis:

a. Wir wählen eine Basis $B' = (x_1, \ldots, x_k)$ von U und ergänzen diese zu einer Basis $B = (x_1, \ldots, x_n)$ von V. Dann ist $B'' = (\overline{x_{k+1}}, \ldots, \overline{x_n})$ eine Basis von V/U und es gilt

$$M_B^B(f) = \left(\begin{array}{c|c} M_{B'}^{B'}(f_U) & * \\ \hline 0 & M_{B''}^{B''}(f_{V/U}) \end{array}\right)$$

nach Proposition 5.19. Für das charakteristische Polynom erhalten wir mit Hilfe des Kästchensatzes 8.24 dann

$$\chi_f = \chi_{M_B^B(f)} = \det\left(\begin{array}{c|c} \mathbb{1}_k - M_{B'}^{B'}(f_U) & ** \\ \hline 0 & \mathbb{1}_{n-k} - M_{B''}^{B''}(f_{V/U}) \end{array}\right)$$

$$= \det\left(\mathbb{1}_k - M_{B'}^{B'}(f_U)\right) \cdot \det\left(\mathbb{1}_{n-k} - M_{B''}^{B''}(f_{V/U})\right) = \chi_{f_U} \cdot \chi_{f_{V/U}}.$$

b. Die Aussage folgt analog zu Teil a. aus Proposition 5.20.

□

Bemerkung 9.12 (Normalformen bezüglich Konjugation als Ziel)
Der Rest der Vorlesung ist folgender Aufgabe gewidmet:

Finde eine Basis B so, daß $M_B^B(f)$ eine besonders einfache Gestalt hat und wichtige Eigenschaften von f direkt aus $M_B^B(f)$ ersichtlich sind!

Alternativ kann man die Frage auch für quadratische Matrizen formulieren:

Finde ein invertierbares $T \in \mathrm{Gl}_n(K)$ so, daß $T^{-1} \circ A \circ T$ eine besonders einfache Gestalt hat und wichtige Eigenschaften von A sofort sichtbar sind!

Solche *einfachen Repräsentanten* der Äquivalenzklassen bezüglich Konjugation nennt man dann *Normalformen bezüglich Konjugation.*

Ich möchte an dieser Stelle daran erinnern, daß wir uns schon mal eine ähnliche Aufgabe gestellt haben. Wir wollten Basen B und D finden, so daß die Matrixdarstellung $M_D^B(f)$ möglichst einfache Gestalt hat, oder alternativ invertierbare Matrizen S und T, so daß $S \circ A \circ T$ möglichst einfach ist. Die Aufgabe haben wir in Satz 5.31 und Korollar 5.32 gelöst und festgestellt, daß wir stets eine Matrix der Form

$$\begin{pmatrix} \mathbb{1}_r & 0 \\ 0 & 0 \end{pmatrix}$$

erhalten können, wobei r der Rang von f bzw. von A ist. Aus dieser Form kann man über die Abbildung bzw. die Matrix außer dem Rang keine interessante Information mehr ablesen. Das ist der Grund, weshalb es wichtig ist, daß wir uns von nun an auf die Situation $B = D$ bei Matrixdarstellungen bzw. $S = T^{-1}$ bei Matrizen beschränken! Und wir haben oben schon gesehen, daß bei solchen Transformationen interessante Eigenschaften wie die Determinante, die Spur und das charakteristische Polynom erhalten bleiben.

B) Eigenwerte

Der Begriff des Eigenwertes ist von zentraler Bedeutung für die in Bemerkung 9.12 angestrebte Klassifikation.

Definition 9.13 (Eigenwerte und Eigenvektoren)
Sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

a. $\lambda \in K$ heißt *Eigenwert* von f, falls es ein $0 \neq x \in V$ mit $f(x) = \lambda x$ gibt.

   Der Vektor x heißt dann *Eigenvektor* zum Eigenwert λ von f.

   $\mathrm{Eig}(f, \lambda) := \{x \in V \mid f(x) = \lambda x\}$ heißt der *Eigenraum* von f zum Eigenwert λ.

   Die Menge $\sigma(f) := \{\lambda \in K \mid \lambda \text{ ist Eigenwert von } f\}$ der Eigenwerte von f heißt das *Spektrum* von f.

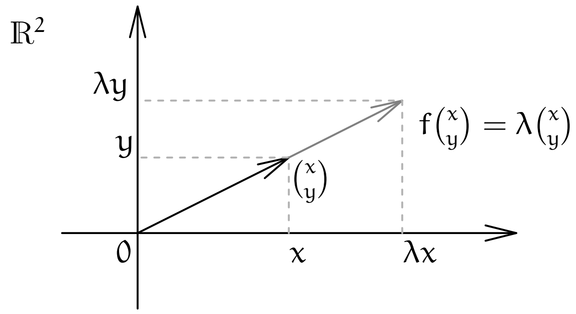
b. $\lambda \in K$ heißt *Eigenwert* von A, falls es ein $0 \neq x \in K^n$ mit $Ax = \lambda x$ gibt.

   Der Vektor x heißt dann *Eigenvektor* zum Eigenwert λ von A.

   $\mathrm{Eig}(A, \lambda) := \{x \in V \mid Ax = \lambda x\}$ heißt der *Eigenraum* von A zum Eigenwert λ.

   Die Menge $\sigma(A) := \{\lambda \in K \mid \lambda \text{ ist Eigenwert von } A\}$ der Eigenwerte von A heißt das *Spektrum* von A.

Bemerkung 9.14 (Geometrische Interpretation von Eigenvektoren)
Ist λ Eigenwert von f mit Eigenvektor x, so bedeutet das anschaulich, daß f in *Richtung* von x durch Multiplikation mit λ wirkt. Diese Anschauung liefert im Fall $V = \mathbb{R}^n$ und $\lambda > 0$, daß f den Vektor x um den Faktor λ streckt, falls $\lambda > 1$, und um den Faktor λ staucht, falls $0 < \lambda < 1$.



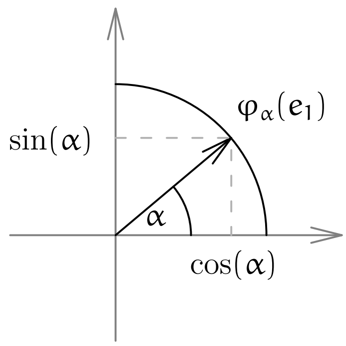
Beispiel 9.15

a. Ist $\dim_K(V) = 1$, so ist jeder Vektor ungleich Null ein Eigenvektor von f, da f schlicht die Multiplikation mit einer Konstanten ist.

b. Ist $\dim_K(V) \geq 2$, so braucht f hingegen keine Eigenwerte und Eigenvektoren zu besitzen. Dabei hängt die Frage der Existenz wesentlich vom Grundkörper K ab. Betrachte etwa die Drehung $\varphi_\alpha = f_{A_\alpha} : \mathbb{R}^2 \to \mathbb{R}^2$ um den Winkel

$\alpha \in \mathbb{R}$ aus Beispiel 2.24. Die Matrixdarstellung bezüglich der kanonischen Basis $E = (e_1, e_2)$ ist

$$A_\alpha = M_E^E(\varphi_\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$



Aus einer rein geometrischen Betrachtung folgt unmittelbar, daß φ_α bzw. A_α nur dann einen Eigenvektor besitzen können, wenn α ein ganzzahliges Vielfaches von π ist.

Bemerkung 9.16 (Eigenräume)
Es seien $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

a. Da $f(x) = \lambda x$ für $x \in V$ und $\lambda \in K$ genau dann erfüllt ist, wenn x im Kern der linearen Abbildung $f - \lambda \,\mathrm{id}_V \in \mathrm{End}_K(V)$ liegt, gilt also

$$\mathrm{Eig}(f, \lambda) = \mathrm{Ker}(f - \lambda \,\mathrm{id}_V) = \mathrm{Ker}(\lambda \,\mathrm{id}_V - f).$$

Analog erhält man:

$$\mathrm{Eig}(A, \lambda) = \mathrm{Ker}(f_A - \lambda \,\mathrm{id}_V) = \mathrm{Lös}(A - \lambda \mathbb{1}_n, 0) = \mathrm{Lös}(\lambda \mathbb{1}_n - A, 0).$$

b. Aus der Definition folgt unmittelbar, daß $\sigma(A) = \sigma(f_A)$ und $\sigma(f) = \sigma\big(M_B^B(f)\big)$.

c. Ebenso folgt unmittelbar, daß der Eigenraum $\mathrm{Eig}(f, \lambda)$ von f zum Eigenwert λ f-invariant ist.

d. Kennt man einen Eigenwert $\lambda \in K$ von A, so kann man das lineare Gleichungssystem

$$(A - \lambda \mathbb{1}_n)x = 0 \quad \text{oder} \quad (\lambda \mathbb{1}_n - A)x = 0$$

lösen und damit eine Basis des Eigenraumes $\mathrm{Eig}(A, \lambda) = \mathrm{Lös}(A - \lambda \mathbb{1}_n, 0)$ bestimmen. D. h., bei Kenntnis des Eigenwertes λ lassen sich die Eigenvektoren von A zu λ durch Lösen eines linearen Gleichungssystems bestimmen. Aber wie kommt man zu den Eigenwerten von A?

Satz 9.17 (Eigenwerte und das charakteristische Polynom)
Es sei $f \in \mathrm{End}_K(V)$ *und* $A \in \mathrm{Mat}_n(K)$.

a. *Die Eigenwerte von* f *sind genau die Nullstellen von* χ_f *in* K.

b. *Die Eigenwerte von* A *sind genau die Nullstellen von* χ_A *in* K.

Insbesondere, f *und* A *haben höchstens* n *paarweise verschiedene Eigenwerte.*

Beweis: Für $\lambda \in K$ gilt unter Berücksichtigung von Korollar 4.22:

$$\begin{aligned}\lambda \text{ ist Eigenwert von } f &\Longleftrightarrow \operatorname{Ker}(\lambda \operatorname{id}_V - f) = \operatorname{Eig}(f, \lambda) \neq \{0\}\\ &\Longleftrightarrow \lambda \operatorname{id}_V - f \text{ ist nicht injektiv}\\ &\overset{4.22}{\Longleftrightarrow} \lambda \operatorname{id}_V - f \text{ ist nicht bijektiv}\\ &\Longleftrightarrow \chi_f(\lambda) = \det(\lambda \operatorname{id}_V - f) = 0.\end{aligned}$$

Der Beweis für die Matrizen geht analog. □

Bevor wir das charakteristische Polynom weiter untersuchen, wollen wir zunächst einige Beispiele betrachten.

Beispiel 9.18

a. Betrachten wir zunächst die folgende Matrix:

$$A = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix} \in \operatorname{Mat}(3, \mathbb{Q}).$$

Mit Hilfe der Regel von Sarrus oder durch den Laplaceschen Entwicklungssatz bestimmen wir das charakteristische Polynom von A als

$$\chi_A = \det \begin{pmatrix} t & -1 & -1 \\ 1 & t-2 & -1 \\ 1 & -1 & t-2 \end{pmatrix} = t^3 - 4t^2 + 5t - 2 = (t-1)^2 \cdot (t-2).$$

Alternativ kann man, da $\mathbb{Q}(t)$ ein Körper ist, die Determinante mittels des Gaußschen Algorithmus' 8.15 bestimmen. Insbesondere dürfen wir dabei durch Polynome dividieren!

$$\begin{aligned}&\begin{pmatrix} t & -1 & -1 \\ 1 & t-2 & -1 \\ 1 & -1 & t-2 \end{pmatrix} \xmapsto[\text{III}\mapsto\text{III}-\frac{1}{t}\text{I}]{\text{II}\mapsto\text{II}-\frac{1}{t}\text{I}} \begin{pmatrix} t & -1 & -1 \\ 0 & t-2+\frac{1}{t} & \frac{1}{t}-1 \\ 0 & \frac{1}{t}-1 & t-2+\frac{1}{t} \end{pmatrix} = \\ &\begin{pmatrix} t & -1 & -1 \\ 0 & \frac{(t-1)^2}{t} & -\frac{t-1}{t} \\ 0 & -\frac{t-1}{t} & \frac{(t-1)^2}{t} \end{pmatrix} \xmapsto{\text{III}\mapsto\text{III}+\frac{1}{t-1}\text{II}} \begin{pmatrix} t & -1 & -1 \\ 0 & \frac{(t-1)^2}{t} & -\frac{t-1}{t} \\ 0 & 0 & t-2 \end{pmatrix}.\end{aligned} \tag{26}$$

Entsprechend erhalten wir für das charakteristische Polynom

$$\chi_A = t \cdot \frac{(t-1)^2}{t} \cdot (t-2) = (t-1)^2 \cdot (t-2).$$

Das charakteristische Polynom hat also die Nullstellen $\lambda = 1$ und $\lambda = 2$, wobei $\lambda = 1$ eine zweifache Nullstelle ist. Insbesondere ist also $\sigma(A) = \{1, 2\}$.

Wir können jetzt für $\lambda = 1$ und für $\lambda = 2$ jeweils den Eigenraum $\mathrm{Eig}(A, \lambda) = \mathrm{Lös}(\lambda\mathbb{1}_n - A, 0)$ mit Hilfe des Gauß-Algorithmus bestimmen.[1]

Der Algorithmus zur Bestimmung von $\mathrm{Eig}(A, 1) = \mathrm{Lös}(\mathbb{1}_n - A, 0)$ sieht vor, daß wir die Matrix zunächst auf reduzierte ZSF bringen und dann in den Nullzeilen die Diagonalelemente durch -1 ersetzen:

$$\begin{pmatrix} -1 & 1 & 1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & -1 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Die letzten beiden Spalten, d.h. die, bei denen eine -1 auf der Diagonalen steht, bilden dann eine Basis des Eigenraumes zum Eigenwert 1:

$$\mathrm{Eig}(A, 1) = \mathrm{Lin}\left((-1, -1, 0)^t, (-1, 0, -1)^t\right).$$

$\mathrm{Eig}(A, 1)$ ist also zweidimensional.

Analog ergibt sich $\mathrm{Eig}(A, 2)$ aus

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

und damit gilt $\mathrm{Eig}(A, 2) = \mathrm{Lin}\left((-1, -1, -1)^t\right)$.

b. Wir hatten schon durch eine geometrische Argumentation gesehen, daß die Drehung um einen Winkel α im allgemeinen keinen reellen Eigenwert besitzt. Den gleichen Sachverhalt prüfen wir nun noch einmal mit algebraischen Methoden. Die Matrixdarstellung der Drehung bezüglich der kanonischen Basis von $\mathbb{R}^2$ ist

$$A_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

Aber

$$\chi_{A_\alpha} = \left(t - \cos(\alpha)\right)^2 + \sin^2(\alpha) = t^2 - 2\cos(\alpha)t + 1.$$

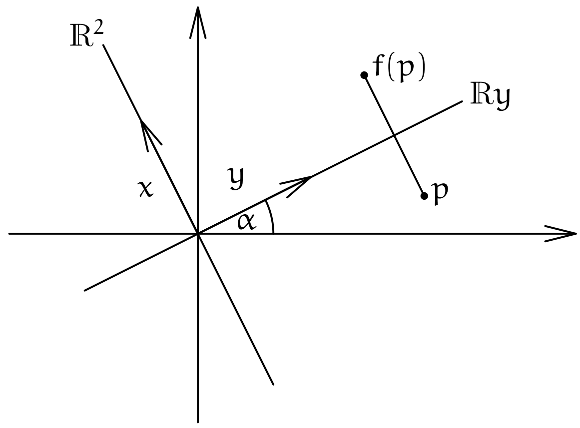
Die Nullstellen von χ_{A_α} sind $\cos(\alpha) + \sqrt{\cos^2(\alpha) - 1}$ und $\cos(\alpha) - \sqrt{\cos^2(\alpha) - 1}$. Für beide Terme gilt, sie sind genau dann reell, wenn α ein ganzzahliges Vielfaches von π ist.

Insbesondere hat die Drehung also nur dann reelle Eigenwerte, wenn α ein ganzzahliges Vielfaches von π ist, d. h. $A_\alpha = \mathbb{1}_2$ oder $A_\alpha = -\mathbb{1}_2$.

c. Es sei $f \in \mathrm{End}_{\mathbb{R}}\left(\mathbb{R}^2\right)$ die *Spiegelung* an einer Geraden $\mathrm{Lin}\,(y) = \mathbb{R} \cdot y \subset \mathbb{R}^2$ mit $0 \neq y = (y_1, y_2)^t \in \mathbb{R}^2$.

---

[1] Man beachte, daß es zur Berechnung der reduzierten Zeilen-Stufen-Form von $\lambda\mathbb{1}_n - A$ für $\lambda = 1$ *nicht* erlaubt ist, in (26) in der letzten Matrix t etwa durch $\lambda = 1$ zu ersetzen, um die ZSF zu erhalten, da wir bei den vorgenommenen Umformungen zur Ermittelung obiger Matrix durch das Polynom $t - 1$ dividiert haben. Dies ist über $\mathbb{Q}(t)$ eine erlaubte Operation gewesen. Ersetzen wir jedoch t durch 1, so ist die Operation nicht mehr erlaubt!



Wir setzen $x = (-y_2, y_1)^t \in \mathbb{R}^2$. Dann steht x senkrecht auf y und $B = (y, x)$ ist eine Basis von $\mathbb{R}^2$. Die Spiegelung f bildet mithin y auf sich selbst und x auf $-x$ ab, da x senkrecht auf $\mathrm{Lin}\,(y)$ steht. Damit hat f die folgende Matrixdarstellung bezüglich B

$$M_B^B(f) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

und das charakteristische Polynom von f ist gerade

$$\chi_f = (t-1)\cdot(t+1).$$

Die Spiegelung von f hat also Spektrum $\sigma(f) = \{-1, 1\}$.
Beschreiben wir f in den Standardkoordinaten $E = (e_1, e_2)$ von $\mathbb{R}^2$, so ist f die Spiegelung an $\mathrm{Lin}\,(e_1) = \mathbb{R}\cdot e_1$ gefolgt von der Drehung um den Winkel 2α, wenn α der Winkel ist, den $\mathrm{Lin}\,(y)$ mit $\mathrm{Lin}\,(e_1)$ einschließt. Wir erhalten also

$$M_E^E(f) = \begin{pmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{pmatrix} = \begin{pmatrix} \cos(2\alpha) & -\sin(2\alpha) \\ \sin(2\alpha) & \cos(2\alpha) \end{pmatrix} \circ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Das charakteristische Polynom errechnet sich aus dieser Matrixdarstellung als

$$\big(t-\cos(2\alpha)\big)\cdot\big(t+\cos(2\alpha)\big) - \sin^2(2\alpha) = t^2 - 1 = (t-1)\cdot(t+1).$$

Korollar 9.19 (Eigenwerte einer Dreiecksmatrix)
Ist $A = (a_{ij}) \in \mathrm{Mat}_n(K)$ eine obere oder untere Dreiecksmatrix, dann ist

$$\chi_A = (t-a_{11})\cdot\ldots\cdot(t-a_{nn})$$

und die Einträge auf der Diagonalen sind genau die Eigenwerte von A.

Beweis: Für eine obere Dreiecksmatrix $A = (a_{ij})$ gilt

$$\chi_A = \det\begin{pmatrix} t-a_{11} & * & \ldots & \ldots & * \\ 0 & t-a_{22} & * & \ldots & * \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \ldots & \ldots & 0 & t-a_{nn} \end{pmatrix} = \prod_{i=1}^n (t-a_{ii}),$$

und dieses Polynom hat genau die Nullstellen $a_{11}, \ldots, a_{nn}$. Der Beweis für untere Dreiecksmatrizen geht analog. □

Für die Definition der Vielfachheit $\operatorname{mult}(p, \lambda)$ einer Zahl λ als Nullstelle eines Polynoms p verweisen wir auf Definition B7.13.

Definition 9.20 (Vielfachheit von Eigenwerten)
Es sei $f \in \operatorname{End}_K(V)$, $A \in \operatorname{Mat}_n(K)$ und $\lambda \in K$.

a. $\operatorname{mult}(\chi_f, \lambda)$ heißt *algebraische Vielfachheit* von λ als Eigenwert von f.
   $\dim_K \operatorname{Eig}(f, \lambda)$ heißt *geometrische Vielfachheit* von λ als Eigenwert von f.
b. $\operatorname{mult}(\chi_A, \lambda)$ heißt *algebraische Vielfachheit* von λ als Eigenwert von A.
   $\dim_K \operatorname{Eig}(A, \lambda)$ heißt *geometrische Vielfachheit* von λ als Eigenwert von A.

Die algebraischen Vielfachheiten nennt man auch *arithmetische Vielfachheiten.*

Beispiel 9.21
Die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \operatorname{Mat}_2(\mathbb{C})$$

aus Beispiel 10.4 hat nur den Eigenwert 0, da $\chi_A = t^2$. Die algebraische Vielfachheit von 0 als Eigenwert von A ist

$$\operatorname{mult}(\chi_A, 0) = \operatorname{mult}(t^2, 0) = 2,$$

während die geometrische Vielfachheit

$$\dim_{\mathbb{C}} \operatorname{Eig}(A, 0) = \dim_{\mathbb{C}} \operatorname{Lös}(A, 0) = \dim_{\mathbb{C}} \operatorname{Lin}\big((1, 0)^t\big) = 1$$

ist.

Lemma 9.22 (Geometrische und algebraische Vielfachheit)
Es sei $f \in \operatorname{End}_K(V)$*,* $A \in \operatorname{Mat}_n(K)$ *und* $\lambda \in K$*. Dann gilt stets*

$$\dim_K \operatorname{Eig}(f, \lambda) \leq \operatorname{mult}(\chi_f, \lambda)$$

und

$$\dim_K \operatorname{Eig}(A, \lambda) \leq \operatorname{mult}(\chi_A, \lambda),$$

d.h. die geometrische Vielfachheit eines Eigenwertes ist stets nach oben durch die algebraische Vielfachheit beschränkt.

Beweis: Man beachte, daß $U := \operatorname{Eig}(f, \lambda)$ ein f-invarianter Unterraum ist und daß $f_U = \lambda \cdot \operatorname{id}_U$ gilt. Mithin ist

$$\chi_{f_U} = \chi_{\lambda \cdot \operatorname{id}_U} = \det(t \cdot \operatorname{id}_U - \lambda \cdot \operatorname{id}_U) = \det\big((t - \lambda) \cdot \operatorname{id}_U\big) = (t - \lambda)^s$$

wobei $s = \dim_K(U) = \dim_K \operatorname{Eig}(f, \lambda)$. Außerdem gilt nach Proposition 9.11

$$\chi_f = \chi_{f_U} \cdot \chi_{f_{V/U}} = (t - \lambda)^s \cdot \chi_{f_{V/U}}.$$

Daraus folgt unmittelbar

$$\operatorname{mult}(f, \lambda) \geq s = \dim_K \operatorname{Eig}(f, \lambda).$$

Die analoge Aussage für A folgt hieraus mit $f = f_A$. $\square$

Lemma 9.23 (Eigenwerte bei konjugierten Matrizen)
Für $A, B \in \operatorname{Mat}_n(K)$ *und* $T \in \operatorname{Gl}_n(K)$ *mit* $B = T^{-1} \circ A \circ T$ *sowie* $\lambda \in K$ *gelten:*

a. $\sigma(A) = \sigma(B)$.

b. $\operatorname{mult}(\chi_A, \lambda) = \operatorname{mult}(\chi_B, \lambda)$.

c. $\dim_K \operatorname{Eig}(A, \lambda) = \dim_K \operatorname{Eig}(B, \lambda)$.

d. $x \in \operatorname{Eig}(A, \lambda) \iff T^{-1}x \in \operatorname{Eig}(B, \lambda)$.

D.h. konjugierte Matrizen haben die gleichen Eigenwerte und für jeden Eigenwert stimmen ihre geometrischen Vielfachheiten ebenso überein wie ihre algebraischen Vielfachheiten.

Beweis: Nach Proposition 9.8 haben A und B die gleichen charakteristischen Polynome. Mithin stimmen wegen Satz 9.17 die Eigenwerte von A und B sowie deren algebraische Vielfachheiten überein. Damit sind a. und b. gezeigt. Ferner gilt

$$\begin{aligned} x \in \operatorname{Eig}(A, \lambda) &\iff \lambda x = Ax = ATT^{-1}x \\ &\iff \lambda T^{-1}x = T^{-1}ATT^{-1}x = BT^{-1}x \iff T^{-1}x \in \operatorname{Eig}(B, \lambda). \end{aligned}$$

Damit ist d. gezeigt und außerdem folgt, daß der Isomorphismus $f_{T^{-1}}$ den Eigenraum $\operatorname{Eig}(A, \lambda)$ isomorph auf den Eigenraum $\operatorname{Eig}(B, \lambda)$ abbildet. Die beiden müssen also die gleiche Dimension haben, womit auch c. gezeigt ist. $\square$

Aufgaben

Aufgabe 9.24 (Nilpotente Endomorphismen und Matrizen)
Es sei $A \in \operatorname{Mat}_n(K)$ und $f \in \operatorname{End}_K(V)$ mit $1 \leq \dim_K(V) < \infty$.

a. Zeige, gibt es ein $r \in \mathbb{N}$ mit $f^r = 0$, so gilt $\operatorname{Spur}(f) = 0$.

b. Zeige, gibt es ein $r \in \mathbb{N}$ mit $A^r = 0$, so gilt $\operatorname{Spur}(A) = 0$.

c. Finde ein Beispiel für eine Matrix wie in Teil b., bei der nicht alle Diagonalelemente Null sind.

Hinweis zum Beweis von a.: Führe Induktion über $n = \dim_K(V)$. Dazu zeige man, daß $M_B^B(f)$ für eine geeignete Wahl von B Blockgestalt mit einem Nullblock in der oberen linken Ecke hat. Proposition 5.19b. mit $U = \operatorname{Ker}(f)$ ist dabei hilfreich.

Aufgabe 9.25
Es sei $1 \leq \dim_K(V) < \infty$, $f \in \operatorname{End}_K(V)$.

a. Ist $U \subseteq V$ ein f-invarianter Unterraum, dann gilt

$$\det(f) = \det(f_U) \cdot \det(f_{V/U}).$$

b. Ist $V = U_1 \oplus \ldots \oplus U_k$, wobei die U_i f-invariant seien, dann gilt

$$\det(f) = \det(f_{U_1}) \cdot \ldots \cdot \det(f_{U_k}).$$

Aufgabe 9.26

Für ein Polynom $p \in K[t]$ und zwei konjugierte Matrizen $A, B \in \mathrm{Mat}_n(K)$ gilt

$$p(A) = 0 \quad \Longleftrightarrow \quad p(B) = 0.$$

Aufgabe 9.27

Bestimme die Eigenwerte und die Eigenräume der folgenden Matrix A:

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 3 & 0 & 0 \\ -1 & 1 & 2 & 1 \\ -1 & 1 & 0 & 3 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Q}).$$

Aufgabe 9.28 (Die Eigenräume bilden eine direkte Summe.)

Es sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

a. Sind $x_1, \ldots, x_r \in V$ Eigenvektoren von f zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_r \in K$, dann ist die Familie $(x_1, \ldots, x_r)$ linear unabhängig. Insbesondere gilt

$$\mathrm{Eig}(f, \lambda_1) + \ldots + \mathrm{Eig}(f, \lambda_r) = \mathrm{Eig}(f, \lambda_1) \oplus \ldots \oplus \mathrm{Eig}(f, \lambda_r).$$

b. Sind $x_1, \ldots, x_r \in K^n$ Eigenvektoren von A zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_r \in K$, dann ist die Familie $(x_1, \ldots, x_r)$ linear unabhängig. Insbesondere gilt

$$\mathrm{Eig}(A, \lambda_1) + \ldots + \mathrm{Eig}(A, \lambda_r) = \mathrm{Eig}(A, \lambda_1) \oplus \ldots \oplus \mathrm{Eig}(A, \lambda_r).$$

Aufgabe 9.29

Sind $f, g \in \mathrm{End}_K(V)$, so gilt $\sigma(f \circ g) = \sigma(g \circ f)$.

§ 10 Diagonalisierbarkeit und Trigonalisierbarkeit

In diesem Abschnitt sei V ein K-Vektorraum mit $1 \leq \dim_K(V) = n < \infty$.

In Bemerkung 9.12 haben wir erläutert, daß unser zentrales Anliegen darin besteht, eine Basis B bzw. eine invertierbare Matrix $T \in \mathrm{Gl}_n(K)$ zu finden, so daß $M_B^B(f)$ bzw. $T^{-1} \circ A \circ T$ eine möglichst einfache Gestalt hat. Dazu zählt sicher, daß die Matrix möglichst viele Nullen enthält.

Definition 10.1 (Diagonalisierbar und trigonalisierbar)
Es sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

a. f heißt *diagonalisierbar* (bzw. *trigonalisierbar*), falls es eine Basis B von V gibt, so daß $M_B^B(f)$ eine Diagonalmatrix (bzw. eine obere Dreiecksmatrix) ist.
b. A heißt *diagonalisierbar* (bzw. *trigonalisierbar*), falls es eine Matrix $T \in \mathrm{Gl}_n(K)$ gibt, so daß $T^{-1} \circ A \circ T$ eine Diagonalmatrix (bzw. eine obere Dreiecksmatrix) ist.

A) Trigonalisierbarkeit

Satz 10.2 (Trigonalisierbarkeit)
Es sei $f \in \mathrm{End}_K(V)$ *und* $A \in \mathrm{Mat}_n(K)$.

a. *Genau dann ist* f *trigonalisierbar, wenn* χ_f *über* K *in Linearfaktoren zerfällt.*
b. *Genau dann ist* A *trigonalisierbar, wenn* χ_A *über* K *in Linearfaktoren zerfällt.*

Beweis: Ist f trigonalisierbar, so gibt es eine Basis B mit

$$M_B^B(f) = \begin{pmatrix} \lambda_1 & * & \ldots & \ldots & * \\ 0 & \lambda_2 & * & \ldots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda_{n-1} & * \\ 0 & \ldots & \ldots & 0 & \lambda_n \end{pmatrix}.$$

Damit folgt, daß das charakteristische Polynom

$$\chi_f = (t - \lambda_1) \cdots (t - \lambda_n)$$

von f über K in Linearfaktoren zerfällt.

Zerfalle nun umgekehrt das charakteristische Polynom von f in Linearfaktoren $\chi_f = (t - \lambda_1) \cdots (t - \lambda_n)$. Wir beweisen mit Induktion über $n = \dim_K(V)$, daß dann f trigonalisierbar ist. Im Fall $n = 1$ ist f nach Beispiel 9.15 sogar diagonalisierbar.

Sei also $n > 1$ und sei $0 \neq x_1 \in V$ ein Eigenvektor von f zum Eigenwert λ_1. Wir setzen $U := \mathrm{Lin}(x_1) \leq V$. Wegen $f(x_1) = \lambda_1 x_1 \in U$ ist U ein f-invarianter Unterraum von V und $\chi_{f_U} = t - \lambda_1$. Mithin folgt aus Proposition 9.11

$$\chi_{f_{V/U}} = (t - \lambda_2) \cdots (t - \lambda_n),$$

d. h. das charakteristische Polynom von $f_{V/U}$ zerfällt über K in Linearfaktoren. Da $\dim_K(V/U) = n - 1 < n$, existiert per Induktion eine Basis $B'' = \left(\overline{x_2}, \ldots, \overline{x_n}\right)$ von V/U, so daß $M_{B''}^{B''}\left(f_{V/U}\right)$ eine obere Dreiecksmatrix ist. Dann ist aber $B = (x_1, \ldots, x_n)$ eine Basis von V und mit $B' = (x_1)$ gilt wegen Proposition 5.19

$$M_B^B(f) = \left(\begin{array}{c|c} M_{B'}^{B'}(f_U) & * \\ \hline 0 & M_{B''}^{B''}(f_{V/U}) \end{array}\right) = \left(\begin{array}{c|c} \lambda_1 & * \\ \hline 0 & M_{B''}^{B''}(f_{V/U}) \end{array}\right).$$

Damit ist $M_B^B(f)$ eine obere Dreiecksmatrix und f ist trigonalisierbar.

Die Aussage für A erhalten wir aus der entsprechenden Aussage für f_A. □

Bemerkung 10.3
Ist K ein algebraisch abgeschlossener Körper, etwa $K = \mathbb{C}$, so sind somit jede Matrix A und jeder Endomorphismus f trigonalisierbar. Eine vergleichbare Aussage für die Diagonalisierbarkeit gilt nicht.

Beispiel 10.4

a. Die Drehmatrix

$$A_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

hat das charakteristische Polynom $\chi_{A_\alpha} = t^2 - 2\cos(\alpha)t + 1 = (t - \lambda) \cdot \left(t - \overline{\lambda}\right)$ mit $\lambda = \cos(\alpha) + i\sin(\alpha) \in \mathbb{C}$, $\alpha \in \mathbb{R}$. Damit hat χ_{A_α} also keine reellen Nullstellen, wenn α kein ganzzahliges Vielfaches von π ist, und somit ist A_α über $\mathbb{R}$ nicht trigonalisierbar.
Hingegen zerfällt χ_{A_α} über $\mathbb{C}$ in Linearfaktoren, so daß A_α über $\mathbb{C}$ trigonalisierbar sein muß. In der Tat ist A_α sogar diagonalisierbar mit Diagonalgestalt

$$\begin{pmatrix} \lambda & 0 \\ 0 & \overline{\lambda} \end{pmatrix}.$$

Ist α kein ganzzahliges Vielfaches von π, so besitzt A_α zwei verschiedene Eigenwerte, so daß zugehörige Eigenvektoren nach Aufgabe 9.28 eine Basis von $\mathbb{C}^2$ bilden müssen und diese transformieren A_α in obige Diagonalmatrix. Ist α hingegen ein ganzzahliges Vielfaches von π, so ist $A_\alpha = \mathbb{1}_2$ oder $A_\alpha = -\mathbb{1}_2$ und hat bereits Diagonalgestalt.

b. Die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{C})$$

ist hingegen auch über $\mathbb{C}$ nicht diagonalisierbar. Denn, gäbe es eine Matrix $T \in \mathrm{Gl}_2(\mathbb{C})$ mit

$$T^{-1} \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{C}),$$

dann wäre

$$\begin{pmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_2^2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}^2 = T^{-1} \circ A^2 \circ T = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

also wären $\lambda_1 = \lambda_2 = 0$. Aber damit würde gelten:

$$0 = \mathrm{rang} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \mathrm{rang}\left(T^{-1} \circ A \circ T\right) = \mathrm{rang}(A) = 1,$$

da $T \in \mathrm{Gl}_2(\mathbb{C})$. Dies ist jedoch ein Widerspruch.

B) Diagonalblockmatrizen

Definition 10.5 (Diagonalblockmatrizen)
Wir werden im Folgenden sehr häufig mit Blockmatrizen der folgenden Form arbeiten:

$$A = \left(\begin{array}{c|c|ccc} A_1 & 0 & \cdots & & 0 \\ \hline 0 & \boxed{A_2} & & & \vdots \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & & \cdots & \boxed{A_r} \end{array}\right) \in \mathrm{Mat}_n(K),$$

wobei $A_i \in \mathrm{Mat}_{n_i}(K)$, $i = 1, \ldots, r$ mit $n = n_1 + \ldots + n_r$. Es empfiehlt sich deshalb, eine Kurzschreibweise für solche *Diagonalblockmatrizen* einzuführen. Wir schreiben kurz:

$$A = A_1 \oplus \ldots \oplus A_r = \bigoplus_{i=1}^r A_i.$$

Bemerkung 10.6 (Diagonalblockmatrizen)

a. Man beachte, daß es bei der obigen Schreibweise für Diagonalblockmatrizen auf die Reihenfolge der Summation ankommt, daß aber Matrizen, die durch Änderung der Summationsreihenfolge entstehen, zueinander konjugiert sind!

b. Mit Hilfe dieser Notation gilt beispielsweise, daß eine Matrix A genau dann diagonalisierbar ist, wenn es Körperelemente $\lambda_1, \ldots, \lambda_r \in K$ und positive natürliche Zahlen $n_1, \ldots, n_r \in \mathbb{N}$ gibt sowie eine invertierbare Matrix $T \in \mathrm{Gl}_n(K)$ mit

$$T^{-1} \circ A \circ T = \bigoplus_{i=1}^r \lambda_i \mathbb{1}_{n_i}.$$

c. Ist $A = \bigoplus_{i=1}^r A_i$ eine Diagonalblockmatrix, so verifiziert man leicht, daß für $k \in \mathbb{N}$ gilt $A^k = \bigoplus_{i=1}^r A_i^k$, und damit, daß für ein Polynom $p \in K[t]$ gilt

$$p(A) = \bigoplus_{i=1}^r p(A_i).$$

Insbesondere gilt also für eine Diagonalmatrix $D = \bigoplus_{i=1}^n \lambda_i \mathbb{1}_1$, daß

$$p(D) = \bigoplus_{i=1}^n p(\lambda_i)\mathbb{1}_1 = \begin{pmatrix} p(\lambda_1) & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & p(\lambda_n) \end{pmatrix}$$

In der Tat kann man sogar zeigen, daß für eine Blockmatrix der Form

$$A = \left(\begin{array}{c|ccc} A_1 & * & \cdots & * \\ \hline 0 & \boxed{A_2} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \boxed{A_r} \end{array}\right) \in \mathrm{Mat}_n(K),$$

gilt, daß

$$p(A) = \left(\begin{array}{c|ccc} p(A_1) & * & \cdots & * \\ \hline 0 & \boxed{p(A_2)} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \boxed{p(A_r)} \end{array}\right) \in \mathrm{Mat}_n(K),$$

wobei sich die Sterne oberhalb der Blöcke verändert haben.

Damit gilt insbesondere, daß $p(A)$ eine obere Dreiecksmatrix ist, falls A eine solche war.

C) Der Satz von Cayley-Hamilton

Da $\dim_K\big(\mathrm{Mat}_n(K)\big) = n^2$ gilt, sind die n^2+1 Matrizen

$$\mathbb{1}_n = A^0, A^1, A^2, \ldots, A^{n^2}$$

in $\mathrm{Mat}_n(K)$ linear abhängig. D. h. es existieren $\lambda_0, \ldots, \lambda_{n^2} \in K$, nicht alle null, mit

$$\lambda_0 A^0 + \lambda_1 A^1 + \ldots + \lambda_{n^2} A^{n^2} = 0 \in \mathrm{Mat}_n(K).$$

Ein einfaches Dimensionsargument zeigt also, es gibt ein Polynom $0 \neq p = \lambda_{n^2} t^{n^2} + \ldots + \lambda_0 \in K[t]$ vom Grad kleiner gleich n^2 mit $p(A) = 0$. Der folgende wichtige Satz von Cayley-Hamilton besagt nun, daß es sogar ein Polynom vom Grad n gibt, das A annulliert.

Satz 10.7 (Cayley-Hamilton)
Für $f \in \mathrm{End}_K(V)$ *und* $A \in \mathrm{Mat}_n(K)$ *gilt* $\chi_f(f) = 0$ *und* $\chi_A(A) = 0$.

Beweis: Da für eine Basis D von V die Abbildung $M_D^D : \mathrm{End}_K(V) \longrightarrow \mathrm{Mat}_n(K)$ ein K-Algebrenhomomorphismus ist, gilt

$$M_D^D\big(\chi_f(f)\big) = \chi_f\big(M_D^D(f)\big).$$

Dann gilt aber $\chi_f(f) = 0$ genau dann, wenn

$$0 = M_D^D\big(\chi_f(f)\big) = \chi_f\big(M_D^D(f)\big) = \chi_{M_D^D(f)}\big(M_D^D(f)\big).$$

Es reicht deshalb, die Aussage für Matrizen zu beweisen.

Betrachte dazu die Matrix

$$B_t := t \cdot \mathbb{1}_n - A \in \mathrm{Mat}_n(K(t))$$

sowie die Adjunkte $B_t^\# = (p_{ij}) \in \mathrm{Mat}_n(K(t))$ von B_t, die auch *Busadjunkte* von A genannt wird. Nach dem Satz über die Adjunkte 8.29 in $\mathrm{Mat}_n(K[t])$ gilt die Adjunktengleichung

$$B_t \circ B_t^\# = (t\mathbb{1}_n - A) \circ (t\mathbb{1}_n - A)^\# = \det(t\mathbb{1}_n - A) \cdot \mathbb{1}_n = \chi_A \cdot \mathbb{1}_n. \qquad (27)$$

Man beachte nun noch, daß die Einträge von $B_t^\#$ Determinanten von gewissen $(n-1) \times (n-1)$-Matrizen von B_t sind, also Polynome vom Grad höchstens $n-1$. Wir können nun $B_t^\#$ auch als Polynom schreiben, dessen Koeffizienten Matrizen sind, und dieses Polynom hat dann höchstens den Grad $n-1$, d. h. es gibt Matrizen $B_0, \ldots, B_{n-1} \in \mathrm{Mat}_n(K)$ mit

$$B_t^\# = B_{n-1}t^{n-1} + \ldots + B_1 t + B_0.$$

Ist $\chi_A = t^n + \alpha_{n-1}t^{n-1} + \ldots + \alpha_0$, so folgt aus der Adjunktengleichung (27)

$$(\mathbb{1}_n t - A) \circ \left(B_{n-1}t^{n-1} + \ldots + B_1 t + B_0\right) \stackrel{!}{=} \mathbb{1}_n t^n + \alpha_{n-1}\mathbb{1}_n t^{n-1} + \ldots + \alpha_0 \mathbb{1}_n \qquad (28)$$

durch Koeffizientenvergleich für die t^i, $i = 0, \ldots, n$:

$$\begin{array}{rcl} B_{n-1} & = & \mathbb{1}_n \\ -AB_{n-1} + B_{n-2} & = & \alpha_{n-1}\mathbb{1}_n \\ -AB_{n-2} + B_{n-3} & = & \alpha_{n-2}\mathbb{1}_n \\ \vdots & & \\ -AB_1 + B_0 & = & \alpha_1 \mathbb{1}_n \\ -AB_0 & = & \alpha_0 \mathbb{1}_n \end{array} \qquad (29)$$

Multipliziert man die i-te Zeile in (29) mit A^{n-i+1} und summiert die beiden Seiten auf, so erhält man die Behauptung:

$$\begin{array}{rcl} A^n B_{n-1} & = & A^n \\ -A^n B_{n-1} + A^{n-1}B_{n-2} & = & \alpha_{n-1}A^{n-1} \\ -A^{n-1}B_{n-2} + A^{n-2}B_{n-3} & = & \alpha_{n-2}A^{n-2} \\ \vdots & & \\ -A^2 B_1 + AB_0 & = & \alpha_1 A \\ -AB_0 & = & \alpha_0 \mathbb{1}_n \\ \hline 0 & = & \chi_A(A). \end{array}$$

□

Bemerkung 10.8

a. Man beachte, daß der folgende *offensichtliche* Beweis für $\chi_A(A) = 0$, nämlich

$$\chi_A(A) = \det(A * \mathbb{1}_n - A) = \det(0) = 0$$

falsch ist, da "$*$" beim Einsetzen von A in $\det(t\mathbb{1}_n - A) \in K[t]$ eben *nicht* die Matrixmultiplikation ist! Man beachte ferner, daß die Gleichung auch schon deshalb keinen Sinn ergeben kann, da $\chi_A(A)$ die Nullmatrix ist, während $\det(0)$ die Null in K ist.

b. Kennt man das charakteristische Polynom $\chi_A = t^n + \alpha_{n-1}t^{n-1} + \ldots + \alpha_0$, so läßt sich daraus mittels (28) und der Rekursionsformel (29) die Busadjunkte

$$(t\mathbb{1}_n - A)^{\#} = B_{n-1}t^{n-1} + \ldots + B_1 t + B_0$$

von A bestimmen. Für die B_{n-k}, $k = 1, \ldots, n$, gilt dabei explizit:

$$B_{n-k} = A^{k-1} + \alpha_{n-1}A^{k-2} + \ldots + \alpha_{n-k+1}A^0,$$

und wegen $\alpha_0 = (-1)^n \cdot \det(A)$ erhalten wir die Adjunkte von A als

$$A^{\#} = (-1)^{n+1} \cdot B_0 = (-1)^{n+1} \cdot \left(A^{n-1} + \alpha_{n-1}A^{n-2} + \ldots + \alpha_1 A^0\right).$$

Diese Formel zur Berechnung der Adjunkten von A ist weit effizienter, als die Determinanten sämtlicher Streichungsmatrizen zu berechnen.

D) Das Minimalpolynom

Satz 10.9 (Das Minimalpolynom)
Es sei $f \in \mathrm{End}_K(V)$ *und* $A \in \mathrm{Mat}_n(K)$.

a. *Es gibt ein eindeutiges normiertes Polynom* $0 \neq \mu_f \in K[t]$*, so daß*

$$\{p \in K[t] \mid p(f) = 0\} = \{\mu_f \cdot q \mid q \in K[t]\}.$$

Insbesondere ist μ_f *das Nicht-Null-Polynom kleinsten Grades mit* $\mu_f(f) = 0$*. Wir nennen* μ_f *das* Minimalpolynom *von* f*.*

b. *Es gibt ein eindeutiges normiertes Polynom* $0 \neq \mu_A \in K[t]$*, so daß*

$$\{p \in K[t] \mid p(A) = 0\} = \{\mu_A \cdot q \mid q \in K[t]\}.$$

Insbesondere ist μ_A *das Nicht-Null-Polynom kleinsten Grades mit* $\mu_A(A) = 0$*. Wir nennen* μ_A *das* Minimalpolynom *von* A*.*

Beweis: Die Aussage folgt unmittelbar aus der allgemeinen Aussage in Proposition B7.11, da $\mathrm{End}_K(V)$ sowie $\mathrm{Mat}_n(K)$ beides K-Algebren sind. Man beachte dabei, daß das Minimalpolynom nicht Null ist, da nach dem Satz von Cayley-Hamilton der Kern des Einsetzhomomorphismus nicht Null ist.

Zum besseren Verständnis skizzieren wir hier die wesentliche Beweisidee für μ_f. Wir betrachten die Menge

$$I := \{p \in K[t] \mid p(f) = 0\}$$

und beachten, daß für zwei Polynome p und q in I sowie ein beliebiges Polynom r stets auch $p + q \in I$ und $r \cdot p \in I$ gilt, wegen

$$(p + q)(f) = p(f) + q(f) = 0 + 0 = 0$$

und

$$(r \cdot p)(f) = r(f) \cdot p(f) = r(f) \cdot 0 = 0.$$

Aufgrund des Satzes von Cayley-Hamilton enthält I ein Nicht-Null-Polynom und wir können deshalb in I ein Nicht-Null-Polynom $0 \neq g \in I$ von minimalem Grad wählen. Normieren wir das Polynom g, indem wir durch seinen Leitkoeffizienten teilen, so erhalten wir ein Polynom $0 \neq \mu_f \in I$ vom selben minimalen Grad. Jedes Vielfache $\mu_f \cdot q$ mit $q \in K[t]$ liegt ebenfalls in I, was die Inklusion

$$I = \{p \in K[t] \mid p(f) = 0\} \supseteq \{\mu_f \cdot q \mid q \in K[t]\}$$

zeigt. Sei nun umgekehrt $p \in I$ gegeben, so können wir p mit Polynomdivision durch μ_f mit Rest teilen

$$p = q \cdot \mu_f + r,$$

wobei der Grad des Restes r echt kleiner ist als der von μ_f. Wegen

$$r = p - q \cdot \mu_f \in I$$

und wegen der Wahl von μ_f mit minimalem Grad, muß dann $r = 0$ gelten und $p = q \cdot \mu_f$ ist ein polynomielles Vielfaches von q, was die umgekehrte Inklusion zeigt. □

Bemerkung 10.10 (Minimalpolynome konjugierter Matrizen)

a. Sei $f \in \mathrm{End}_K(V)$, B eine Basis von V und $p \in K[t]$, dann gilt

$$M_B^B\big(p(f)\big) = p\big(M_B^B(f)\big),$$

da M_B^B ein K-Algebrenhomomorphismus ist.
Insbesondere gilt daher $p(f) = 0$ genau dann, wenn $p\big(M_B^B(f)\big) = 0$, und deshalb

$$\mu_f = \mu_{M_B^B(f)}.$$

Entsprechend gilt dann auch $\mu_{f_A} = \mu_{M_E^E(f_A)} = \mu_A$, wobei E die kanonische Basis von K^n bezeichnet.

b. Konjugierte Matrizen haben dasselbe Minimalpolynom, denn wegen Aufgabe 9.26 und Satz 10.9 gilt für konjugierte Matrizen $A, B \in \mathrm{Mat}_n(K)$

$$\mu_A \cdot K[t] = \{p \in K[t] \mid p(A) = 0\} = \{p \in K[t] \mid p(B) = 0\} = \mu_B \cdot K[t].$$

Korollar 10.11 (Die Eigenwerte sind die Nullstellen des Minimalpolynoms.)
Es sei $f \in \mathrm{End}_K(V)$ *und* $A \in \mathrm{Mat}_n(K)$.

a. *Die Eigenwerte von* f *sind genau die Nullstellen des Minimalpolynoms* μ_f.

b. *Die Eigenwerte von* A *sind genau die Nullstellen des Minimalpolynoms* μ_A.

Beweis: Sei zunächst $\lambda \in K$ eine Nullstelle des Minimalpolynoms μ_f. Wegen des Satzes von Cayley-Hamilton 10.7 und Satz 10.9 gibt es ein Polynom q mit $\chi_f = q \cdot \mu_f$, und mithin

$$\chi_f(\lambda) = q(\lambda) \cdot \mu_f(\lambda) = q(\lambda) \cdot 0 = 0,$$

so daß λ ein Eigenwert von f ist.

Ist umgekehrt $\lambda \in K$ ein Eigenwert von f, so gibt es einen Eigenvektor $0 \neq x \in V$ von f zu λ. Ist $\mu_f = \sum_{j=0}^m a_j \cdot t^j$, so folgt

$$\mu_f(f)(x) = \sum_{j=0}^m a_j \cdot f^j(x) = \sum_{j=0}^m a_j \cdot \lambda^j \cdot x = \mu_f(\lambda) \cdot x.$$

Da $\mu_f(f)$ die Nullabbildung ist, ist die linke Seite der Gleichung der Nullvektor, und da x nicht der Nullvektor ist, erzwingt dies

$$\mu_f(\lambda) = 0.$$

Die entsprechenden Aussagen für eine Matrix A folgen aus a. mit $f = f_A$.

□

Korollar 10.12 (Die Eigenwerte sind die Nullstellen des Minimalpolynoms.)
Es sei $f \in \mathrm{End}_K(V)$ *und* $A \in \mathrm{Mat}_n(K)$.

a. *Zerfällt das charakteristische Polynom von* f *in Linearfaktoren*

$$\chi_f = (t-\lambda_1)^{n_1} \cdot \ldots \cdot (t-\lambda_r)^{n_r}$$

für paarweise verschiedene $\lambda_i \in K$, *so gilt für das Minimalpolynom*

$$\mu_f = (t-\lambda_1)^{m_1} \cdot \ldots \cdot (t-\lambda_r)^{m_r}$$

mit $1 \leq m_i \leq n_i$ *für* $i = 1, \ldots, r$.

b. *Zerfällt das charakteristische Polynom von* A *in Linearfaktoren*

$$\chi_A = (t-\lambda_1)^{n_1} \cdot \ldots \cdot (t-\lambda_r)^{n_r}$$

für paarweise verschiedene $\lambda_i \in K$, *so gilt für das Minimalpolynom*

$$\mu_A = (t-\lambda_1)^{m_1} \cdot \ldots \cdot (t-\lambda_r)^{m_r}$$

mit $1 \leq m_i \leq n_i$ *für* $i = 1, \ldots, r$.

Beweis: Aus dem Satz zum Minimalpolynom 10.9 und dem Satz von Cayley-Hamilton 10.7 folgt, daß es ein Polynom $h \in K[t]$ gibt, so daß

$$\mu_f \cdot h = \chi_f = (t-\lambda_1)^{n_1} \cdot \ldots \cdot (t-\lambda_r)^{n_r}.$$

Die Primfaktorzerlegung von μ_f muß also von der Form

$$\mu_f = (t-\lambda_1)^{m_1} \cdot \ldots \cdot (t-\lambda_r)^{m_r}$$

sein, wobei für $i = 1, \ldots, r$

$$0 \leq m_i \leq n_i$$

gilt. Aus Korollar 10.11 folgt zudem $m_i > 0$ für alle $i = 1, \ldots, r$. □

Beispiel 10.13

a. Ist $A = \lambda \mathbb{1}_n \in \mathrm{Mat}_n(K)$ eine Diagonalmatrix mit gleichen Diagonalelementen, so gilt wegen $\lambda \mathbb{1}_n - A = 0$ offenbar $\chi_A = (t-\lambda)^n$ und $\mu_A = t - \lambda$.

b. Sei $\lambda \in K$ und

$$J := J_n(\lambda) := \begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \ldots & \ldots & 0 & \lambda \end{pmatrix} \in \mathrm{Mat}_n(K),$$

d. h. $J_n(\lambda)$ hat auf der Hauptdiagonalen den Wert λ und auf der oberen Nebendiagonalen Einsen stehen, ansonsten nur Nullen. Wir nennen $J_n(\lambda)$ einen *Jordanblock* (oder eine *Jordanzelle* oder ein *Jordankästchen*) der Größe n zum Eigenwert λ .

Offenbar gilt wieder

$$\chi_J = (t-\lambda)^n.$$

Nach Korollar 10.12 ist mithin $\mu_J = (t-\lambda)^m$ für ein $1 \leq m \leq n$. Dabei ist m die kleinste natürliche Zahl mit $(J - \lambda \mathbb{1}_n)^m = 0$. Nun ist aber

$$J - \lambda \mathbb{1}_n = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \ldots & \ldots & 0 & 0 \end{pmatrix} =: N$$

und man sieht mittels einer einfachen Induktion, daß $N^k \neq 0$ für $k < n$, aber $N^n = 0$ (vgl. Aufgabe 1.16). Also gilt

$$\mu_J = (t-\lambda)^n.$$

c. Ist $A = A_1 \oplus \ldots \oplus A_r \in \mathrm{Mat}_n(K)$ eine Diagonalblockmatrix mit $A_i \in \mathrm{Mat}_{n_i}(K)$, so folgt aus der Definition des charakteristischen Polynoms unmittelbar (vgl. Proposition 9.11)

$$\chi_A = \prod_{i=1}^r \chi_{A_i}.$$

Die entsprechende Formel für das Minimalpolynom gilt nicht. Sei etwa $A_1 = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in \mathrm{Mat}_2(K)$ und $A_2 = (1) \in \mathrm{Mat}_1(K)$, dann gilt für $A = A_1 \oplus A_2$

$$\mu_A = (t-1)^2 \neq (t-1)^3 = \mu_{A_1} \cdot \mu_{A_2}.$$

Man kann zeigen, daß für eine Diagonalblockmatrix wie oben μ_A ein kleinstes gemeinsames Vielfaches von $\mu_{A_1}, \ldots, \mu_{A_r}$ ist (siehe auch Aufgabe 10.29). Darauf wollen wir hier nicht näher eingehen.

Bemerkung 10.14 (Berechnung des Minimalpolynoms)

Zur praktischen Berechnung des Minimalpolynoms von $A \in \mathrm{Mat}_n(K)$ kann man wie folgt vorgehen. Aufgrund des Satzes von Cayley-Hamilton wissen wir, daß die Matrizen $A^0, \ldots, A^n$ linear abhängig sind. Fassen wir die Matrix A^i als einen *langen*

Spaltenvektor in K^{n^2} auf und bezeichnen wir diesen mit x_i, dann suchen wir das minimale m, so daß $x_0, \ldots, x_m$ linear abhängig sind, und wir suchen ferner geeignete $\beta_0, \ldots, \beta_{m-1}$ mit

$$x_m + \beta_{m-1}x_{m-1} + \ldots + \beta_0 x_0 = 0.$$

Dies ist dann gleichbedeutend damit, daß

$$t^m + \beta_{m-1}t^{m-1} + \ldots + \beta_0 \in K[t]$$

das gesuchte Minimalpolynom von A ist.

Bezeichne $X = (x_0 \ldots x_n) \in \mathrm{Mat}\big(n^2 \times (n+1), K\big)$ die Matrix, deren Spalten $x_0, \ldots, x_n$ sind, dann suchen wir eine Lösung des linearen Gleichungssystems

$$X \cdot \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_n \end{pmatrix} = 0 \in K^{n^2} \tag{30}$$

mit $\beta_{m+1} = \ldots = \beta_n = 0$ und $\beta_m = 1$ und so, daß m minimal mit dieser Eigenschaft ist. Da $(x_0, \ldots, x_{m-1})$ nach Definition von m linear unabhängig, $(x_0, \ldots, x_m)$ aber linear abhängig ist, bedeutet dies, daß in einer ZSF von X die Zahlen $1, \ldots, m$ Pivotindizes sind, während $m + 1$ kein Pivotindex mehr ist.

Berechnet man eine Basis des Lösungsraums von (30) mittels des Algorithmus 7.10, so erhalten wir den gesuchten Koeffizientenvektor β als das Negative des ersten Basisvektors, d.h. des ersten Vektors mit einer -1 auf der Diagonalen.

Dies führt zu folgendem Algorithmus zur Berechnung des Minimalpolynoms einer Matrix $A \in \mathrm{Mat}_n(K)$.

Algorithmus 10.15 (Algorithmus zur Berechnung des Minimalpolynoms)
INPUT: $A \in \mathrm{Mat}_n(K)$

OUTPUT: μ_A

1. **Schritt:** Falls A nicht quadratisch ist, gib 0 zurück.
2. **Schritt:** Bilde die Potenzen $A^0, \ldots, A^n$ und schreibe die Matrizen in Form von Spaltenvektoren der Länge n^2 in eine Matrix $X \in \mathrm{Mat}\big(n^2 \times (n+1), K\big)$.
3. **Schritt:** Berechne eine Basis von $\mathrm{Lös}(X, 0)$.
4. **Schritt:** Verwende die Negativen der Koeffizienten des ersten Basisvektors als Koeffizienten eines Polynoms und gib dieses zurück.

E) Die Hauptraumzerlegung

Für unsere weiteren Betrachtungen brauchen wir einen neuen Begriff, der auch im folgenden Abschnitt für die Jordansche Normalform von Bedeutung sein wird. Für $\lambda \in K$ haben wir aufsteigende Ketten von Unterräumen von V (vgl. Aufgabe 4.27)

$$\mathrm{Ker}\big(f - \lambda\,\mathrm{id}_V\big) \subseteq \mathrm{Ker}\big((f - \lambda\,\mathrm{id}_V)^2\big) \subseteq \mathrm{Ker}\big((f - \lambda\,\mathrm{id}_V)^3\big) \subseteq \ldots \subseteq V$$

und

$$\text{Lös}\big(A-\lambda\mathbb{1}_n,0\big)\subseteq\text{Lös}\big((A-\lambda\mathbb{1}_n)^2,0\big)\subseteq\text{Lös}\big((A-\lambda\mathbb{1}_n)^3,0\big)\subseteq\ldots\subseteq V$$

Die Vereinigung all dieser Unterräume ist offenbar wieder ein Unterraum und führt zu folgender Definition.

Definition 10.16 (Hauptraum)
Es sei $f\in\text{End}_K(V)$, $A\in\text{Mat}_n(K)$ und $\lambda\in K$. Dann heißen

$$\text{Hau}(f,\lambda)=\bigcup_{k\in\mathbb{N}}\text{Ker}\big((f-\lambda\,\text{id}_V)^k\big)\quad\text{und}\quad\text{Hau}(A,\lambda)=\bigcup_{k\in\mathbb{N}}\text{Lös}\big((A-\lambda\mathbb{1}_n)^k,0\big)$$

der *Hauptraum* oder *verallgemeinerte Eigenraum* von f bzw. A zu λ.

Lemma 10.17 (Fitting-Lemma)
Es sei $\lambda\in K$ gegeben.

a. *Für jedes $k\in\mathbb{N}$ sind* $\text{Ker}\big((f-\lambda\,\text{id}_V)^k\big)$ *sowie* $\text{Im}\big((f-\lambda\,\text{id}_V)^k\big)$ *f-invariant. Insbesondere sind also Eigenräume und Haupträume von f auch f-invariant.*

b. *Es gibt eine natürliche Zahl $0\leq m\leq n$ mit*

$$\text{Ker}\big((f-\lambda\,\text{id}_V)^0\big)\subsetneqq\text{Ker}\big((f-\lambda\,\text{id}_V)^1\big)\subsetneqq\ldots\subsetneqq\text{Ker}\big((f-\lambda\,\text{id}_V)^m\big)$$

und für $k>m$

$$\text{Hau}(f,\lambda)=\text{Ker}\big((f-\lambda\,\text{id}_V)^m\big)=\text{Ker}\big((f-\lambda\,\text{id}_V)^k\big).$$

Die Zahl m heißt Nilpotenzindex *von $f-\lambda\,\text{id}_V$ und erfüllt*

$$V=\text{Ker}\big((f-\lambda\,\text{id}_V)^m\big)\oplus\text{Im}\big((f-\lambda\,\text{id}_V)^m\big)$$

und

$$\mu_{f_{\text{Hau}(f,\lambda)}}=(t-\lambda)^m$$

sowie $m\leq\text{mult}(\mu_f,\lambda)$.

Die entsprechenden Aussagen für eine Matrix $A\in\text{Mat}_n(K)$ gelten analog.

Beweis: Durch Betrachtung von f_A ergibt sich die Aussage für eine Matrix A unmittelbar aus der entsprechenden Aussage für Endomorphismen.

a. Da f mit Potenzen von f und mit der Identität vertauschbar ist, gilt für $k\in\mathbb{N}$ und $x\in\text{Ker}\big((f-\lambda\,\text{id}_V)^k\big)$

$$(f-\lambda\,\text{id}_V)^k\big(f(x)\big)=f\big((f-\lambda\,\text{id}_V)^k(x)\big)=f(0)=0,$$

woraus der erste Teil der Behauptung folgt. Ferner gilt für $x=(f-\lambda\,\text{id}_V)^k(y)\in\text{Im}\big((f-\lambda\,\text{id}_V)^k\big)$ auch

$$f(x)=f\big((f-\lambda\,\text{id}_V)^k(y)\big)=(f-\lambda\,\text{id}_V)^k\big(f(y)\big)\in\text{Im}\big((f-\lambda\,\text{id}_V)^k\big),$$

woraus der zweite Teil der Behauptung folgt.

b. Wir setzen für den Beweis $g = f - \lambda \,\mathrm{id}_V$. Aus Aufgabe 4.27 wissen wir, daß es eine natürliche Zahl $0 \leq m \leq n$ gibt mit

$$\mathrm{Ker}(g^0) \subsetneqq \mathrm{Ker}(g^1) \subsetneqq \ldots \subsetneqq \mathrm{Ker}(g^m) \tag{31}$$

und

$$\mathrm{Ker}(g^m) = \mathrm{Ker}(g^k)$$

für $k > m$. Aus der Definition des Hauptraumes folgt dann unmittelbar

$$\mathrm{Hau}(f, \lambda) = \bigcup_{k \in \mathbb{N}} \mathrm{Ker}(g^k) = \mathrm{Ker}(g^m).$$

Ist nun $x \in \mathrm{Ker}(g^m) \cap \mathrm{Im}(g^m)$, so gilt $x = g^m(y)$ für ein $y \in V$ und zudem gilt

$$0 = g^m(x) = g^{2m}(y),$$

woraus

$$y \in \mathrm{Ker}(g^{2m}) = \mathrm{Ker}(g^m)$$

und damit

$$x = g^m(y) = 0$$

folgt. Damit haben wir

$$\mathrm{Ker}(g^m) \cap \mathrm{Im}(g^m) = \{0\}$$

gezeigt, und wegen

$$\dim_K(V) = \dim_K \mathrm{Ker}(g^m) + \dim_K \mathrm{Im}(g^m)$$

folgt dann schon $V = \mathrm{Ker}(g^m) \oplus \mathrm{Im}(g^m)$.
Nun beachten wir, daß

$$U := \mathrm{Ker}(g^m) = \mathrm{Ker}\left((f - \lambda \,\mathrm{id}_V)^m\right) = \mathrm{Hau}(f, \lambda)$$

ein f-invarianter Unterraum mit

$$(f_U - \lambda \,\mathrm{id}_U)^m = (f - \lambda \,\mathrm{id}_V)_U^m = 0$$

ist, woraus für das Minimalpolynom von f_U

$$\mu_{f_U} = (t - \lambda)^k$$

für ein $1 \leq k \leq m$ folgt. Wäre $k < m$, so würde aus $(f - \lambda \,\mathrm{id}_V)_U^k = 0$ schon

$$\mathrm{Ker}(g^m) = U \subseteq \mathrm{Ker}(g^k)$$

folgen, im Widerspruch zur Minimalität von m in (31). Also gilt

$$\mu_{f_U} = (t - \lambda)^m, \tag{32}$$

und wegen Aufgabe 10.29 ist dann $(t-\lambda)^m$ ein Teiler des Minimalpolynoms μ_f und damit

$$m \leq \mathrm{mult}(\mu_f, \lambda).$$

$\square$

Satz 10.18 (Hauptraumzerlegung)
Es sei $f \in \mathrm{End}_K(V)$ *so, daß* χ_f *über* K *in Linearfaktoren zerfällt, d. h. es gibt paarweise verschiedene* $\lambda_1, \ldots, \lambda_r \in K$ *und* $0 < m_i \leq n_i$*, so daß*

$$\chi_f = (t - \lambda_1)^{n_1} \cdots (t - \lambda_r)^{n_r} \quad \textit{und} \quad \mu_f = (t - \lambda_1)^{m_1} \cdots (t - \lambda_r)^{m_r}.$$

Dann gelten:

a. $V = \mathrm{Hau}(f, \lambda_1) \oplus \ldots \oplus \mathrm{Hau}(f, \lambda_r)$,

b. $n_i = \mathrm{mult}(\chi_f, \lambda_i) = \dim_K\big(\mathrm{Hau}(f, \lambda_i)\big)$ *und*

c. $m_i = \mathrm{mult}(\mu_f, \lambda_i)$ *ist der Nilpotenzindex von* $f - \lambda_i \,\mathrm{id}_V$.

Die analogen Aussagen für eine Matrix $A \in \mathrm{Mat}_n(K)$*, deren charakteristisches Polynom zerfällt, gelten analog.*

Beweis: Für den Beweis setzen wir $V_i := \mathrm{Hau}(f, \lambda_i) = \mathrm{Ker}\big((f - \lambda_i \,\mathrm{id}_V)^{k_i}\big)$, wobei k_i der Nilpotenzindex von $f - \lambda_i \,\mathrm{id}_V$ ist.

b. Betrachten wir $W_i := \mathrm{Im}\big((f - \lambda_i \,\mathrm{id}_V)^{k_i}\big)$, so sind V_i und W_i nach dem Fitting-Lemma 10.17 f-invariant und es gilt

$$V = V_i \oplus W_i.$$

Daraus folgt mit Proposition 9.11

$$(t - \lambda_i)^{n_i} \cdot \frac{\chi_f}{(t - \lambda_i)^{n_i}} = \chi_f = \chi_{f_{V_i}} \cdot \chi_{f_{W_i}}.$$

Aus dem Fitting-Lemma 10.17 wissen wir

$$\mu_{f_{V_i}} = (t - \lambda_i)^{k_i}, \tag{33}$$

so daß f_{V_i} wegen Korollar 10.11 nur den Eigenwert λ_i hat. Da das Polynom $\chi_{f_{V_i}}$ als Faktor von χ_f zerfällt, muß es

$$\chi_{f_{V_i}} = (t - \lambda_i)^{\dim_K(V_i)}$$

sein und es folgt

$$\dim_K(V_i) \leq n_i.$$

Wäre die Ungleichung strikt, wäre λ_i eine Nullstelle von $\chi_{f_{W_i}}$ und mithin gäbe es einen Eigenvektor $0 \neq x \in W_i$ von f zum Eigenwert λ_i, woraus der Widerspruch

$$0 \neq x \in \mathrm{Ker}(f - \lambda_i \,\mathrm{id}_V) \cap W_i \subseteq V_i \cap W_i = \{0\}$$

folgen würde. Mithin ist die Gleichheit gezeigt.

a. Wir zeigen die Aussage mittels Induktion nach der Anzahl r der Eigenwerte von f, wobei die Aussage für $r = 1$ unmittelbar aus Teil b. folgt, da dann $\dim_K(V) = n_1 = \dim_K(V_1)$ gilt.
Sei nun also $r \geq 2$. Aus dem Fitting-Lemma erhalten wir die Zerlegung

$$V = V_r \oplus W$$

und sowohl V_r, als auch W sind f-invariant. Wegen Teil b. erhalten wir mittels Proposition 9.11 dann

$$\chi_{f_W} = (t-\lambda_1)^{n_1} \cdot \ldots \cdot (t-\lambda_{r-1})^{n_{r-1}}.$$

Mit Induktion folgt dann aber

$$W = \mathrm{Hau}(f_W, \lambda_1) \oplus \ldots \oplus \mathrm{Hau}(f_W, \lambda_{r-1}),$$

wobei wegen b.

$$\dim_K \mathrm{Hau}(f_W, \lambda_i) = n_i = \dim_K \mathrm{Hau}(f, \lambda_i)$$

gilt. Beachtet man nun noch, daß

$$\mathrm{Hau}(f_W, \lambda_i) = \mathrm{Ker}\left((f_W - \lambda_i \,\mathrm{id}_W)^{l_i}\right) \subseteq \mathrm{Ker}\left((f - \lambda_i \,\mathrm{id})^{l_i}\right) \subseteq \mathrm{Hau}(f, \lambda_i)$$

für ein geeignetes l_i gilt, so folgt notwendigerweise

$$\mathrm{Hau}(f_W, \lambda_i) = \mathrm{Hau}(f, \lambda_i)$$

und Teil a. ist gezeigt.

c. Für den Nilpotenzindex k_i von $f - \lambda_i \,\mathrm{id}_V$ gilt nach Lemma 10.17 $m_i \geq k_i$. Wegen (33) gilt

$$p := \mu_{f_{V_1}} \cdot \ldots \cdot \mu_{f_{V_r}} = (t-\lambda_1)^{k_1} \cdot \ldots \cdot (t-\lambda_r)^{k_r} = p_i \cdot (t-\lambda_i)^{k_i}$$

für

$$p_i = \prod_{j \neq i} (t-\lambda_j)^{k_j},$$

und für jeden Vektor $x = x_1 + \ldots + x_r \in V$ mit $x_i \in V_i$ für $i = 1, \ldots, r$ gilt dann

$$p(f)(x) = \sum_{i=1}^{r} p_i(f) \circ (f - \lambda_i \,\mathrm{id}_V)^{k_i}(x_i) = \sum_{i=1}^{r} p_i(f)(0) = 0.$$

Also ist μ_f ein Teiler von p, woraus notwendigerweise $m_i \leq k_i$ für alle $i = 1, \ldots, r$ folgt.

Die entsprechende Aussage für eine Matrix A läßt sich unmittelbar auf die Aussage für f_A zurückführen. □

Aus Satz 10.18 Teil b. und c. folgt, da die Haupträume von f f-invariant sind, unmittelbar das folgende Korollar.

Korollar 10.19
Sei f *wie in Satz 10.18, dann gilt*

$$\chi_{f_{\mathrm{Hau}(f,\lambda_i)}} = (t-\lambda_i)^{n_i}$$

und

$$\mu_{f_{\mathrm{Hau}(f,\lambda_i)}} = (t-\lambda_i)^{m_i}.$$

F) Diagonalisierbarkeit

Satz 10.20 (Diagonalisierbarkeit von Endomorphismen)
Für einen Endomorphismus $f \in \mathrm{End}_K(V)$ *sind die folgenden Aussagen äquivalent:*

a. f *ist diagonalisierbar.*

b. V *hat eine Basis aus Eigenvektoren von* f.

c. *Sind* $\lambda_1, \ldots, \lambda_r$ *die paarweise verschiedenen Eigenwerte von* f *in* K*, dann gilt*

$$V = \bigoplus_{i=1}^{r} \mathrm{Eig}(f, \lambda_i).$$

d. *Das charakteristische Polynom von* f *zerfällt über* K *in Linearfaktoren und für jeden Eigenwert* λ *stimmen algebraische und geometrische Vielfachheit überein.*

e. *Das Minimalpolynom von* f *zerfällt über* K *in paarweise verschiedene Linearfaktoren und das charakteristische Polynom zerfällt über* K.

Beweis:
a. $\Rightarrow$ e.: Ist f diagonalisierbar, dann gibt es eine Basis B von V, so daß

$$M_B^B(f) = \bigoplus_{i=1}^{r} \lambda_i \mathbb{1}_{n_i},$$

mit $\lambda_i \neq \lambda_j$ für $i \neq j$. Setzen wir $p = (t - \lambda_1) \cdot \ldots \cdot (t - \lambda_r) \in K[t]$, so ist

$$p\big(M_B^B(f)\big) = p(\lambda_1 \mathbb{1}_{n_1}) \oplus \ldots \oplus p(\lambda_r \mathbb{1}_{n_r})$$

wegen Bemerkung 10.6 eine Diagonalblockmatrix, und für die Blöcke gilt

$$p(\lambda_i \mathbb{1}_{n_i}) = p(\lambda_i) \cdot \mathbb{1}_{n_i} = 0.$$

Also ist schon $p(f) = 0$ erfüllt und p ist ein Vielfaches von μ_f. Dann zerfällt μ_f aber in paarweise verschiedene Linearfaktoren. Außerdem zerfällt auch χ_f wegen Satz 10.2.

e. $\Rightarrow$ d.: Aus dem Satz zur Hauptraumzerlegung 10.18 folgt zudem $\mathrm{Hau}(f, \lambda_i) = \mathrm{Eig}(f, \lambda_i)$, da der Nilpotenzindex von $f - \lambda_i \,\mathrm{id}_V$ eins ist, und

$$\mathrm{mult}(\chi_f, \lambda_i) = n_i = \dim_K \mathrm{Hau}(f, \lambda_i) = \dim_K \mathrm{Eig}(f, \lambda_i),$$

d.h. die geometrische und die algebraische Vielfachheit jedes Eigenwertes stimmen überein.

d. $\Rightarrow$ c.: Das charakteristische Polynom habe die Primfaktorzerlegung

$$\chi_f = (t - \lambda_1)^{n_1} \cdot \ldots \cdot (t - \lambda_r)^{n_r}.$$

Aus dem Satz über die Hauptraumzerlegung und der Voraussetzung folgt dann

$$\dim_K \mathrm{Hau}(f, \lambda_i) = n_i = \mathrm{mult}(\chi_f, \lambda_i) = \dim_K \mathrm{Eig}(f, \lambda_i),$$

und da stets $\mathrm{Eig}(f,\lambda_i) \subseteq \mathrm{Hau}(f,\lambda_i)$ gilt, folgt

$$\mathrm{Eig}(f,\lambda_i) = \mathrm{Hau}(f,\lambda_i).$$

Aus dem Satz über die Hauptraumzerlegung folgt dann aber wiederum

$$V = \mathrm{Hau}(f,\lambda_1) \oplus \ldots \oplus \mathrm{Hau}(f,\lambda_r) = \mathrm{Eig}(f,\lambda_1) \oplus \ldots \oplus \mathrm{Eig}(f,\lambda_r).$$

c. $\Rightarrow$ b.: Es sei B_i eine Basis von $\mathrm{Eig}(f,\lambda_i)$, dann ist nach Aufgabe 3.30 $B = B_1 \cup \ldots \cup B_r$ eine Basis von V, die aus Eigenvektoren besteht.

b. $\Rightarrow$ a.: Ist $B = (x_1,\ldots,x_n)$ eine Basis von V aus Eigenvektoren, so ist $f(x_i) = \lambda_i \cdot x_i$ für ein geeignetes $\lambda_i \in K$. Damit ist dann aber $M_B^B(f)$ eine Diagonalmatrix mit den Diagonaleinträgen $\lambda_1,\ldots,\lambda_n$. □

Korollar 10.21 (Diagonalisierbarkeit von Matrizen)
Für eine quadratische Matrix $A \in \mathrm{Mat}_n(K)$ sind die folgenden Aussagen äquivalent:

a. *A ist diagonalisierbar.*

b. *K^n hat eine Basis aus Eigenvektoren von A.*

c. *Sind $\lambda_1,\ldots,\lambda_r$ die paarweise verschiedenen Eigenwerte von A, dann gilt*

$$K^n = \bigoplus_{i=1}^{r} \mathrm{Eig}(A,\lambda_i).$$

d. *Das charakteristische Polynom von A zerfällt über K in Linearfaktoren und für jeden Eigenwert λ stimmen algebraische und geometrische Vielfachheit überein.*

e. *Das Minimalpolynom von A zerfällt über K in paarweise verschiedene Linearfaktoren und das charakteristische Polynom zerfällt über K.*

Insbesondere, genau dann ist $T \in \mathrm{Gl}_n(K)$ so, daß $T^{-1} \circ A \circ T$ eine Diagonalmatrix ist, wenn die Spalten von T eine Basis des K^n aus Eigenvektoren von A sind.

Beweis: Wende Satz 10.20 auf die Abbildung f_A an. □

Bemerkung 10.22
In den beiden letzten Sätzen kann in Bedingnung e. jeweils darauf verzichtet werden, zu prüfen, daß das charakteristische Polynom zerfällt, da dies in der Tat aus dem Zerfallen des Minimalpolynoms folgt. Um dies zu zeigen, muß man aber zu Körpererweiterungen übergehen, worauf wir in dieser Vorlesung verzichten wollen.

Falls ein Endomorphismus oder eine Matrix hinreichend viele verschiedene Eigenwerte hat, so folgt aus den obigen Überlegungen unmittelbar deren Diagonalisierbarkeit.

Korollar 10.23 (Diagonalisierbarkeit)
Es sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

a. *Hat f genau n paarweise verschiedene Eigenwerte, so ist f diagonalisierbar.*

b. *Hat A genau n paarweise verschiedene Eigenwerte, so ist A diagonalisierbar.*

Beweis: Hat f genau n paarweise verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_n \in K$, so muß

$$\chi_f = \mu_f = (t - \lambda_1) \cdot \ldots \cdot (t - \lambda_n)$$

gelten. D.h. μ_f zerfällt in paarweise verschiedene Linearfaktoren und f ist diagonalisierbar. Der Beweis für A geht analog. □

Aus Korollar 10.21 können wir ein Verfahren ableiten, das es uns erlaubt, eine Matrix zu diagonalisieren und die Transformationsmatrix T zu berechnen.

Algorithmus 10.24 (Algorithmus zur Diagonalisierung)

INPUT: $A \in \mathrm{Mat}_n(K)$.

OUTPUT: 0, falls A über K nicht diagonalisierbar ist,
$1, D, T$, falls A diagonaliserbar ist, wobei D eine zu A konjugierte Diagonalmatrix ist, und T die zugehörige Transformationsmatrix mit $T^{-1} \circ A \circ T = D$.

1. **Schritt:** Berechne das charakteristische Polynom von A.
2. **Schritt:** Faktorisiere das charakteristische Polynom über K. Ist einer der Faktoren nicht linear, ist A nicht diagonalisierbar (nicht einmal trigonalisierbar) und man gebe 0 zurück. Sind alle Faktoren linear, so liefert die Faktorisierung die Eigenwerte $\lambda_1, \ldots, \lambda_r$ sowie ihre algebraischen Vielfachheiten $n_1, \ldots, n_r$.
3. **Schritt:** Bestimme für jeden Eigenwert λ_i eine Basis des Eigenraums $\mathrm{Eig}(A, \lambda_i)$ als $\mathrm{Lös}(A - \lambda_i \mathbb{1}_n, 0)$ - vgl. Algorithmus 7.10 - sowie seine Dimension - vgl. Algorithmus 6.20 -, d. h. die geometrische Vielfachheit von λ_i.
4. **Schritt:** Stimmt für jeden Eigenwert die algebraische Vielfachheit mit der geometrischen überein, so schreibe man die im 3. Schritt bestimmten Basen als Spalten in eine Matrix und erhält so T. Ferner erhält man D, indem man die Eigenwerte $\lambda_1, \ldots, \lambda_r$ entsprechend ihren algebraischen Vielfachheiten in der Diagonalen einer Nullmatrix einträgt.

Beispiel 10.25

Gegeben sei die Matrix

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & -1 & -1 & 1 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Q}).$$

Das charakteristische Polynom von A berechnet man mit Hilfe zweifacher Laplace-Entwicklung nach der jeweils letzten Spalte als

$$\chi_A = \begin{vmatrix} t-2 & 1 & 0 & 0 \\ 0 & t-1 & 0 & 0 \\ 0 & 0 & t-2 & 0 \\ -1 & 1 & 1 & t-1 \end{vmatrix} = (t-1) \cdot (t-2) \cdot \begin{vmatrix} t-2 & 1 \\ 0 & t-1 \end{vmatrix} = (t-1)^2 \cdot (t-2)^2.$$

Damit ist also $\sigma(A) = \{1, 2\}$ mit $\mathrm{mult}(\chi_A, 1) = \mathrm{mult}(\chi_A, 2) = 2$.

Als nächstes berechnen wir den Eigenraum $\text{Lös}(2\mathbb{1}_4 - A, 0)$ zum Eigenwert $\lambda = 2$:

$$2 \cdot \mathbb{1}_4 - A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & 1 \end{pmatrix} \xmapsto[\substack{IV \mapsto IV - II \\ I \mapsto I + II}]{\substack{I \leftrightarrow IV \\ I \mapsto -I}} \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xmapsto{-1\text{'en einfügen}} \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Mithin ist

$$\text{Eig}(A, 2) = \text{Lin}\left((-1, 0, -1, 0)^t, (-1, 0, 0, -1)^t\right)$$

und

$$\dim_{\mathbb{Q}} \text{Eig}(A, 2) = 2 = \text{mult}(\chi_A, 2).$$

Dann berechnen wir den Eigenraum $\text{Lös}(\mathbb{1}_4 - A, 0)$ zum Eigenwert $\lambda = 1$:

$$1 \cdot \mathbb{1}_4 - A = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ -1 & 1 & 1 & 0 \end{pmatrix} \xmapsto[III \mapsto -III]{\substack{IV \mapsto IV - I + III \\ I \mapsto -I}} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xmapsto{-1\text{'en einfügen}} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Mithin ist

$$\text{Eig}(A, 1) = \text{Lin}\left((-1, -1, 0, 0)^t, (0, 0, 0, -1)^t\right)$$

und

$$\dim_{\mathbb{Q}} \text{Eig}(A, 1) = 2 = \text{mult}(\chi_A, 1).$$

Also zerfällt χ_A über $\mathbb{Q}$ in Linearfaktoren und die geometrischen Vielfachheiten der Eigenwerte stimmen mit den algebraischen überein, so daß A diagonalisierbar ist. Zudem gilt für

$$T = \begin{pmatrix} -1 & -1 & -1 & 0 \\ 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 \end{pmatrix},$$

daß

$$T^{-1} \circ A \circ T = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Aufgaben

Aufgabe 10.26 (Zyklische Unterräume)
Zeige, $\chi_{f_U} = \mu_{f_U} = t^m$ für den Endomorphismus f_U aus Aufgabe 5.38.

Aufgabe 10.27
Es sei $E = (E_{11}, E_{12}, E_{21}, E_{22})$ die kanonische Basis von $V = \text{Mat}_2(K)$ und $T = E_{11} + E_{12} + E_{22} \in \text{Gl}_2(K)$. Zeige, daß der Endomorphismus $f : V \to V : A \mapsto T \circ A \circ T^{-1}$ trigonalisierbar, aber nicht diagonalisierbar ist, und bestimme eine Basis B von V, so daß $M_B^B(f)$ eine obere Dreiecksmatrix ist.

Aufgabe 10.28
Zeige, ist $A \in \mathrm{Gl}_n(K)$, so gibt es ein Polynom $p \in K[t]$ mit $A^{-1} = p(A)$.

Aufgabe 10.29
Ist $f \in \mathrm{End}_K(V)$ und ist $U \leq V$ ein f-invarianter Unterraum von V, so ist das Minimalpolynom μ_{f_U} von f_U ein Teiler des Minimalpolynoms μ_f von f.

Aufgabe 10.30
Zeige, ist $1 \leq \dim_K(V) = n < \infty$, so sind für $\mathcal{A} \subseteq \mathrm{End}_K(V)$ die folgenden beiden Aussagen gleichwertig:

a. $\mathcal{A}$ ist simultan diagonalisierbar, d. h. es gibt eine Basis B von V, so daß für alle $f \in \mathcal{A}$ gilt $M_B^B(f)$ ist eine Diagonalmatrix.

b. Für alle $f \in \mathcal{A}$ gilt, f ist diagonalisierbar, und für alle $f, g \in \mathcal{A}$ gilt, $f \circ g = g \circ f$.

Hinweis: Führe für "b. $\Rightarrow$ a." Induktion über n und zerlege dazu V in zwei invariante Unterräume kleinerer Dimension.

§ 11 Die Jordansche Normalform

Eine Matrix $A \in \mathrm{Mat}_n(K)$, deren charakteristisches Polynom in Linearfaktoren zerfällt, was etwa für einen algebraisch abgeschlossenen Körper wie $\mathbb{C}$ stets der Fall ist, ist zu einer Matrix konjugiert, die besonders einfach gebaut ist, der sog. Jordanschen Normalform von A. Aus der Jordanschen Normalform lassen sich Invarianten von A einfach ablesen und diese Invarianten bestimmen die Matrix A bis auf Konjugation eindeutig.

Satz 11.1 (Jordansche Normalform eines Endomorphismus)
Es sei $f \in \mathrm{End}_K(V)$ *ein Endomorphismus, dessen charakteristisches Polynom über* K *zerfällt,* $\chi_f = (t-\lambda_1)^{n_1} \cdot \ldots \cdot (t-\lambda_r)^{n_r}$*, und es sei* $\mu_f = (t-\lambda_1)^{m_1} \cdot \ldots \cdot (t-\lambda_r)^{m_r}$*.*

Dann gibt es für jedes $i = 1, \ldots, r$ *und* $1 \leq j \leq m_i$*, je eine natürliche Zahl* t_{ij} *und es gibt eine Basis* B *so, daß*

(1) $\sum_{j=1}^{m_i} j \cdot t_{ij} = n_i = \dim_K \mathrm{Hau}(f, \lambda_i)$,

(2) $\sum_{j=1}^{m_i} t_{ij} = \dim_K \mathrm{Eig}(f, \lambda_i)$,

(3) $t_{im_i} \geq 1$ *und*

$$J_f := M_B^B(f) = \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i).$$

J_f *heißt* Jordansche Normalform *von* f*, und die* t_{ij} *werden* Elementarteiler *von* f *zum Eigenwert* λ_i *genannt.*

Korollar 11.2 (Jordansche Normalform einer quadratischen Matrix)
Es sei $A \in \mathrm{Mat}_n(K)$ *eine quadratische Matrix, deren charakteristisches Polynom über* K *zerfällt,* $\chi_A = (t-\lambda_1)^{n_1} \cdot \ldots \cdot (t-\lambda_r)^{n_r}$*, und es sei* $\mu_A = (t-\lambda_1)^{m_1} \cdot \ldots \cdot (t-\lambda_r)^{m_r}$*.*

Dann gibt es für jedes $i = 1, \ldots, r$ *und* $1 \leq j \leq m_i$*, je eine natürliche Zahl* t_{ij} *und es gibt ein invertierbare Matrix* $T \in \mathrm{Gl}_n(K)$ *so, daß*

(1) $\sum_{j=1}^{m_i} j \cdot t_{ij} = n_i = \dim_K \mathrm{Hau}(A, \lambda_i)$,

(2) $\sum_{j=1}^{m_i} t_{ij} = \dim_K \mathrm{Eig}(A, \lambda_i)$,

(3) $t_{im_i} \geq 1$ *und*

$$J_A := T^{-1} \circ A \circ T = \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i).$$

J_A *heißt* Jordansche Normalform *von* A*, und die* t_{ij} *werden* Elementarteiler *von* A *zum Eigenwert* λ_i *genannt.*

Beweis: Der Beweis folgt aus Satz 11.1 mit $f = f_A$. □

Es scheint angebracht, den Satz zunächst etwas zu erläutern, um ihn verständlicher zu machen.

Bemerkung 11.3 (Jordansche Normalform)

a. Ziel des Abschnittes ist es, zu zeigen, daß eine Matrix A, deren charakteristisches Polynom zerfällt, konjugiert zu einer Matrix von besonders einfacher Gestalt ist. Der obige Satz sagt nun, daß in der Tat A konjugiert ist zu einer Diagonalblockmatrix, deren Diagonalblöcke, die $J_j(\lambda_i)$, alle Jordanblöcke sind, also obere Dreiecksmatrizen, die auf der Diagonalen stets den gleichen Wert λ_i stehen haben, auf der oberen Nebendiagonalen nur Einsen und ansonsten nur Nullen (vgl. Beispiel 10.13).
Dabei gelten:
   - Die natürlichen Zahlen t_{ij} geben an, wieviele Jordanblöcke der Größe $j \times j$ zum Eigenwert λ_i denn vorkommen.
   - $j \leq m_i$ bedeutet, die maximale Größe eines Jordanblockes ist $m_i \times m_i$.
   - $t_{im_i} \geq 1$ besagt, daß auch mindestens ein Block der maximalen Größe $m_i \times m_i$ vorkommt. D. h. die Vielfachheit von λ_i als Nullstelle von μ_A gibt die maximale Größe eines vorkommenden Jordanblockes in J_A zum Eigenwert λ_i an.
   - Die Summe $\sum_{j=1}^{m_i} j \cdot t_{ij}$ gibt gerade an, wie oft der Eigenwert λ_i auf der Diagonalen der Diagonalblockmatrix vorkommt, und da diese das gleiche charakteristische Polynom wie A besitzt, muß die Summe mithin n_i, also die algebraische Vielfachheit von λ_i als Eigenwert von A, sein.
   - Und $\sum_{j=1}^{m_i} t_{ij} = \dim_K \operatorname{Eig}(A, \lambda_i)$ bedeutet schließlich, daß die Anzahl der Jordanblöcke zum Eigenwert λ_i, die in J_A vorkommen, der Dimension des Eigenraumes von A zum Eigenwert λ_i entspricht, d.h. seiner geometrischen Vielfachheit.

b. Schon die direkte Summenschreibweise der Jordanschen Normalform bringt zum Ausdruck, daß die Jordansche Normalform nur bis auf die Reihenfolge der Jordanblöcke eindeutig bestimmt sein kann, und in der Tat ist sie es auch, d. h.:

Zwei Jordansche Normalformen sind genau dann konjugiert, wenn ihre Eigenwerte und die zugehörigen Elementarteiler übereinstimmen.

   Es ist leicht einsichtig, daß eine Vertauschung der Blöcke durch Konjugation mit einer Reihe von Permutationsmatrizen erreicht werden kann, daß mithin zwei Jordansche Normalformen, deren Eigenwerte mit zugehörigen Elementarteilern übereinstimmen, zueinander konjugiert sind.
   Seien umgekehrt zwei Jordansche Normalformen zueinander konjugiert, dann stimmen zunächst die charakteristischen Polynome und damit die Eigenwerte

überein. Ferner folgt aus Aufgabe 11.20, daß die Elementarteiler übereinstimmen, da für eine invertierbare Matrix $T \in \mathrm{Gl}_n(K)$ und ein $k \in \mathbb{N}$ gilt

$$\begin{aligned}\operatorname{rang}\Big(\big(T^{-1} \circ A \circ T - \lambda \mathbb{1}_n\big)^k\Big) &= \operatorname{rang}\big(T^{-1} \circ (A - \lambda \mathbb{1}_n)^k \circ T\big) \\ &= \operatorname{rang}\big((A - \lambda \mathbb{1}_n)^k\big).\end{aligned}$$

Damit ist natürlich auch die Jordansche Normalform eines Endomorphismus bis auf die Reihenfolge der Jordanblöcke eindeutig bestimmt.

c. Wir wollen folgende Notation einführen, die die Jordanblöcke von A (bzw. f) zu einem Eigenwert λ_i zusammenfaßt:

$$J_A(\lambda_i) := \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i) \quad \text{bzw.} \quad J_f(\lambda_i) := \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i)$$

Dann gilt

$$J_A = \bigoplus_{i=1}^{r} J_A(\lambda_i) \quad \text{bzw.} \quad J_f = \bigoplus_{i=1}^{r} J_f(\lambda_i).$$

Beispiel 11.4 (Jordansche Normalform)
Wir wollen nun in einigen einfachen Fällen die Jordansche Normalform bestimmen.

a. Die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 6 & 7 \\ 0 & 0 & 8 & 9 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Q})$$

ist eine obere Dreiecksmatrix und ihr charakteristisches Polynom

$$\chi_A = (t-1) \cdot (t-5) \cdot (t-8) \cdot (t-2)$$

zerfällt in paarweise verschiedene Linearfaktoren. Da zu jedem der Eigenwerte ein Jordanblock gehören muß und da die Matrix J_A nicht mehr als vier Jordanblöcke aufnehmen kann, gilt also

$$J_A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Über die Transformationsmatrix T ist damit noch nichts gesagt. Da die Matrix J_A aber eine Diagonalmatrix ist, wissen wir aus Korollar 10.21 bereits, daß die Spalten von T Eigenvektoren zu den vier Eigenwerten sein müssen. Wir könnten T also leicht berechnen.

b. Die Matrix

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Q})$$

hat offenbar den Rang eins. Deshalb gilt für die geometrische Vielfachheit von 0 als Eigenwert von A

$$\dim_K \mathrm{Eig}(A, 0) = \dim_K \mathrm{Lös}(A, 0) = 4 - \mathrm{rang}(A) = 3. \tag{34}$$

Da die algebraische Vielfachheit von 0 als Eigenwert von A mindestens so groß sein muß wie die geometrische, besitzt im charakteristischen Polynom χ_A von A der Linearfaktor t also mindestens Vielfachheit 3. Deshalb gibt es ein $\lambda \in \mathbb{Q}$ mit

$$\chi_A = t^3 \cdot (t - \lambda) = t^4 - \lambda \cdot t^3.$$

Aus Lemma 9.6 wissen wir aber, daß der zweithöchste Koeffizient des charakteristischen Polynoms das Negative der Spur der Matrix ist, d.h. $\lambda = \mathrm{Spur}(A) = 4$. Wir haben also

$$\chi_A = t^3 \cdot (t - 4).$$

Aus (34) folgt, daß es drei Jordanblöcke zum Eigenwert 0 geben muß, und außerdem muß es einen Jordanblock zum Eigenwert 4 geben. Da aber wieder höchstens vier Jordanblöcke in J_A passen, gilt

$$J_A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

Die Transformationsmatrix T enthält als Spalten also auch wieder Eigenvektoren und läßt sich so leicht berechnen.

c. Wir betrachten wieder die Matrix A aus dem vorherigen Teil, fassen sie nun aber als Matrix über dem Körper $\mathbb{F}_2$ auf, d.h.

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{F}_2).$$

Wie oben sieht man, daß die Matrix Rang eins hat und somit 0 die geometrische Vielfachheit 3 besitzt. Und mit den gleichen Argumenten erhalten wir

$$\chi_A = t^3 \cdot \big(t - \mathrm{Spur}(A)\big).$$

Allerdings ist die Spur diesmal

$$\mathrm{Spur}(A) = 1 + 1 + 1 + 1 = 0 \in \mathbb{F}_2,$$

so daß wir

$$\chi_A = t^4$$

erhalten. 0 hat die geometrische Vielfachheit 3 und hat somit exakt drei Jordanblöcke zum Eigenwert 0, und da A keine anderen Eigenwerte besitzt, muß einer dieser Jordanblöcke diesmal die Größe zwei haben! Wir erhalten also

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

Die Matrix A ist diesmal also *nicht* diagonalisierbar und wir wissen deshalb auch noch nicht, wie wir die Transformationsmatrix T bestimmen sollten!

d. Es sei $A \in \mathrm{Mat}_3(\mathbb{Q})$ mit der Eigenschaft $A^3 - A^2 = 0$. Was können wir über die Jordansche Normalform von A sagen?
A ist eine Nullstelle des Polynoms

$$p = t^3 - t^2 = t^2 \cdot (t-1).$$

Das Minimalpolynom von A muß nach Satz 10.9 ein Teiler von p sein, so daß für μ_A nur folgende Möglichkeiten in Betracht kommen:

$$\mu_A \in \{t, t-1, t\cdot(t-1), t^2, t^2\cdot(t-1)\}.$$

Daraus ergeben sich für die Jordansche Normalform bis auf die Reihenfolge der Jordanblöcke folgenden Möglichkeiten:

μ_A	t	$t-1$	$t\cdot(t-1)$	t^2	$t^2\cdot(t-1)$
J_A	$\begin{pmatrix} 0&0&0\\0&0&0\\0&0&0 \end{pmatrix}$	$\begin{pmatrix} 1&0&0\\0&1&0\\0&0&1 \end{pmatrix}$	$\begin{pmatrix} 0&0&0\\0&1&0\\0&0&1 \end{pmatrix}$ $\begin{pmatrix} 0&0&0\\0&0&0\\0&0&1 \end{pmatrix}$	$\begin{pmatrix} 0&1&0\\0&0&0\\0&0&0 \end{pmatrix}$	$\begin{pmatrix} 0&1&0\\0&0&0\\0&0&1 \end{pmatrix}$

Dabei ist die Situation für $\mu_A = t$ oder $\mu_A = t-1$ klar, da dann schon A selbst die angegebene Jordansche Normalform sein muß, wie man durch einsetzen von A in die Gleichung sieht.
Ist $\mu_A = t\cdot(t-1)$, so zerfällt das Minimalpolynom in paarweise verschiedene Linearfaktoren und A ist nach Korollar 10.21 diagonalisierbar. Zudem muß für jeden Eigenwert mindestens ein Jordanblock vorkommen, so daß genau die beiden angegebenen Matrizen in Frage kommen.
Wenn $\mu_A = t^2$ ist, so muß ein Jordanblock der Größe zwei zum Eigenwert 0 vorkommen und da nur Blöcke zum Eigenwert 0 vorkommen können, sind wir dann auch schon fertig. $\mu_A = t^2\cdot(t-1)$ geht analog.

Wir werden Satz 11.1 zunächst für nilpotente Endomorphismen zeigen, d. h. für Endomorphismen, die nur einen Eigenwert, nämlich $\lambda = 0$, besitzen, und den allgemeinen Fall dann auf diesen zurückführen.

Definition 11.5 (Nilpotent)
Wir nennen einen Endomorphismus $f \in \mathrm{End}_K(V)$ bzw. eine Matrix $A \in \mathrm{Mat}_n(K)$ *nilpotent*, wenn es ein $r \in \mathbb{N}$ gibt, so daß $f^r = 0$ bzw. $A^r = 0$. Offenbar gilt dann $\mu_f = t^m$ bzw. $\mu_A = t^m$ für ein $1 \leq m \leq r$.

Beispiel 11.6 (Ein nilpotentes Jordankästchen)
Sei $f \in \mathrm{End}_K(V)$ und sei $B = (x_1, \ldots, x_n)$ eine Basis von V, so daß

$$M_B^B(f) = J_n(0) := \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \ldots & \ldots & 0 & 0 \end{pmatrix}$$

ein Jordankästchen der Größe n zum Eigenwert 0 ist, dann folgt aus der Matrixdarstellung zunächst

$$f(x_{i+1}) = x_i$$

und damit

$$x_i = f^{n-i}(x_n)$$

für $i = 1, \ldots, n-1$. Das heißt, V ist ein zyklischer Unterraum mit seiner kanonischen Basis

$$B = \left(f^{n-1}(x_n), f^{n-2}(x_n), \ldots, f(x_n), x_n\right)$$

wie wir sie in Aufgabe 5.38 betrachtet haben. Man beachte auch, daß die Matrix $M_B^B(f)$ und damit der Endomorphismus f nilpotent mit Nilpotenzindex n ist (siehe Aufgabe 1.16).

Wir wollen im folgenden zeigen, daß die Jordansche Normalform eines nilpotenten Endomorphismus eine Blockdiagonalmatrix ist, deren Diagonalblöcke Jordankästchen der obigen Form sind. Das Beispiel sagt uns also, welche Gestalt der Anteil der Basis haben muß, der zu einem solchen Kästchen gehört!

Definition 11.7 (Partitionen)
Eine *Partition* der positiven natürlichen Zahl n ist ein Tupel $P = (k_1, \ldots, k_m)$ natürlicher Zahlen, so daß $k_1 \geq k_2 \geq \ldots \geq k_m \geq 1$ und $k_1 + k_2 + \ldots + k_m = n$.

Lemma 11.8 (Die duale Partition)
Ist $P = (k_1, \ldots, k_m)$ *eine Partition von* n*, ist auch* $P^* = (l_1, \ldots, l_s)$ *mit* $s = k_1$ *und*

$$l_i = |\{j \mid 1 \leq j \leq m, k_j \geq i\}|$$

eine Partition von n*, die sogenannte* duale Partition *zu* P.

Beweis: Man kann die Partition P durch ihr *Young-Diagramm* veranschaulichen. Dieses besteht aus n Kästen, die in r Reihen übereinander angeordnet sind, wobei in der i-ten Reihe genau k_i Kästchen sind (siehe Abbildung 1).

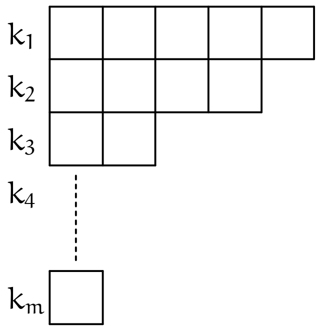


ABBILDUNG 1. Young-Diagramm von $P = (k_1, \ldots, k_m)$

Die Anzahl an Kästchen in der i-ten Spalte ist dann genau l_i. Damit ist die Summe der l_i gerade n und $l_1 \geq l_2 \geq \ldots \geq l_s$. □

Beispiel 11.9
$P = (5, 4, 2, 2, 2)$ ist eine Partition von $n = 15$ mit dem folgenden Young-Diagramm (siehe Abbildung 2).

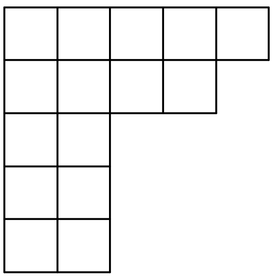


ABBILDUNG 2. Young-Diagramm von $P = (5, 4, 2, 2, 2)$

Das Young-Diagramm der dualen Partition $P^* = (5, 5, 2, 2, 1)$ entsteht durch Spiegelung an der Winkelhalbierenden (siehe Abbildung 3).

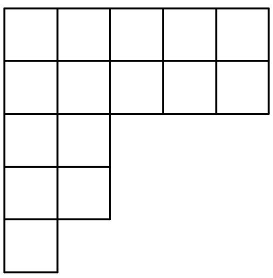


ABBILDUNG 3. Young-Diagramm von $P^* = (5, 5, 2, 2, 1)$

Bemerkung 11.10 (Anzahl der Partitionen von n)
Die Funktion
$$\pi : \mathbb{Z}_{>0} \longrightarrow \mathbb{Z}_{>0},$$
die einer natürlichen Zahl n die Anzahl der Partitionen von n zuordnet, ist eine interessante zahlentheoretische Funktion. Wir wollen einige Werte von π zur Veranschaulichung ihrer Komplexität angeben:

n	1	2	3	4	5	6	7	8	9	10	100
$\pi(n)$	1	2	3	5	7	11	15	22	30	42	190569292

Für große n gilt
$$\pi(n) \approx \tfrac{1}{4\sqrt{3}n} e^{\pi\sqrt{\frac{2n}{3}}}.$$
Das folgende Lemma zusammen mit Bemerkung 11.3 besagt, daß $\pi(n)$ zugleich die Anzahl der Konjugationsklassen nilpotenter Matrizen der Größe $n \times n$ angibt.

Lemma 11.11 (Jordansche Normalform nilpotenter Endomorphismen)
Sei $f \in \mathrm{End}_K(V)$ *ein nilpotenter Endomorphismus mit* $m_f = t^m$.

a. *Setzen wir* $U_i = \mathrm{Ker}(f^i)$, $i = 0, \ldots, m$, *dann induziert* f *für* $i = 2, \ldots, m$ *eine injektive lineare Abbildung*
$$f_i : U_i/U_{i-1} \longrightarrow U_{i-1}/U_{i-2} : \overline{x} \mapsto \overline{f(x)}.$$
Zudem ist $P = (k_1, \ldots, k_m)$ *eine Partition von* n *mit*
$$k_i = \dim_K(U_i/U_{i-1}) = \mathrm{rang}(f^{i-1}) - \mathrm{rang}(f^i),$$
die wir die Jordan-Partition *des nilpotenten Endomorphismus nennen wollen.*

b. *Ist* $P^* = (l_1, \ldots, l_s)$ *die zu* P *duale Partition, dann gibt es eine Basis* B *von* V, *so daß*
$$M_B^B(f) = J_{l_1}(0) \oplus J_{l_2}(0) \oplus \ldots \oplus J_{l_s}(0).$$

Die analogen Aussagen für Matrizen $A \in \mathrm{Mat}(n \times n, K)$ *gelten ebenfalls.*

Beweis: Wir beweisen zunächst Teil a. und beachten dazu, daß wir aus Lemma 10.17
$$U_0 \subsetneqq U_1 \subsetneqq \ldots \subsetneqq U_m$$
wissen.

Wir müssen zunächst zeigen, daß f_i wohldefiniert ist. Für $\overline{x} = \overline{y} \in U_i/U_{i-1}$ ist
$$x - y \in U_{i-1} = \mathrm{Ker}(f^{i-1}),$$
so daß
$$f(x) - f(y) = f(x - y) \in \mathrm{Ker}(f^{i-2}) = U_{i-2}$$
folgt, d.h. $\overline{f(x)} = \overline{f(y)} \in U_{i-1}/U_{i-2}$. Da mit $x \in U_i = \mathrm{Ker}(f^i)$ zudem $f(x) \in \mathrm{Ker}(f^{i-1}) = U_{i-1}$ gilt, ist f_i wohldefiniert.

Mit f ist dann aber auch f_i eine lineare Abbildung und für die Injektivität reicht es, zu zeigen

$$\operatorname{Ker}(f_i) = \{\overline{0}\}.$$

Nun ist aber $\overline{x} \in \operatorname{Ker}(f_i)$ gleichwertig zu $f(x) \in U_{i-2} = \operatorname{Ker}(f^{i-2})$, was wiederum nur für $x \in \operatorname{Ker}(f^{i-1}) = U_{i-1}$, d.h. für $\overline{x} = \overline{0} \in U_i/U_{i-1}$, zutrifft.

Wir müssen noch zeigen, daß $P = (k_1, \ldots, k_m)$ eine Partition von n ist. Aus

$$0 \neq U_m/U_{m-1} \hookrightarrow U_{m-1}/U_{m-2} \hookrightarrow \ldots \hookrightarrow U_1/U_0 = U_1$$

folgt auch

$$1 \leq \dim_K(U_m/U_{m-1}) \leq \dim_K(U_{m-1}/U_{m-2}) \leq \ldots \leq \dim_K(U_1/U_0),$$

d.h.

$$1 \leq k_m \leq k_{m-1} \leq \ldots \leq k_1.$$

Man beachte dabei, daß $U_m/U_{m-1} \neq 0$ gilt, weil m der Nilpotenzindex von f ist!

Außerdem folgt aus der Dimensionsformel für Vektorräume

$$n = \dim_K(V) = \dim_K(V/U_{m-1}) + \dim_K(U_{m-1}) = \dim_K(U_m/U_{m-1}) + \dim_K(U_{m-1}).$$

Mit Induktion nach m folgt dann

$$\begin{aligned} n =& \dim_K(U_m/U_{m-1}) + \dim_K(U_{m-1}) \\ =& \dim_K(U_m/U_{m-1}) + \dim_K(U_{m-1}/U_{m-2}) + \ldots + \dim_K(U_1/U_0) + \dim_K(U_0) \\ =& k_m + k_{m-1} + \ldots + k_1 + 0. \end{aligned}$$

Damit ist P eine Partition von n und Teil a. ist bewiesen.

Wenden wir uns nun Teil b. zu und konstruieren die Basis B.

Dazu wählen wir zunächst Vektoren

$$x_1^m, \ldots, x_{k_m}^m, \tag{35}$$

deren Restklassen eine Basis von U_m/U_{m-1} bilden. Dann wenden wir f auf diese an und erhalten Vektoren

$$x_1^{m-1} := f(x_1^m), \ldots, x_{k_m}^{m-1} := f(x_{k_m}^m) \in U_{m-1},$$

deren Restklassen in U_{m-1}/U_{m-2} linear unabhängig sind, weil die Abbildung

$$f_m : U_m/U_{m-1} \hookrightarrow U_{m-1}/U_{m-2}$$

eine injektive lineare Abbildung ist. Nun ergänzen wir die Restklassen von diesen durch die Restklassen von Vektoren

$$x_{k_m+1}^{m-1}, \ldots, x_{k_{m-1}}^{m-1},$$

zu einer Basis von U_{m-1}/U_{m-2}. Mit den so gewonnenen Vektoren

$$x_1^{m-1}, \ldots, x_{k_{m-1}}^{m-1}$$

verfahren wir analog und konstruieren so rekursiv Vektoren

$$x_1^i, \ldots, x_{k_i}^i, \tag{36}$$

deren Restklassen jeweils eine Basis für U_i/U_{i-1} sind, für $i = 1, \ldots, m$. Diese $n = k_1 + \ldots + k_m$ Vektoren ordnen wir der Übersichtlichkeit halber in dem Young-Diagramm von P an (siehe Abbildung 4). Wir können die Vektoren im Young-

$$\begin{array}{cccccccccc}
x_1^1 & x_2^1 & ---- & x_{k_m}^1 & x_{k_m+1}^1 & ---- & x_{k_{m-1}}^1 & ---------------- & x_{k_1}^1 \\
\vdots & \vdots & & \vdots & \vdots & & & \vdots & \\
x_1^{m-2} & x_2^{m-2} & ---- & x_{k_m}^{m-2} & x_{k_m+1}^{m-2} & ---- & x_{k_{m-1}}^{m-2} & & \\
x_1^{m-1} & x_2^{m-1} & ---- & x_{k_m}^{m-1} & x_{k_m+1}^{m-1} & ---- & x_{k_{m-1}}^{m-1} & & \\
x_1^{m} & x_2^{m} & ---- & x_{k_m}^{m} & & & & &
\end{array}$$

ABBILDUNG 4. Anordnung der Basis B im Young-Diagramm zu P

Diagramm auch als Bilder unter der Abbildung f schreiben und erhalten Abbildung 5. Schließlich benennen wir die Vektoren um, wie in Abbildung 6 angegeben,

$$\begin{array}{cccccc}
f^{m-1}(x_1^m) & -- & f^{m-1}(x_{k_m}^m) & f^{m-2}(x_{k_m+1}^{m-1}) & -------------- & x_{k_1}^1 \\
\vdots & & \vdots & \vdots & \vdots & \\
f^2(x_1^m) & -- & f^2(x_{k_m}^m) & f(x_{k_m+1}^{m-1}) -- & & \\
f(x_1^m) & -- & f(x_{k_m}^m) & x_{k_m+1}^{m-1} \quad -- & & \\
x_1^m & -- & x_{k_m}^m & & &
\end{array}$$

ABBILDUNG 5. Anordnung der Basis B im Young-Diagramm zu P

d.h. wir lesen das Diagramm aus, indem wir, in der linken oberen Ecke beginnend, die Spalten sukzessive von oben nach unten durchlaufen. Wir müssen nun nur noch zeigen, daß

$$B = (x_1, \ldots, x_n)$$

linear unabhängig ist, dann ist B eine Basis des n-dimensionalen Vektorraums V und die Matrix-Darstellung hat offenbar die gewünschte Gestalt, wie wir aus Abbildung 5

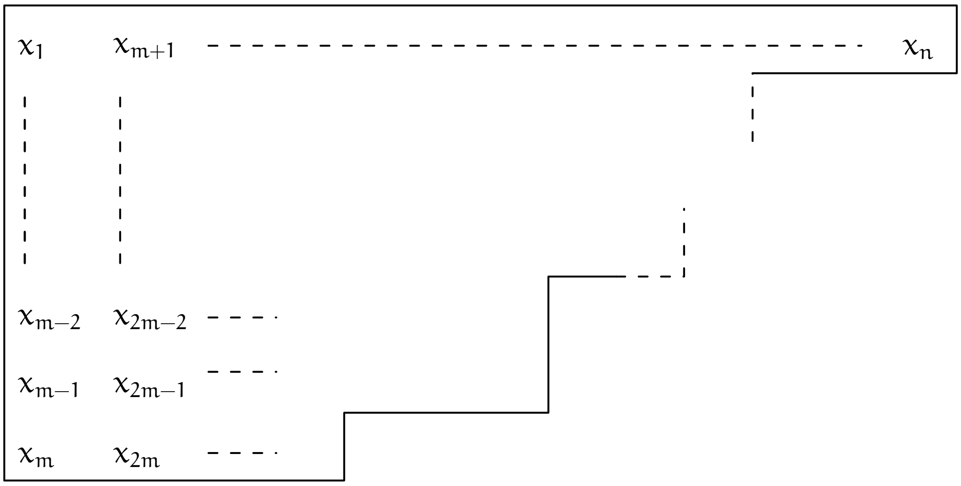


Abbildung 6. Anordnung der Basis B im Young-Diagramm zu P

sehen. Dazu beachten wir, daß die Spalten des Diagramms jeweils die kanonische Basis eines zyklischen Unterraums sind und somit einen Jordanblock liefern. Das zeigt insbesondere, daß die zu P duale Partition die Größen der Jordanblöcke liefert.

Um zu zeigen, daß B eine Basis ist, betrachten wir eine Linearkombination

$$\sum_{i=1}^{m} \sum_{j=1}^{k_j} \lambda_{i,j} \cdot x_j^i = 0 \tag{37}$$

der Vektoren in B, die den Nullvektor ergibt. Beachten wir, daß $x_j^i \in U_{m-1}$ für $i < m$ gilt, und betrachten wir die Gleichung in U_m/U_{m-1}, so reduziert sie sich auf

$$\sum_{j=1}^{k_m} \lambda_{m,j} \cdot \overline{x_j^m} = \overline{0} \in U_m/U_{m-1}.$$

Da die Vektoren $\overline{x_1^m}, \ldots, \overline{x_{k_m}^m}$ linear unabhängig sind (siehe (35)), gilt also

$$\lambda_{m,1} = \ldots = \lambda_{m,k_m} = 0$$

und (37) reduziert sich zu

$$\sum_{i=1}^{m-1} \sum_{j=1}^{k_j} \lambda_{i,j} \cdot x_j^i = 0.$$

Diese Gleichung können wir mit demselben Argument in U_{m-1}/U_{m-2} betrachten und erhalten

$$\sum_{j=1}^{k_{m-1}} \lambda_{m-1,i} \cdot \overline{x_j^{m-1}} = \overline{0} \in U_{m-1}/U_{m-2}.$$

Die Restklassen der beteiligten Vektoren sind nach Konstruktion (siehe (36)) linear unabhängig in U_{m-1}/U_{m-2} und somit gilt

$$\lambda_{m-1,1} = \ldots = \lambda_{m-1,k_{m-1}} = 0.$$

Fahren wir so fort erhalten wir insgesamt, daß alle $\lambda_{i,j}$ Null sein müssen und B ist linear unabhängig. $\square$

Wir können damit nun auch Satz 11.1 für nilpotente Endomorphismen beweisen.

Lemma 11.12 (Jordansche Normalform nilpotenter Endomorphismen)
Es sei $f \in \mathrm{End}_K(V)$ *ein nilpotenter Endomorphismus mit* $\mu_f = t^m$*. Dann gibt es für jedes* $1 \leq j \leq m$ *je eine natürliche Zahl* t_j *und es gibt eine Basis* B *so, daß*

(1) $\sum_{j=1}^{m} j \cdot t_j = n = \dim_K \mathrm{Hau}(f, 0) = \dim_K(V)$,

(2) $\sum_{j=1}^{m} t_j = \dim_K \mathrm{Eig}(f, 0)$,

(3) $t_m \geq 1$ *und*

$$J_f := M_B^B(f) = \bigoplus_{j=1}^{m} \bigoplus_{k=1}^{t_j} J_j(0).$$

Beweis: Sei die Partition $P = (k_1, \ldots, k_m)$ wie in Lemma 11.11 gegeben. Setzen wir

$$t_j := k_j - k_{j+1}$$

für $j = 1, \ldots, m$ mit der Konvention $k_{m+1} = 0$, dann ist t_j gerade die Anzahl der Jordanblöcke der Größe $j \times j$ in der Matrixdarstellung

$$M_B^B(f) = J_{l_1}(0) \oplus \ldots \oplus J_{l_s}(0) \tag{38}$$

in Lemma 11.11 (siehe Abbildung 4). Mithin gilt

$$\sum_{j=1}^{m} t_j = k_1 - k_{m+1} = k_1 = \dim_K(U_1) = \dim_K \mathrm{Ker}(f) = \dim_K \mathrm{Eig}(f, 0)$$

und

$$\sum_{j=1}^{m} j \cdot t_j = n = \dim_K(V) = \dim_K \mathrm{Hau}(f, 0),$$

weil die Summe der Größen der Kästchen mit ihren Vielfachheiten die Größe der Gesamtmatrix ist. Außerdem ist

$$t_m = k_m - k_{m+1} = k_m - 0 = k_m \geq 1$$

und die Matrixdarstellung in (38) kann dann auch geschrieben werden als

$$M_B^B(f) = \bigoplus_{j=1}^{m} \bigoplus_{k=1}^{t_j} J_j(0).$$

□

Bemerkung 11.13 (Jordanbasis einer nilpotenten Matrix)
Ist A eine nilpotente Matrix mit $\mu_A = t^m$ und bestimmt man wie im Beweis von Lemma 11.11 (siehe auch Abbildung 5) linear unabhängige Familien

$$B_{j,l} = \left(A^{j-1}x_l^j, A^{j-2}x_l^j, \ldots, Ax_l^j, x_l^j\right) \subset K^n$$

in Lös$(A^j, 0)$ für $j = 1, \ldots, m$ und $l = k_{j+1} + 1, \ldots, k_j$, dann ist die Matrix $T \in \mathrm{Gl}_n(K)$, deren Spalten gerade all diese Vektoren sind, eine Transformationsmatrix, die A in Jordansche Normalform überführt.

Beispiel 11.14 (Jordansche Normalform einer nilpotenten Matrix)
Wir wollen nun für die folgende nilpotente Matrix

$$A = \begin{pmatrix} -2 & 0 & 0 & 1 & 0 \\ -2 & -1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ -4 & 0 & -1 & 2 & 0 \\ -2 & -1 & 0 & 1 & 1 \end{pmatrix} \in \mathrm{Mat}_5(\mathbb{Q})$$

die Jordansche Normalform sowie die Transformationsmatrix $T \in \mathrm{Gl}_5(\mathbb{Q})$ bestimmen.

Dazu berechnen wir zunächst den Nilpotenzindex von A und merken uns die Potenzen A^k von A, da wir sie anschließend benötigen:

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

und

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

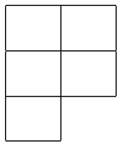
Der Nilpotenzindex von A ist also $m = 3$, $\mu_A = t^3$ und

$$\mathrm{Hau}(A, 0) = \mathbb{Q}^5.$$

Damit muß in der Jordanschen Normalform von A also ein Jordanblock $J_3(0)$ der Größe $m = 3$ vorkommen, und aufgrund der geringen Größe der Matrix A bleiben damit nur die beiden folgenden Möglichkeiten für die Jordansche Normalform übrig:

$$J_A = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \quad \text{oder} \quad J_A = \left(\begin{array}{ccc|c|c} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \cline{4-5} 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Bestimmen wir die Ränge der Potenzen $A^0, \ldots, A^3$ von A (siehe weiter unten), so können wir die Jordan-Partition



von A dann als

$$\begin{aligned} P =& \big(\operatorname{rang}(A^0) - \operatorname{rang}(A^1), \operatorname{rang}(A^1) - \operatorname{rang}(A^2), \operatorname{rang}(A^2) - \operatorname{rang}(A^3)\big) \\ =& (5-3, 3-1, 1-0) = (2,2,1) \end{aligned}$$

berechnen, und wir erhalten als duale Partition

$$P^* = (3,2),$$

woraus sich unmittelbar die Jordansche Normalform

$$J_A = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

ergibt sowie die Elementarteiler

$$t_1 = 0, \quad t_2 = 1, \quad t_3 = 1.$$

Um die Jordanbasis B von f_A oder alternativ die Transformationsmatrix T von A zu bestimmen, reicht es gemäß Lemma 11.11 im wesentlichen, geeignete Basen der Vektorräume U_j/U_{j-1} für $j = 3, 2, 1$ zu bestimmen, wobei $U_j = \text{Lös}(A^j, 0) = \operatorname{Ker}(f_A^j)$ ist.

Wir beginnen mit $j = 3$ und $U_3/U_2 = \mathbb{Q}^5/U_2$, wobei wir $U_3 = \mathbb{Q}^5$ beachten.

Um eine Basis von U_3 zu finden, muß man einerseits eine Basis von U_2 berechnen und diese dann zu einer Basis von U_3 ergänzen, indem man sie mit Steinitz in eine Basis von U_3 hineintauscht. Bestimmen wir also zunächst eine Basis B_2' von $U_2 = \text{Lös}(A^2, 0)$. Aufgrund der einfachen Form von A^2 mit Rang 1 geschieht dies durch einfaches Draufschauen — vier der Einheitsvektoren tun es offenbar:

$$B_2' = \big((1,0,0,0,0)^t, (0,1,0,0,0)^t, (0,0,0,1,0)^t, (0,0,0,0,1)^t\big).$$

Für $U_3 = \text{Lös}(A^3, 0) = \mathbb{Q}^5$ ist es noch einfacher, eine Basis zu bestimmen, die kanonische Basis tut's:

$$B_3' = \big((1,0,0,0,0)^t, (0,1,0,0,0)^t, (0,0,1,0,0)^t, (0,0,0,1,0)^t, (0,0,0,0,1)^t\big).$$

Damit ist es in dem vorliegenden Beispiel auch denkbar einfach, die Basis B_2' in die Basis B_3' hineinzutauschen, es fehlt nämlich einfach der Vektor e_3, und wir setzen deshalb

$$x_1^3 = e_3 = (0,0,1,0,0)^t.$$

Damit erhalten wir $k_3 = \dim_{\mathbb{Q}}(U_3/U_2) = 1$ und die erste Teilbasis der Jordanbasis:

$$B_{3,1} = \left(A^2x_1^3, Ax_1^3, x_1^3\right) = \left((-1,0,0,-2,0)^t, (0,-1,0,-1,0)^t, (0,0,1,0,0)^t\right).$$

Diese können wir in das Young-Diagramm der Jordan-Partition eintragen:

$A^2x_1^3$	
Ax_1^3	
x_1^3	

Als nächstes betrachten wir $j = 2$ und U_2/U_1.

Um eine geeignete Basis von U_2/U_1 zu bestimmen, müssen wir eine Basis von U_1 berechnen und diese zusammen mit der zweiten Ebene des bereits befüllten Young-Diagramms, d.h. mit der $(Ax_1^3) = (Ae_3)$, zu einer Basis von U_2 ergänzen. Dazu berechnen wir zunächst eine Basis B_1' von $U_1 = \text{Lös}(A, 0)$:

$$A = \begin{pmatrix} -2 & 0 & 0 & 1 & 0 \\ -2 & -1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ -4 & 0 & -1 & 2 & 0 \\ -2 & -1 & 0 & 1 & 1 \end{pmatrix} \xmapsto{\text{Gauß}} \begin{pmatrix} 1 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xmapsto[\text{einfügen}]{-1\text{'en}} \begin{pmatrix} 1 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}$$

Die letzten beiden Spaltenvektoren bilden also eine Basis von $\text{Lös}(A, 0)$. Wir dürfen die Vektoren aber auch mit einem Skalar multiplizieren, um schönere Vektoren zu erhalten, und tun dies. Unsere Basis von $U_1 = \text{Eig}(A, 0) = \text{Lös}(A, 0)$ ist dann

$$B_1' = \left((1,0,0,2,0)^t, (0,1,0,0,1)^t\right).$$

Wir müssen also die Familie

$$B_2'' = B_1' \cup (Ae_3) = \left((1,0,0,2,0)^t, (0,1,0,0,1)^t, (0,-1,0,-1,0)^t\right),$$

zu einer Basis von U_2 ergänzen. Dazu können wir sie mit Hilfe des Austauschsatzes von Steinitz in die Basis B_2' von U_2 hineintauschen, oder alternativ kann man auch einfach genau hinschauen. Man sieht nämlich leicht, daß der Vektor

$$x_2^2 = (0,0,0,0,1)^t$$

von den drei Vektoren in B_2'' linear unabhängig ist, und somit ergänzt er B_2'' zu einer Basis von U_2. Von der linearen Unabhängigkeit der vier Vektoren kann man sich auch überzeugen, indem man die Vektoren in eine Matrix schreibt und den Rang bestimmt, was schneller ist als dreimal Steinitz und trotzdem ausreicht. Wir überlassen die Rechnung dem Leser. Nachdem wir nun x_2^2 bestimmt haben, erhalten wir die zweite Teilbasis

$$B_{2,2} = \left(Ax_2^2, x_2^2\right) = \left((0,1,0,0,1)^t, (0,0,0,0,1)^t\right)$$

der Jordanbasis und das fertig ausgefüllte Young-Diagramm der Jordan-Partition:

$$\begin{array}{|c|c|}\hline A^2x_1^3 & Ax_2^2 \\ \hline Ax_1^3 & x_2^2 \\ \hline x_1^3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \quad = \quad \begin{array}{|c|c|}\hline x_1 & x_4 \\ \hline x_2 & x_5 \\ \hline x_3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}$$

Im Prinzip bliebe noch der Fall $j = 1$ zu untersuchen, aber da die zu P duale Partition nur zwei Spalten hat und damit $t_1 = 0$ gilt, sind wir fertig.

Wir haben also die Jordanbasis $B = B_{3,1} \cup B_{2,2}$ bestimmt und damit auch die Transformationsmatrix T, deren Spalten die Vektoren in B sind. Wir haben

$$T = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ -2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \in \mathrm{Gl}_5(\mathbb{Q})$$

mit

$$T^{-1} = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 2 & 1 & 0 & -1 & 0 \\ -2 & -1 & 0 & 1 & 1 \end{pmatrix},$$

und es gilt

$$J_A = T^{-1} \circ A \circ T = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Kommen wir nun zum Beweis von Satz 11.1.

Beweis von Satz 11.1: Nach Satz 10.18 zerfällt V in die direkte Summe der Haupträume $V_i := \mathrm{Hau}(f, \lambda_i)$, $i = 1, \ldots, r$, und diese sind nach Lemma 10.17 invariant unter f und $f - \lambda_i \,\mathrm{id}_V$. Betrachten wir nun die Abbildungen

$$(f - \lambda_i \,\mathrm{id}_V)_{V_i} : V_i \to V_i$$

für $i = 1, \ldots, r$, so sind diese nilpotent mit $\chi_{(f-\lambda_i \,\mathrm{id}_V)_{V_i}} = t^{n_i}$ und $\mu_{(f-\lambda_i \,\mathrm{id}_V)_{V_i}} = t^{m_i}$ (vgl. Korollar 10.19). Nach Lemma 11.12 gibt es dann aber für jedes $i = 1, \ldots, r$ Basen B_i von V_i und natürliche Zahlen t_{ij}, $j = 1, \ldots, m_i$, so daß gilt

(1) $\sum_{j=1}^{m_i} j \cdot t_{ij} = n_i = \dim_K \mathrm{Hau}(f, \lambda_i)$,

(2) $\sum_{j=1}^{m_i} t_{ij} = \dim_K \mathrm{Eig}\big((f - \lambda_i \,\mathrm{id}_V)_{V_i}, 0\big) = \dim_K \mathrm{Eig}(f, \lambda_i)$,

(3) $t_{im_i} \geq 1$ und

$$\begin{aligned} M_{B_i}^{B_i}(f_{V_i}) &= \lambda_i \mathbb{1}_{n_i} + M_{B_i}^{B_i}\big((f - \lambda_i \operatorname{id}_V)_{V_i}\big) \\ &= \lambda_i \mathbb{1}_{n_i} + \left(\bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(0) \right) = \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i). \end{aligned}$$

Damit folgt die Behauptung, da für $B = B_1 \cup \ldots \cup B_r$ gilt

$$M_B^B(f) = \bigoplus_{i=1}^{r} M_{B_i}^{B_i}(f_{V_i}).$$

□

Wie wir schon gesehen haben, ist der Beweis zur Berechnung der Jordanschen Normalform algorithmisch. Wir wollen nun den Algorithmus beschreiben, mit Hilfe dessen man die Jordansche Normalform einer Matrix A inklusive der zugehörigen Transformationsmatrix bestimmen kann.

Algorithmus 11.15 (Jordansche Normalform - I)

INPUT: $A \in \mathrm{Mat}_n(\mathbb{Q})$ mit μ_A zerfällt in Linearfaktoren.

OUTPUT: J_A und eine Transformationsmatrix $T \in \mathrm{Gl}_n(K)$ mit $T^{-1} \circ A \circ T = J_A$.

1. **Schritt:** Bestimme das Minimalpolynom μ_A von A und faktorisiere es.
2. **Schritt:** Wenn μ_A nicht in Linearfaktoren zerfällt, gebe man eine Fehlermeldung zurück, andernfalls gilt $\mu_A = \prod_{i=1}^{r}(t - \lambda_i)^{m_i}$.
3. **Schritt:** Für $i = 1, \ldots, r$ bilde man die Matrix $A_i = A - \lambda_i \mathbb{1}_n$ und führe folgende Schritte aus:
   - **Schritt a.:** Berechne die Partition $P = (k_1, \ldots, k_{m_i})$ von $n - \mathrm{rang}(A_i^{m_i})$ mit $k_j = \mathrm{rang}(A_i^{j-1}) - \mathrm{rang}(A_i^j)$ gemäß Lemma 11.11 sowie das zugehörige Young-Diagramm.
   - **Schritt b.:** Bestimme eine Basis B_{m_i} von $\mathrm{Lös}\big(A_i^{m_i}, 0\big)$ sowie eine Basis B_{m_i-1} von $\mathrm{Lös}\big(A_i^{m_i-1}, 0\big)$.
   - **Schritt c.:** Tausche B_{m_i-1} mittels des Satzes von Steinitz in B_{m_i} hinein und bestimme die in B_{m_i} verbliebenen Vektoren $x_1^{m_i}, \ldots, x_{k_{m_i}}^{m_i}$.
   - **Schritt d.:** Dann fülle man die ersten k_{m_i} Spalten des Young-Diagramms von P durch die Vektoren $A_i^{m_i-1} x_l^{m_i}, \ldots, A_i^0 x_l^{m_i}$ auf, $l = 1, \ldots, k_{m_i}$, wie in Abbildung 7.
   - **Schritt e.:** Für $j = m_i - 1, \ldots, 1$ führe man folgendes aus:
     - bestimme eine Basis B_{j-1} von $\mathrm{Lös}\big(A_i^{j-1}, 0\big)$;
     - tausche B_{j-1} sowie die auf der j-ten Ebene des Young-Diagramms bereits eingetragenen Vektoren mittels des Satzes von Steinitz in B_j hinein;
     - bestimme die in B_j verbliebenen Vektoren $x_{k_{j+1}+1}^j, \ldots, x_{k_j}^j$;
     - für $l = k_{j+1} + 1, \ldots, k_j$ fülle die Spalten des Young-Diagramms von P mit den Vektoren $A_i^{j-1} x_l^j, \ldots, A_i^0 x_l^j$.

Schritt f.: Füge die Vektoren aus dem Young-Diagramm als Spalten in die Matrix T ein, beginnend in der linken oberen Ecke und die Spalten des Young-Diagramms von oben nach unten nacheinander durchlaufend.

4. Schritt: Gib $T^{-1} \circ A \circ T$ und T zurück.

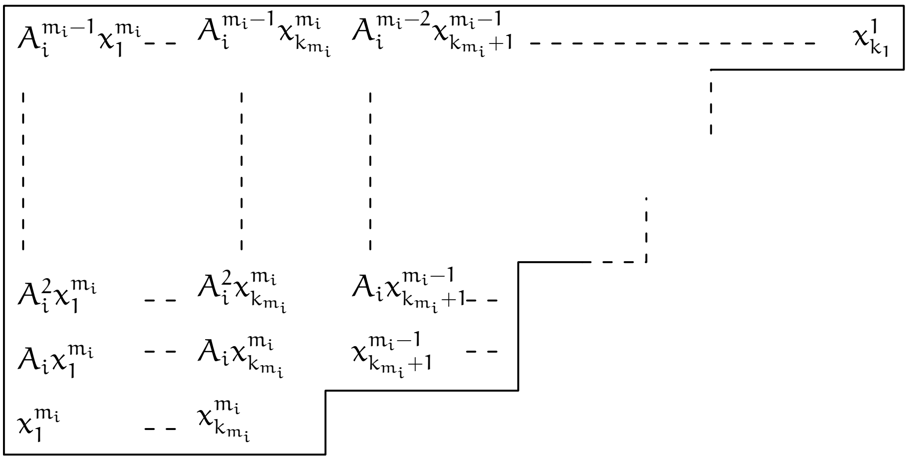


ABBILDUNG 7. Anordnung der Basis B_i im Young-Diagramm zu P

Beispiel 11.16 (Jordansche Normalform)
Wir wollen nun die Jordansche Normalform und die Transformationsmatrix von

$$A = \begin{pmatrix} 2 & 1 & 1 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Q})$$

bestimmen.

Das charakteristische Polynom berechnet man mit Hilfe des Kästchensatzes als

$$\chi_A = \begin{vmatrix} t-2 & -1 & -1 & -2 \\ 0 & t-1 & 0 & 0 \\ 0 & -1 & t-1 & -1 \\ 0 & 0 & 0 & t-2 \end{vmatrix} = (t-2) \cdot \begin{vmatrix} t-1 & 0 \\ -1 & t-1 \end{vmatrix} \cdot (t-2) = (t-2)^2 \cdot (t-1)^2.$$

Dann berechnen wir eine Basis von $\mathrm{Eig}(A, 1) = \mathrm{Lös}(A - \mathbb{1}_4, 0)$:

$$A - \mathbb{1}_4 = \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xmapsto{\text{Gauß}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xmapsto[\text{einfügen}]{-1\text{'en}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Also ist $\mathrm{Eig}(A, 1) = \mathrm{Lin}\left((1, 0, -1, 0)^t\right)$, und damit stimmen die geometrische und die algebraische Vielfachheit von 1 als Eigenwert von A nicht überein. Wir müssen auch noch $\mathrm{Hau}(A, 1) = \mathrm{Lös}\big((A - \mathbb{1}_4)^2, 0\big)$ bestimmen:

$$(A-\mathbb{1}_4)^2 = \begin{pmatrix} 1 & 2 & 1 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xmapsto{\text{Gauß}} \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xmapsto[\text{einfügen}]{-1\text{'en}} \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Als Basis für $\mathrm{Hau}(A, 1)$ erhalten wir also

$$B_1 = \big((2, -1, 0, 0)^t, (1, 0, -1, 0)^t\big).$$

Der erste der beiden Vektoren ist nicht in $\mathrm{Eig}(A, 1)$, so daß wir ihn als x_2 wählen können. Damit erhalten wir

$$x_1 = (A - \mathbb{1}_4)x_2 = (1, 0, -1, 0)^t, \quad x_2 = (2, -1, 0, 0)^t$$

als die ersten beiden Spalten von T.

Nun wenden wir uns der Berechnung von $\mathrm{Eig}(A, 2)$ zu:

$$A - 2\cdot\mathbb{1}_4 = \begin{pmatrix} 0 & 1 & 1 & 2 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xmapsto{\text{Gauß}} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xmapsto[\text{einfügen}]{-1\text{'en}} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Also ist $\mathrm{Eig}(A, 2) = \mathrm{Lin}\left((-1, 0, 0, 0)^t\right)$ und somit stimmen wieder die geometrische und die algebraische Vielfachheit des Eigenwertes nicht überein. Wir müssen also wieder $\mathrm{Hau}(A, 2) = \mathrm{Lös}\big((A - \mathbb{1}_4)^2, 0\big)$ berechnen:

$$(A-2\cdot\mathbb{1}_4)^2 = \begin{pmatrix} 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xmapsto{\text{Gauß}} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xmapsto[\text{einfügen}]{-1\text{'en}} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Wir erhalten als Basis für $\mathrm{Hau}(A, 2)$ also

$$B_2\big((-1, 0, 0, 0)^t, (0, 0, -1, -1)^t\big),$$

und somit ist $x_4 = (0, 0, 1, 1)^t$ im Hauptraum, aber nicht im Eigenraum von 2. Wir erhalten deshalb

$$x_3 = (A - 2\cdot\mathbb{1}_4)x_4 = (3, 0, 0, 0)^t, \quad x_4 = (0, 0, 1, 1)^t$$

als die Spalten 3 und 4 der Matrix T.

Insgesamt haben wir also

$$T = \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in \mathrm{Gl}_4(\mathbb{Q})$$

mit

$$T^{-1} = \begin{pmatrix} 0 & 0 & -1 & 1 \\ 0 & -1 & 0 & 0 \\ \frac{1}{3} & \frac{2}{3} & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

und für die Jordansche Normalform erhalten wir

$$T^{-1} \circ A \circ T = \left(\begin{array}{cc|cc} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{array}\right).$$

Will man nur die Normalform von A, aber nicht die Transformationsmatrix wissen, dann reicht es, die Elementarteiler zu bestimmen, was mit Hilfe von Aufgabe 11.20 sehr viel einfacher zu bewerkstelligen ist. Dies führt auf folgenden Algorithmus zur Bestimmung der Jordanschen Normalform einer Matrix A, deren charakteristisches Polynom zerfällt.

Algorithmus 11.17 (Jordansche Normalform - II)

INPUT: $A \in \mathrm{Mat}_n(\mathbb{Q})$ mit μ_A zerfällt in Linearfaktoren

OUTPUT: Liste mit den Eigenwerten von A und den Elementarteilern

1. **Schritt:** Bestimme das Minimalpolynom μ_A von A und faktorisiere es.
2. **Schritt:** Wenn μ_A nicht in Linearfaktoren zerfällt, gib eine Fehlermeldung zurück.
3. **Schritt:** Für jeden Eigenwert λ_i mit $\mathrm{mult}(\mu_A, \lambda_i) = m_i$ bestimme man für $j = 0, \ldots, m_i + 1$ die Zahlen $\mathrm{rang}\big((A - \lambda_i \mathbb{1}_n)^j\big)$ und berechne daraus den Vektor der Elementarteiler $(t_{i1}, \ldots, t_{im_i})$. Den Eigenwert und den Vektor der Elementarteiler speichere man als i-ten Eintrag in einer Liste `nf`.
4. **Schritt:** Man gebe die Liste `nf` zurück.

Bemerkung 11.18 (Jordanzerlegung einer Matrix)

Es sei $J = (a_{ij})$ eine Matrix in Jordanscher Normalform. $S = (s_{ij})$ bezeichne die Diagonalmatrix, die entsteht, wenn man in J alle Nicht-Diagonalelemente zu Null setzt, d. h. $s_{ii} = a_{ii}$ und $s_{ij} = 0$ für $i \neq j$. Ferner setzen wir $N = J - S$, d. h. N ist eine Matrix, die nur auf der oberen Nebendiagonalen Elemente ungleich Null besitzen kann.

Dann ist N *nilpotent*, und es gelten

$$J = S + N \quad \text{mit} \quad N \circ S = S \circ N.$$

Man nennt dies auch die *Jordan-Zerlegung* von J.

Um die Aussage einzusehen, beachte man, daß für $i = 1, \ldots, r$ und $1 \leq j \leq m_i$ gilt

$$J_j(\lambda_i) = \lambda_i \mathbb{1}_j + J_j(0).$$

Damit gilt

$$S = \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} \lambda_i \mathbb{1}_j$$

und

$$N = \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(0).$$

Aber damit folgt unmittelbar

$$N \circ S = \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} \lambda_i J_j(0) = S \circ N.$$

Allgemeiner nennt man die Darstellung einer Matrix $A \in \mathrm{Mat}_n(K)$ als $A = S + N$ mit N nilpotent und S diagonalisierbar (auch *halbeinfach* genannt, engl. semi-simple, daher das S) und $S \circ N = N \circ S$ eine Jordan-Zerlegung von A. Solche Zerlegungen von Objekten in einen halbeinfachen und einen nilpotenten Anteil spielen auch in anderen Bereichen der Mathematik eine Rolle - siehe etwa Lie-Algebren oder Jordan-Algebren.

Bemerkung 11.19 (Anwendungsmöglichkeit der Jordanschen Normalform)
Anwendung findet die Jordansche Normalform zum Beispiel in der Theorie der linearen Differentialgleichungen, wo ein Fundamentalsystem mit Hilfe der Exponentialabbildung einer Matrix beschrieben wird. Diese kann mit Hilfe der Jordanschen Normalform von A berechnet werden.

Aufgaben

Aufgabe 11.20 (Berechnung der Elementarteiler)
Mit den Bezeichnungen aus Satz 11.1 zeige man, für $i = 1, \ldots, r$ und $1 \leq j \leq m_i$ gilt:

$$t_{ij} = \mathrm{rang}\left((f - \lambda_i \,\mathrm{id}_V)^{j-1}\right) - 2 \cdot \mathrm{rang}\left((f - \lambda_i \,\mathrm{id}_V)^{j}\right) + \mathrm{rang}\left((f - \lambda_i \,\mathrm{id}_V)^{j+1}\right)$$

bzw.

$$t_{ij} = \mathrm{rang}\left((A - \lambda_i \mathbb{1}_n)^{j-1}\right) - 2 \cdot \mathrm{rang}\left((A - \lambda_i \mathbb{1}_n)^{j}\right) + \mathrm{rang}\left((A - \lambda_i \mathbb{1}_n)^{j+1}\right).$$

Hinweise: 1. Zeige, $J_j(0)^l = \left(\delta_{\mu+l,\nu}\right)_{\mu,\nu=1,\ldots,j}$ und $\mathrm{rang}\left(J_j(0)^l\right) = \max\{0, j-l\}$ für $l \in \mathbb{N}$. 2. Man betrachte zunächst den Fall $r = 1$ und $\lambda_1 = 0$. 3. Den allgemeinen Fall führe man auf die Abbildungen $g_i := (f - \lambda_i \,\mathrm{id}_V)_{\mathrm{Hau}(f,\lambda_i)}$ zurück.

Aufgabe 11.21
Bestimme die Jordansche Normalform und die zugehörige Transformationsmatrix

für die Matrix A

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 & 0 \\ -2 & 0 & 1 & 0 & 0 \\ 0 & -4 & 0 & 0 & -2 \end{pmatrix} \in \mathrm{Mat}_5(\mathbb{Q}).$$

Aufgabe 11.22

Bestimme die Jordansche Normalform und die zugehörige Transformationsmatrix $T^{-1} \in \mathrm{Gl}_4(\mathbb{Q})$ für

$$A = \begin{pmatrix} 3 & 6 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 9 & 3 & 0 \\ 0 & 12 & 0 & 3 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Q}).$$

Aufgabe 11.23

Es sei $A \in \mathrm{Mat}(5, K)$ mit $\chi_A = t(t-1)^4$, $\mu_A = t(t-1)^2$ und $\mathrm{rang}(A - \mathbb{1}_5) = 2$. Bestimme die Jordansche Normalform von A.

Aufgabe 11.24

Zeige, ist $A \in \mathrm{Mat}_n(K)$ so, daß χ_A über K in Linearfaktoren zerfällt, so sind A und A^t konjugiert.

Aufgabe 11.25

Beweise oder widerlege die folgende Aussage für zwei Matrizen $A, B \in \mathrm{Mat}_n(K)$:

$$A \text{ ist konjugiert zu } B \iff \chi_A = \chi_B,\ \mu_A = \mu_B \text{ und } \mathrm{rang}(A) = \mathrm{rang}(B).$$

KAPITEL III

Spektralsätze

Im folgenden sei stets $\mathbb{K}$ einer der beiden Körper $\mathbb{R}$ oder $\mathbb{C}$.

§ 12 Euklidische und unitäre Räume

Zur Motivation beginnen wir den Abschnitt mit einigen Überlegungen zur euklidischen Geometrie in der reellen Ebene $\mathbb{R}^2$.

Wir definieren uns zunächst zwei Abbildungen

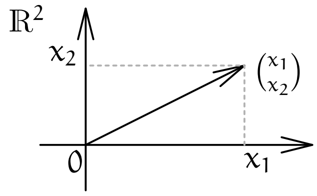
$$||\cdot|| : \mathbb{R}^2 \to \mathbb{R}_{\geq 0},$$

die einem Vektor $x = (x_1, x_2)^t \in \mathbb{R}^2$ seine Länge $||x||$ zuordnet, sowie

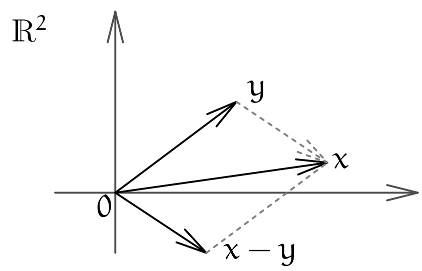
$$d : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}_{\geq 0},$$

die zwei Punkten $x \in \mathbb{R}^2$ und $y \in \mathbb{R}^2$ ihren Abstand $d(x, y)$ zuweist.

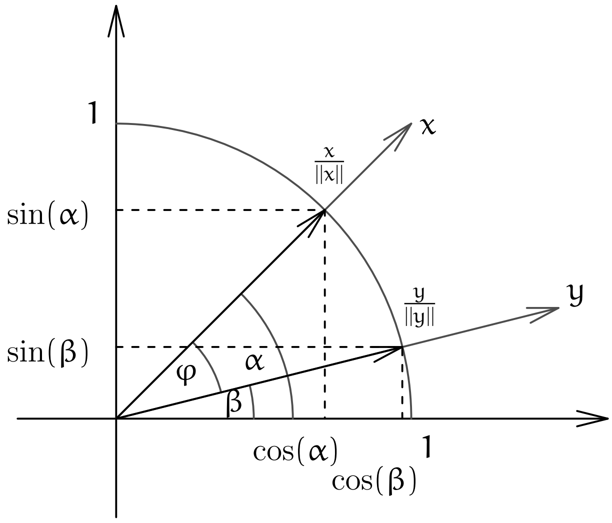
Der Satz von Pythagoras liefert dann $||x|| = \sqrt{x_1^2 + x_2^2}$.



Wir nennen $||x||$ auch die *Norm* des Vektors x. Da der Abstand der Punkte $x = (x_1, x_2)^t$ und $y = (y_1, y_2)^t$ gerade die Länge des Vektors $x - y$ ist, folgt somit $d(x, y) = ||x - y|| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$.



Mit Hilfe der Norm können wir - nach einigen geometrischen Überlegungen - auch den Winkel $\measuredangle(x, y)$, den zwei Vektoren x und y miteinander einschließen, bestimmen.



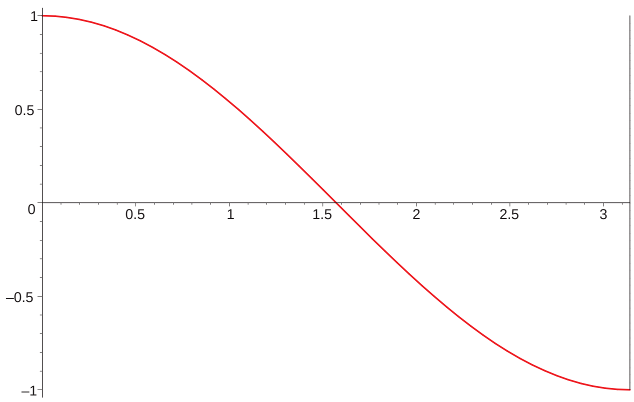
Dazu gehen wir zunächst zu den normierten Vektoren $\frac{x}{||x||}$ und $\frac{y}{||y||}$ über, die beide die Länge eins haben, wobei wir $x \neq 0$ und $y \neq 0$ voraussetzen. Mit den Bezeichnungen in der Skizze gilt dann

$$\measuredangle(x, y) = \measuredangle\left(\frac{x}{||x||}, \frac{y}{||y||}\right) = \alpha - \beta = \varphi.$$

Um φ selbst (im Bogenmaß) auszudrücken, müßte man die Länge des Kreisbogens zwischen $\frac{x}{||x||}$ und $\frac{y}{||y||}$ messen, also einer gekrümmten Linie. Dazu greifen wir auf unsere Analysiskenntnisse zurück.

Zur anschaulichen Herleitung des Winkels φ mit $0 \leq \varphi \leq \pi$, benötigen wir nur, daß die Funktion

$$\cos : [0, \pi] \longrightarrow \mathbb{R} : \varphi \mapsto \cos(\varphi)$$



injektiv ist. Also reicht es, $\cos(\varphi)$ zu kennen, um den Winkel φ eindeutig beschrieben zu haben. Unter Zuhilfenahme der obigen Skizze und des Additionstheorems für den Cosinus erhalten wir

$$\cos(\varphi) = \cos(\alpha - \beta) = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta) = \frac{x_1y_1 + x_2y_2}{||x|| \cdot ||y||}.$$

Dies führt zur Definition einer weiteren Abbildung

$$\langle\cdot,\cdot\rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R} : (x,y) = \big((x_1,x_2)^t,(y_1,y_2)^t\big) \mapsto \langle x,y\rangle := x_1y_1 + x_2y_2,$$

welche wir *Skalarprodukt* nennen. Mit deren Hilfe erhalten wir

$$\cos(\varphi) = \frac{\langle x,y\rangle}{||x|| \cdot ||y||}$$

oder alternativ

$$\measuredangle(x,y) = \varphi = \arccos\left(\frac{\langle x,y\rangle}{||x|| \cdot ||y||}\right).$$

Wir sind also mittels recht einfacher Abbildungen in der Lage, Längen und Winkel auszudrücken. Dieses Beispiel motiviert die folgenden Begriffsbildungen.

A) Skalarprodukte

Definition 12.1 (Skalarprodukt)
Es sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung

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

heißt ein *Skalarprodukt* auf V, falls für alle $\lambda,\mu \in \mathbb{K}$ und $x,y,z \in V$ gilt:

(1) $\langle x, \lambda y + \mu z\rangle = \lambda \cdot \langle x,y\rangle + \mu \cdot \langle x,z\rangle$,

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

(3) $\langle x,x\rangle \in \mathbb{R}$ und $\langle x,x\rangle > 0$ für $x \neq 0$.

Ist $\mathbb{K} = \mathbb{R}$, so nennen wir das Quadrupel $(V,+,\cdot,\langle\cdot,\cdot\rangle)$ einen *euklidischen Raum.*
Ist $\mathbb{K} = \mathbb{C}$, so nennen wir das Quadrupel $(V,+,\cdot,\langle\cdot,\cdot\rangle)$ einen *unitären Raum.*
Wir werden meist nur V statt $(V,+,\cdot,\langle\cdot,\cdot\rangle)$ schreiben.

Bemerkung 12.2 (Skalarprodukte als Bilinear- bzw. Sesquilinearformen)
Es sei V ein $\mathbb{K}$-Vektorraum.

a. Ist $\mathbb{K} = \mathbb{R}$, so ist $\langle\cdot,\cdot\rangle$ genau dann ein Skalarprodukt, wenn $\langle\cdot,\cdot\rangle$ eine positiv-definite, symmetrische Bilinearform auf V ist (siehe Anhang C1).

b. Ist $\mathbb{K} = \mathbb{C}$, so ist $\langle\cdot,\cdot\rangle$ genau dann ein Skalarprodukt, wenn $\langle\cdot,\cdot\rangle$ eine positiv-definite, hermitesche Sesquilinearform auf V ist (siehe Anhang C1).

Beispiel 12.3 (Skalarprodukte)

a. Wir nennen

$$\langle\cdot,\cdot\rangle : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R} : (x,y) \mapsto x^t \circ y = \sum_{i=1}^{n} x_i \cdot y_i$$

das *kanonische Skalarprodukt* oder *Standardskalarprodukt* auf $\mathbb{R}^n$. $\big(\mathbb{R}^n,+,\cdot,\langle\cdot,\cdot\rangle\big)$ ist ein euklidischer Raum.

b. Analog definieren wir das *kanonische Skalarprodukt* auf $\mathbb{C}^n$ durch

$$\langle\cdot,\cdot\rangle : \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C} : (x,y) \mapsto \overline{x}^t \circ y = x^* \circ y = \sum_{i=1}^{n} \overline{x_i} \cdot y_i,$$

und $\big(\mathbb{C}^n, +, \cdot, \langle\cdot,\cdot\rangle\big)$ ist ein unitärer Raum.

c. Sei $V = \mathcal{C}([0,1],\mathbb{R})$ der $\mathbb{R}$-Vektorraum der auf dem Intervall $[0,1]$ stetigen Funktionen. Dann wird durch

$$\langle f, g\rangle := \int_0^1 f(x)g(x)dx \in \mathbb{R}$$

für $f, g \in V$ ein Skalarprodukt definiert (siehe Aufgabe 12.34) und $\big(\mathcal{C}([0,1],\mathbb{R}), +, \cdot, \langle\cdot,\cdot\rangle\big)$ ist ein euklidischer Raum.

Definition 12.4 (Symmetrische und hermitesche Matrizen)

a. Eine Matrix $A \in \mathrm{Mat}_n(\mathbb{R})$ heißt *symmetrisch*, falls $A = A^t$.

b. Für eine Matrix $A = (a_{ij})_{i,j} \in \mathrm{Mat}_n(\mathbb{C})$ heißt die Matrix

$$A^* = \overline{A}^t = \big(\overline{a_{ji}}\big)_{i,j} \in \mathrm{Mat}_n(\mathbb{C})$$

die *adjungierte Matrix* zu A.

c. Eine Matrix $A \in \mathrm{Mat}_n(\mathbb{C})$ heißt *hermitesch* oder *selbstadjungiert*, falls $A = A^*$.

Beachte, eine reelle Matrix ist genau dann symmetrisch, wenn sie hermitesch ist.

Bemerkung 12.5

Ist $A \in \mathrm{Mat}_n(\mathbb{K})$, so erfüllt die Abbildung

$$\langle\cdot,\cdot\rangle_A : \mathbb{K}^n \times \mathbb{K}^n \longrightarrow \mathbb{K} : (x,y) \mapsto x^* \circ A \circ y = \overline{x}^t \circ A \circ y$$

die Bedingnung (1) eines Skalarproduktes, weil die Matrixmultiplikation linear ist. Ist die Matrix A zudem hermitesch, so erfüllt $\langle\cdot,\cdot\rangle_A$ auch die Bedingung (2) eines Skalarproduktes, weil

$$\langle x, y\rangle_A = \overline{x}^t \circ A \circ y = (\overline{x}^t \circ A \circ y)^t = \overline{\overline{y}^t \circ \overline{A}^t \circ x} = \overline{\overline{y}^t \circ A \circ x} = \overline{\langle y, x\rangle_A}.$$

Betrachten wir konkret $\mathbb{K} = \mathbb{R}$ und die symmetrische Matrix

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R}),$$

dann gilt zudem für $x = (x_1, x_2)^t \in \mathbb{R}^2$ mit $x \neq (0,0)$

$$\langle x, x\rangle_A = x_1^2 + 2x_1x_2 + 4x_2^2 = (x_1 + x_2)^2 + 3x_2^2 > 0$$

und damit Bedinung (3), so daß $\langle\cdot,\cdot\rangle_A$ in dem Fall ein Skalarprodukt auf $\mathbb{R}^3$ ist.

Bemerkung 12.6

Wenn wir im Folgenden den Fall eines euklidischen und eines unitären Raumes parallel behandeln wollen, dann werden wir uns häufig zunutze machen, daß für eine reelle Zahl λ gilt $\lambda = \overline{\lambda}$. Mithin sind auf einem *reellen* Vektorraum V die Bedingungen

$$\langle \lambda x + \mu y, z\rangle = \lambda\langle x, z\rangle + \mu\langle y, z\rangle$$

oder

$$\langle \lambda x + \mu y, z\rangle = \overline{\lambda}\langle x, z\rangle + \overline{\mu}\langle y, z\rangle.$$

gleichwertig, und für eine reelle Matrix $A \in \mathrm{Mat}(n, \mathbb{R})$ gilt genau dann $A = A^t$, wenn $A = \overline{A}^t$ erfüllt ist. Dies erspart uns viele Fallunterscheidungen!

B) Die euklidische Norm zu einem Skalarprodukt

Definition 12.7 (Normierter Raum)
Es sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung

$$||\cdot|| : V \to \mathbb{R}_{\geq 0},$$

heißt eine *Norm* auf V, falls für alle $x, y \in V$ und $\lambda \in \mathbb{K}$

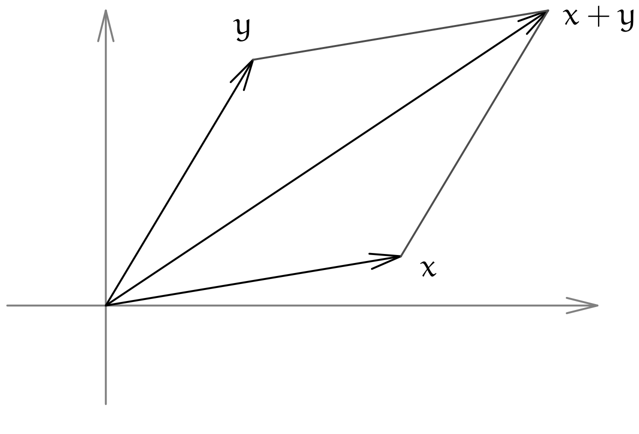
(1) $||x|| = 0 \quad \Leftrightarrow \quad x = 0$, ("Positive Definitheit")

(2) $||\lambda x|| = |\lambda| \cdot ||x||$, und ("Homogenität")

(3) $||x + y|| \leq ||x|| + ||y||$. ("Dreiecksungleichung")

Das Quadrupel $\big(V, +, \cdot, ||\cdot||\big)$ heißt dann ein *normierter Raum*.

Bemerkung 12.8 (Norm als Längenmaß)
Wir erinnern uns, daß eine Norm die Länge von Vektoren sowie Abstände messen soll. Bedingung (1) kann dann so interpretiert werden, daß jeder Vektor eine nicht-negative Länge hat und daß nur der Nullvektor die Länge null hat. Bedingung (2) bedeutet, daß die Streckung eines Vektors um den Faktor λ seine Länge um $|\lambda|$ strecken möge. Und Bedingung (3) kann dahingehend interpretiert werden, daß der Weg vom Ursprung über den Punkt x hin zum Punkt $x + y$ unter gar keinen Umständen kürzer ist, als der direkte Weg vom Ursprung zum Punkt $x + y$.



Diese Forderungen scheinen allesamt für eine Funktion, die die Länge von Vektoren beziehungsweise Abstände von Punkten messen soll, nicht unbillig. Und in der Tat reichen diese Forderungen auch bereits aus, um einen vernünftigen Längenbegriff zu erhalten.

Satz 12.9 (Cauchy-Schwarzsche Ungleichung)
Ist V *ein euklidischer oder unitärer Raum, dann gilt für alle* $x, y \in V$

$$\left|\langle x, y\rangle\right| \leq \sqrt{\langle x, x\rangle} \cdot \sqrt{\langle y, y\rangle}, \tag{39}$$

zudem gilt die Gleichheit genau dann, wenn x *und* y *linear abhängig sind.*

Beweis: Für $x = 0$ oder $y = 0$ ist die Aussage offensichtlich richtig. Wir können also $x, y \neq 0$ annehmen. Dann gilt für $\lambda \in \mathbb{K}$

$$0 \leq \langle x - \lambda y, x - \lambda y\rangle = \langle x, x\rangle - \overline{\lambda}\overline{\langle x, y\rangle} - \lambda\langle x, y\rangle + \overline{\lambda}\lambda\langle y, y\rangle. \tag{40}$$

Wählen wir nun speziell $\lambda = \frac{\overline{\langle x,y\rangle}}{\langle y,y\rangle} \in \mathbb{K}$, dann folgt

$$\begin{aligned} 0 &\leq \langle x, x\rangle - \tfrac{\langle x,y\rangle}{\langle y,y\rangle} \cdot \overline{\langle x, y\rangle} - \tfrac{\overline{\langle x,y\rangle}}{\langle y,y\rangle} \cdot \langle x, y\rangle + \tfrac{\overline{\langle x,y\rangle}}{\langle y,y\rangle} \tfrac{\langle x,y\rangle}{\langle y,y\rangle} \cdot \langle y, y\rangle \\ &= \langle x, x\rangle - \tfrac{|\langle x,y\rangle|^2}{\langle y,y\rangle} - \tfrac{|\langle x,y\rangle|^2}{\langle y,y\rangle} + \tfrac{|\langle x,y\rangle|^2}{\langle y,y\rangle}, \end{aligned}$$

also

$$\left|\langle x, y\rangle\right|^2 \leq \langle x, x\rangle \cdot \langle y, y\rangle. \tag{41}$$

Durch Ziehen der positiven Wurzel folgt die gesuchte Ungleichung (39).

Nun sind x und y genau dann linear abhängig, wenn es ein $\lambda \in \mathbb{K}$ gibt, für das $x = \lambda y$ gilt. Das wiederum ist wegen der positiven Definitheit von $\langle\cdot,\cdot\rangle$ gleichbedeutend dazu, daß es ein $\lambda \in \mathbb{K}$ gibt, für das in (40) das Gleichheitszeichen gilt. Dieses λ ist eindeutig bestimmt, und erfüllt

$$\lambda = \lambda \cdot \frac{\langle y, y\rangle}{\langle y, y\rangle} = \frac{\langle y, \lambda y\rangle}{\langle y, y\rangle} = \frac{\langle y, x\rangle}{\langle y, y\rangle} = \frac{\overline{\langle x, y\rangle}}{\langle y, y\rangle}.$$

Damit ist die Gleichheit in (40) gleichwertig zur Gleichheit in (41). □

Satz 12.10 (Die euklidische Norm zu einem Skalarprodukt)
Es sei V *ein euklidischer oder unitärer Raum. Dann wird durch*

$$\|\cdot\| : V \to \mathbb{R}_{\geq 0} : x \mapsto \sqrt{\langle x, x\rangle}$$

eine Norm auf V *definiert. Wir werden euklidische und unitäre Räume im folgenden stets mit dieser zugehörigen euklidischen Norm als normierte Räume betrachten, so daß die Cauchy-Schwarzsche Ungleichung die folgende Form hat:*

$$\left|\langle x, y\rangle\right| \leq \|x\| \cdot \|y\|.$$

Beweis: Seien $x \in V$ und $\lambda \in \mathbb{K}$. Aus der positiven Definitheit von $\langle\cdot,\cdot\rangle$ folgt, daß $\langle x, x\rangle \geq 0$ und somit $\|x\|$ definiert und stets nicht-negativ ist. Ferner folgt, daß $\|x\| = 0$ genau dann gilt, wenn x der Nullvektor ist. Aus der Bilinearität bzw. Sesquilinearität von $\langle\cdot,\cdot\rangle$ leiten wir her, daß

$$\langle\lambda x, \lambda x\rangle = \overline{\lambda}\lambda\langle x, x\rangle = |\lambda|^2\langle x, x\rangle,$$

und somit $||\lambda x|| = |\lambda| \cdot ||x||$.

Allein, die Dreiecksungleichung ist etwas schwieriger zu zeigen. Wir verwenden hierfür die Cauchy-Schwarzsche Ungleichung aus Satz 12.9. Beachten wir noch, daß für eine komplexe Zahl $c \in \mathbb{C}$ stets

$$\mathrm{Re}(c) = \tfrac{1}{2} \cdot \big(c + \overline{c}\big)$$

gilt, so erhalten wir für $x, y \in V$

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

Hieraus folgt dann die Dreiecksungleichung. □

Bemerkung 12.11 (Winkel in euklidischen Räumen)
Die Cauchy-Schwarzsche Ungleichung erlaubt es uns nun, in einem beliebigen *euklidischen Raum* V Winkel zwischen zwei Vektoren zu definieren. Denn aus der Ungleichung (39) folgt für $0 \neq x, y \in V$

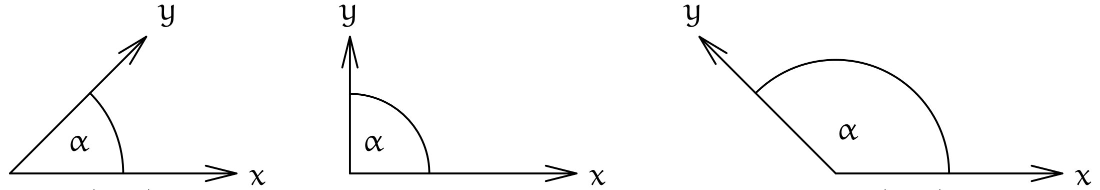
$$-1 \leq \frac{\langle x,y\rangle}{||x|| \cdot ||y||} \leq 1. \tag{42}$$

Vom Cosinus ist bekannt, daß es zu jeder reellen Zahl $-1 \leq r \leq 1$ genau einen Winkel $\alpha \in [0, \pi]$ gibt mit $r = \cos(\alpha)$, nämlich $\alpha = \arccos(r)$. Man definiert deshalb

$$\measuredangle(x,y) = \arccos\left(\frac{\langle x,y\rangle}{||x|| \cdot ||y||}\right) \in [0, \pi]$$

als den *Winkel* zwischen x und y.

Ist $\langle x,y\rangle > 0$, also $\measuredangle(x,y) \in \left[0, \frac{\pi}{2}\right[$, so spricht man von einem *spitzen Winkel*.
Ist $\langle x,y\rangle < 0$, also $\measuredangle(x,y) \in \left]\frac{\pi}{2}, \pi\right]$, so spricht man von einem *stumpfen Winkel*.
Ist $\langle x,y\rangle = 0$, also $\measuredangle(x,y) = \frac{\pi}{2}$, so spricht man von einem *rechten Winkel*.



$\langle x,y\rangle > 0$ "spitzer Winkel" | $\langle x,y\rangle = 0$ "rechter Winkel" | $\langle x,y\rangle < 0$ "stumpfer Winkel"

C) Orthonormalbasen und Parsevalsche Gleichung

Definition 12.12 (Orthogonal)
Sei V ein euklidischer oder unitärer Raum, $x, y \in V$, $M, N \subseteq V$ und $U \leq V$.

a. x heißt *orthogonal* zu y, falls $\langle x, y\rangle = 0$. Wir schreiben dann $x \perp y$.
b. M heißt *orthogonal* zu N, falls $m \perp n$ für alle $m \in M$ und $n \in N$. Wir schreiben dann $M \perp N$.
c. Wir nennen $U^\perp := \{z \in V \mid z \perp U\}$ das *orthogonale Komplement* von U.

Lemma 12.13 (Orthogonales Komplement)
Ist V ein euklidischer oder unitärer Raum und $U \leq V$, dann ist $U^\perp \leq V$.

Beweis: Wegen $0 \in U^\perp$ ist $U^\perp \neq \emptyset$. Sind $x, y \in U^\perp$ und $\lambda, \mu \in \mathbb{K}$, so gilt für $z \in U$

$$\langle \lambda x + \mu y, z\rangle = \overline{\lambda}\langle x, z\rangle + \overline{\mu}\langle y, z\rangle = 0,$$

Also $\lambda x + \mu y \in U^\perp$. Damit ist $U^\perp$ ein Unterraum von V. □

Definition 12.14 (Orthonormalbasis)
Sei V ein euklidischer oder unitärer Raum und $B = (x_i \mid i \in I)$ eine Familie in V.

a. B heißt *orthogonal*, falls $x_i \perp x_j$ für alle $i, j \in I$ mit $i \neq j$ gilt.
b. B heißt *orthonormal*, falls B orthogonal ist und $||x_i|| = 1$ für alle $i \in I$ gilt.
c. Ist B eine Basis und orthonormal, so heißt B eine *Orthonormalbasis*, kurz *ONB*.

Beispiel 12.15 (ONB)
Betrachten wir $\mathbb{K}^n$ mit dem kanonischen Skalarprodukt, dann ist die kanonische Basis $E = (e_1, \ldots, e_n)$ offenbar eine Orthonormalbasis von $\mathbb{K}^n$, da $\langle e_i, e_j\rangle = \delta_{ij}$ für $i, j \in \{1, \ldots, n\}$.

Lemma 12.16 (Orthogonal impliziert linear unabhängig.)
Sei V ein euklidischer oder unitärer Raum und $B = (x_i \mid i \in I)$ eine orthogonale Familie in $V \setminus \{0\}$. Dann ist B linear unabhängig.

Beweis: Aus $\sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i x_i = 0$ folgt für jedes $j \in I$

$$0 = \langle x_j, 0\rangle = \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i \langle x_j, x_i\rangle = \lambda_j \langle x_j, x_j\rangle.$$

Da $\langle x_j, x_j\rangle \neq 0$, muß also $\lambda_j = 0$ gelten. □

Proposition 12.17 (Parsevalsche Gleichung)
Sei V ein euklidischer oder unitärer Raum, $B = (x_i \mid i \in I)$ eine ONB und $x \in V$, dann gilt

$$x = \sum_{\substack{i\in I\\ \text{endlich}}} \langle x_i, x\rangle \cdot x_i. \tag{43}$$

Insbesondere sind nur endlich viele $\langle x_i, x\rangle$, $i \in I$, ungleich null.

Beweis: Da die Darstellung $x = \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i x_i$ von x als endliche Linearkombination von B eindeutig ist, folgt die Behauptung aus

$$\langle x_j, x\rangle = \left\langle x_j, \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i x_i \right\rangle = \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i \langle x_j, x_i\rangle = \lambda_j \langle x_j, x_j\rangle = \lambda_j.$$

□

Bemerkung 12.18

Ist B eine ONB von V, so erlaubt es die Gleichung (43), einen Vektor aus V als Linearkombination von B darzustellen, ohne hierzu eigens ein LGS lösen zu müssen, durch simples Einsetzen der Vektoren in das Skalarprodukt. Dieses Verfahren ist sehr effizient und von hoher praktischer Bedeutung. Die Tatsache, daß sich die Koordinaten eines Vektors bezüglich einer ONB mit Hilfe des Skalarproduktes so einfach ausdrücken lassen, spielt aber auch in vielen Beweisen eine Rolle, und ist somit ebenfalls für die Theorie von Bedeutung.

D) Das Orthonormalisierungsverfahren von Gram-Schmidt

Wir beweisen jetzt, daß jeder endlich-dimensionale euklidische bzw. unitäre Raum eine ONB besitzt. Etwas allgemeiner gilt der folgende Satz.

Satz 12.19 (Gram-Schmidt)

Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und B eine orthonormale Familie in V, dann läßt sich B zu einer ONB von V ergänzen.

Beweis: Ist B schon eine Basis von V, so sind wir fertig. Wir dürfen also annehmen, daß B keine Basis und wegen Lemma 12.16 dann auch kein Erzeugendensystem von V ist. Wir zeigen nun konstruktiv, wie wir die orthonormale Familie $B = (z_1, \ldots, z_r)$ zu einer orthonormalen Familie $(z_1, \ldots, z_{r+1})$ ergänzen können. Wenden wir dieses Verfahren dann $\dim_{\mathbb{K}}(V) - r$ mal an, so haben wir die Aussage bewiesen.

Dazu wählen wir zunächst einen Vektor x_{r+1}, der linear unabhängig von B ist. Dann setzen wir

$$y_{r+1} := x_{r+1} - \sum_{i=1}^{r} \langle z_i, x_{r+1}\rangle \cdot z_i. \tag{44}$$

Da x_{r+1} linear unabhängig von B ist, ist $y_{r+1} \neq 0$, und wir können deshalb

$$z_{r+1} := \tfrac{1}{||y_{r+1}||} \cdot y_{r+1} \tag{45}$$

setzen. Dann ist $||z_{r+1}|| = 1$ und außerdem gilt für $i = 1, \ldots, r$

$$\begin{aligned}
\langle z_i, z_{r+1}\rangle &= \frac{1}{||y_{r+1}||} \cdot \langle z_i, y_{r+1}\rangle \\
&= \frac{1}{||y_{r+1}||} \cdot \left(\langle z_i, x_{r+1}\rangle - \sum_{j=1}^{r} \langle z_j, x_{r+1}\rangle \cdot \langle z_i, z_j\rangle \right) \\
&= \frac{1}{||y_{r+1}||} \cdot \big(\langle z_i, x_{r+1}\rangle - \langle z_i, x_{r+1}\rangle \big) = 0.
\end{aligned}$$

Dann ist aber $(z_1, \ldots, z_{r+1})$ orthonormal und wir sind fertig. $\square$

Korollar 12.20 (Existenz einer ONB)
Jeder endlich-dimensionale euklidische bzw. unitäre Raum besitzt eine ONB.

Beweis: Wende Satz 12.19 mit $B = \emptyset$ an. $\square$

Der Beweis von Satz 12.19 ist konstruktiv und wird auch das *Gram-Schmidtsche Orthonormalisierungsverfahren* genannt. Es erlaubt, aus einem gegebenen Erzeugendensystem eine ONB zu konsturieren.

Algorithmus 12.21 (Gram-Schmidtsches Orthonormalisierungsverfahren)
INPUT: $M \subseteq \mathbb{K}^n$ und ein Skalarprodukt $\langle\cdot,\cdot\rangle$ auf $\mathbb{K}^n$

OUTPUT: ONB B von $\langle M\rangle$

1. **Schritt:** Bestimme eine Basis $B = (x_1, \ldots, x_r)$ von $\langle M\rangle$, z. B. mittels Algorithmus 6.20.
2. **Schritt:** Für $i = 1, \ldots, r$ führe man folgende Schritte aus:
   Schritt a.: berechne die Summe $y_i = x_i - \sum_{j=1}^{i-1} \langle z_j, x_i\rangle \cdot z_j$;
   Schritt b.: berechne $z_i = \frac{1}{\|y_i\|} \cdot y_i$;
3. **Schritt:** Gib die veränderte Basis $(z_1, \ldots, z_r)$ zurück.

Bemerkung 12.22

a. Will man in der Praxis ein Skalarprodukt übergeben, so wird man im reellen Fall eine symmetrische Matrix übergeben und im komplexen Fall eine hermitesche. Das Skalarprodukt wird dann gemäß Beispiel C1.2 bzw. Beispiel C1.29 gebildet.

b. Um zu normieren, ist in Algorithmus 12.21 das Ziehen von Wurzeln notwendig. Verzichtet man jedoch auf die Normierung der Vektoren, so kommt man ohne Wurzelziehen aus. Läßt man im obigen Algorithmus Schritt 2.b. weg und ersetzt dafür in Schritt 2.a. die rechte Seite der Gleichung durch

$$y_i = x_i - \sum_{j=1}^{i-1} \frac{\langle y_j, x_i\rangle}{\langle y_j, y_j\rangle} \cdot y_j,$$

dann liefert Algorithmus 12.21 eine orthogonale Basis $(y_1, \ldots, y_r)$ von $\langle M\rangle$. Das hat den Vorteil, daß man exakt rechnen kann - etwa in SINGULAR, wenn die Eingabedaten rationale Zahlen waren.

Beispiel 12.23 (Gram-Schmidtsches Orthonormalisierungsverfahren)
Es sei $B = (x_1, x_2, x_3) = \left\{(1,0,1)^t, (1,1,1)^t, (0,0,4)^t\right\} \subseteq \mathbb{R}^3$, wobei wir $\mathbb{R}^3$ mit dem kanonischen Skalarprodukt versehen betrachten. Man sieht leicht, daß B bereits eine Basis von $\mathbb{R}^3$ ist. Wir wollen hier B in eine ONB von $\mathbb{R}^3$ überführen.

Wir setzen nun $y_1 := (1,0,1)^t$, dann ist $\langle y_1, y_1\rangle = 2$ und somit ersetzen wir x_1 in B durch

$$z_1 = \frac{1}{||y_1||} \cdot y_1 = \frac{1}{\sqrt{2}} \cdot y_1 = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)^t.$$

Im nächsten Schritt setzen wir

$$y_2 = x_2 - \langle z_1, x_2\rangle \cdot z_1 = (1,1,1)^t - \frac{2}{\sqrt{2}} \cdot \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)^t = (0,1,0)^t.$$

Dann ist $\left\langle y_2, y_2\right\rangle = 1$ und somit ersetzen wir x_2 in B durch $z_2 = y_2$.

Schließlich bilden wir

$$\begin{aligned} y_3 &= x_3 - \langle z_1, x_3\rangle \cdot z_1 - \langle z_2, x_3\rangle \cdot z_2 \\ &= (0,0,4)^t - \frac{4}{\sqrt{2}} \cdot \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)^t - 0 \cdot (0,1,0)^t \\ &= (-2,0,2)^t, \end{aligned}$$

und erhalten $\langle y_3, y_3\rangle = 8$. Somit müssen wir x_3 durch den Vektor

$$z_3 = \frac{1}{||y_3||} \cdot y_3 = \frac{1}{\sqrt{8}} \cdot (-2,0,2)^t = \left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)^t$$

ersetzen. Damit ergibt sich aus dem Gram-Schmidtschen Orthonormalisierungsverfahren die ONB

$$\left(\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)^t, (0,1,0)^t, \left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)^t\right).$$

E) Orthogonale und unitäre Matrizen

Definition 12.24 (Orthogonale / unitäre Matrizen)

a. Eine Matrix $A \in \mathrm{Mat}_n(\mathbb{R})$ heißt *orthogonal*, wenn $A^t \circ A = \mathbb{1}_n$ gilt. Wir nennen $O(n) := \{A \in \mathrm{Mat}_n(\mathbb{R}) \mid A \text{ ist orthogonal}\}$ *orthogonale Gruppe* vom Grad n.
b. Eine Matrix $A \in \mathrm{Mat}_n(\mathbb{C})$ heißt *unitär*, wenn $A^* \circ A = \mathbb{1}_n$ gilt, und wir nennen $U(n) := \{A \in \mathrm{Mat}_n(\mathbb{C}) \mid A \text{ ist unitär}\}$ die *unitäre Gruppe* vom Grad n.

Proposition 12.25 (Die Determinante orthogonaler / unitärer Matrizen)
Ist $A \in \mathrm{Mat}_n(\mathbb{K})$ orthogonal oder unitär, so gilt $|\det(A)| = 1$.

Beweis: Wegen $A^* \circ A = \mathbb{1}_n$ gilt

$$\begin{aligned} 1 = \det(\mathbb{1}_n) &= \det(A^* \circ A) = \det(A^*) \cdot \det(A) \\ &= \overline{\det(A^t)} \cdot \det(A) = \overline{\det(A)} \cdot \det(A) = |\det(A)|^2. \end{aligned}$$

Daraus folgt $|\det(A)| = 1$. □

Proposition 12.26 (Orthogonale / unitäre Matrizen)
Für eine quadratische Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ sind äquivalent:

a. A *ist orthogonal bzw. unitär.*
b. A *ist invertierbar mit* $A^* = A^{-1}$.
c. *Die Spalten von A sind eine ONB von $\mathbb{K}^n$ mit kanonischem Skalarprodukt.*

d. *Die Zeilen von* A *sind eine ONB von* $\mathbb{K}^n$ *mit kanonischem Skalarprodukt.*

Beweis: Die Äquivalenz von a. und b. folgt unmittelbar aus der Definition, denn $A^* \circ A = \mathbb{1}_n$ heißt, daß A^* die Inverse von A ist.

Ist a^i der i-te Spaltenvektor von A, so ist $\overline{a^i}^t$ der i-te Zeilenvektor von A^* und deshalb ist

$$\langle a^i, a^j \rangle = \overline{a^i}^t \circ a^j$$

der Eintrag an der Stelle (i,j) von $A^* \circ A$. Deshalb sind die Spalten von A genau dann eine ONB von $\mathbb{K}^n$, wenn $A^* = A^{-1}$ die Inverse von A ist. Dies zeigt die Äquivalenz von b. und c..

Ist a_i der i-te Zeilenvektor von A, so ist $\overline{a_i}^t$ der i-te Spaltenvektor von A^*. Also ist

$$\overline{\langle a_i^t, a_j^t \rangle} = a_i \circ \overline{a_j}^t$$

der Eintrag von $A \circ A^*$ an der Stelle (i,j). Dies zeigt schließlich, daß b. und d. äquivalent sind. □

Beispiel 12.27 (Orthogonale Matrix)
Die Matrix

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R})$$

ist orthogonal, da ihre Spalten nach Beispiel 12.23 eine ONB von $\mathbb{R}^3$ bezüglich des kanonischen Skalarproduktes sind.

Korollar 12.28 (Die orthogonale und die unitäre Gruppe)
$(\mathrm{O}(n), \circ)$ *und* $(\mathrm{U}(n), \circ)$ *sind Gruppen.*

Beweis: Es reicht, zu zeigen, daß sie Untergruppen von $\mathrm{Gl}_n(\mathbb{C})$ sind. Offenbar sind $\mathrm{O}(n)$ und $\mathrm{U}(n)$ nicht-leere Teilmengen von $\mathrm{Gl}_n(\mathbb{C})$. Sind nun A und B in $\mathrm{O}(n)$ bzw. in $\mathrm{U}(n)$, so gilt

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

und

$$\left(A^{-1}\right)^* = A^{**} = A = \left(A^{-1}\right)^{-1}.$$

Mithin liegen auch $A \circ B$ und A^{-1} in $\mathrm{O}(n)$ bzw. in $\mathrm{U}(n)$. Damit ist gezeigt, daß $\mathrm{O}(n)$ und $\mathrm{U}(n)$ Untergruppen von $\mathrm{Gl}_n(\mathbb{C})$ sind. □

Bemerkung 12.29 (Die orthogonale Gruppe $\mathrm{O}(2)$)
Die Determinante

$$\det : \mathrm{O}(2) \longrightarrow \{1, -1\}$$

ist wegen des Determinantenmultiplikationssatzes ein Gruppenepimorphismus. Der Kern von det ist der Normalteiler

$$\mathrm{SO}(2) := \{A \in \mathrm{O}(2) \mid \det(A) = 1\}$$

von O(2) und wird die *spezielle orthogonale Gruppe* vom Grad 2 genannt. Wir werden unten zeigen, daß

$$\mathrm{SO}(2) = \{T(\alpha) \mid \alpha \in [0, 2\pi)\}$$

und

$$\mathrm{O}(2) \setminus \mathrm{SO}(2) = \{S(\alpha) \mid \alpha \in [0, 2\pi)\},$$

wobei

$$T(\alpha) := \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

eine Drehung um den Winkel α ist und

$$S(\alpha) := \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix}$$

eine Spiegelung an der Geraden $\mathrm{Lin}\left(\left(\cos\left(\frac{\alpha}{2}\right), \sin\left(\frac{\alpha}{2}\right)\right)^t\right)$. Insbesondere ist im Fall $n = 2$ also jede orthogonale Matrix eine Drehung oder eine Spieglung.

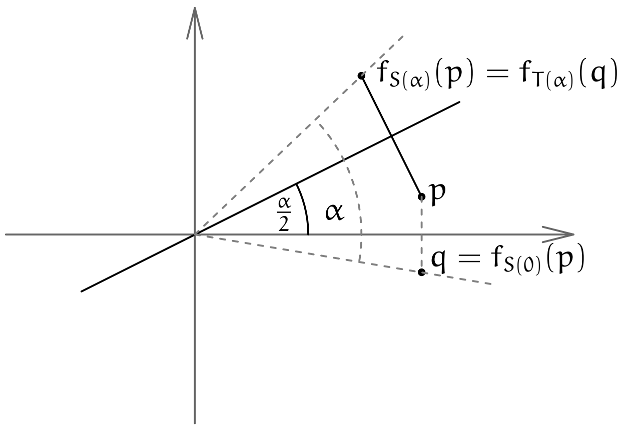


Abbildung 1. Die Spielung $S(\alpha) = T(\alpha) \circ S(0)$

Man beachte auch daß $S(\alpha) = T(\alpha) \circ S(0)$ d. h. die von $S(\alpha)$ induzierte Spiegelung ist Komposition der Spiegelung an der x-Achse gefolgt von einer Drehung um den Winkel α. Damit gilt zugleich, daß jede Drehung im $\mathbb{R}^2$ Komposition von zwei Spiegelungen ist.

Beweis: Die Matrix

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$$

ist genau dann orthogonal, wenn die beiden Spaltenvektoren $x = (a, b)^t$ und $y = (c, d)^t$ eine Orthonormalbasis von $\mathbb{R}^2$ bezüglich des Standardskalarproduktes sind. Insbesondere muß y also senkrecht auf x stehen. In der Ebene ist ein Vektor, der senkrecht steht auf x aber bis auf einen Skalarfaktor eindeutig bestimmt und $(-b, a)^t$ ist ein solcher Vektor. Es muß also

$$y = \lambda \cdot (-b, a)^t$$

gelten. Aus

$$1 = ||y|| = |\lambda| \cdot \sqrt{(-b)^2 + a^2} = |\lambda| \cdot ||x|| = |\lambda|$$

folgt dann $\lambda = 1$ oder $\lambda = -1$. Also ist die Matrix A von der Form

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad \text{oder} \quad A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix},$$

wobei die Zahlen $a, b \in \mathbb{R}$ nur die Bedingung

$$a^2 + b^2 = ||x||^2 = 1$$

erfüllen müssen. Aus dem Satz von Pythagoras wissen wir aber, daß es dann genau einen Winkel $\alpha \in [0, 2\pi)$ gibt mit

$$a = \cos(\alpha) \quad \text{und} \quad b = \sin(\alpha),$$

und somit

$$A = T(\alpha) \quad \text{oder} \quad A = S(\alpha).$$

Beachten wir nun noch, daß

$$\det(T(\alpha)) = \cos(\alpha)^2 + \sin(\alpha)^2 = 1$$

und

$$\det(S(\alpha)) = -\cos(\alpha)^2 - \sin(\alpha)^2 = -1$$

ist, so ist

$$\mathrm{SO}(2) = \{T(\alpha) \mid \alpha \in [0, 2\pi)\}$$

und

$$\mathrm{O}(2) \setminus \mathrm{SO}(2) = \{S(\alpha) \mid \alpha \in [0, 2\pi)\},$$

gezeigt. □

F) Orthogonale Summe und orthogonale Projektion

Definition 12.30 (Orthogonale Summe)
Sei V ein euklidischer oder unitärer Raum. Wir nennen V die *orthogonale Summe* der Unterräume $U_1, \ldots, U_r$, falls $V = U_1 \oplus \ldots \oplus U_r$ und $U_i \perp U_j$ für $i \neq j$. In diesem Fall schreiben wir $V = U_1 \perp \ldots \perp U_r$.

Proposition 12.31 (Orthogonales Komplement)
Ist V ein endlich-dimensionaler euklidischer oder unitärer Raum und $U \leq V$, so gilt

$$V = U \perp U^{\perp}.$$

Insbesondere, $U^{\perp}$ ist ein Komplement von U.

Beweis: Da nach Voraussetzung $U \perp U^\perp$ gilt, bleibt $U \cap U^\perp = \{0\}$ und $V = U + U^\perp$ zu zeigen, wobei für letzteres auch $V \subseteq U + U^\perp$ reicht.

Ist $x \in U \cap U^\perp$, so gilt $\langle x, x\rangle = 0$ und damit $x = 0$, da das Skalarprodukt positiv definit ist. Also ist $U \cap U^\perp = \{0\}$.

Zudem können wir wegen Satz 12.19 eine ONB $(x_1, \ldots, x_r)$ von U wählen und diese zu einer ONB $(x_1, \ldots, x_n)$ von V ergänzen. Dann gilt aber

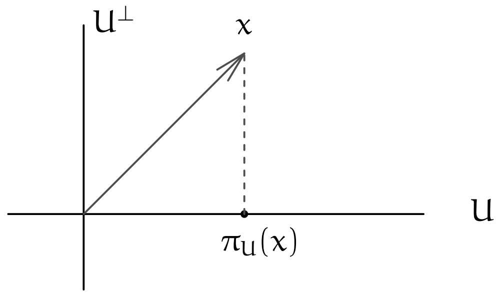
$$V = \mathrm{Lin}\,(x_1, \ldots, x_r) + \mathrm{Lin}\,(x_{r+1}, \ldots, x_n) \subseteq U + U^\perp,$$

da nach Wahl $x_{r+1}, \ldots, x_n \in U^\perp$. Hierbei beachte man, daß ein Vektor, der orthogonal zu einer Basis von U ist, automatisch orthogonal zu jedem Vektor in U ist. □

Bemerkung 12.32
Es sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $U \leq V$. Da sich jeder Vektor $x \in V$ in eindeutiger Weise darstellen läßt als $x = u + u'$ mit $u \in U$ und $u' \in U^\perp$, können wir die *orthogonale Projektion* von V auf U

$$\pi_U : V \to V$$

definieren durch $\pi\big(u + u'\big) = u$ für $u \in U$ und $u' \in U^\perp$.



In Aufgabe 12.35 wird gezeigt, daß π_U in der Tat eine Projektion ist, d.h. π_U ist linear mit $\pi_U^2 = \pi_U$. Außerdem ist $\mathrm{Im}(\pi_U) = U$ das Bild von π_U und $\mathrm{Ker}(\pi_U) = U^\perp$ der Kern.

Bemerkung 12.33 (Determinante als Volumenform)
In Bemerkung 8.12 haben wir das Parallelotop

$$P\big(x, y, z\big) := \big\{\lambda x + \mu y + \nu z \in \mathbb{R}^3 \mid 0 \leq \lambda, \mu, \nu \leq 1\big\}$$

betrachtet, das von den Vektoren $0 \neq x, y, z \in \mathbb{R}^3$ aufgespannt wird (siehe Abbildung 2).

Wir wollen zeigen, daß das Volumen

$$\text{Volumen}\,\big(P(x, y, z)\big) = |\det(x \;\; y \;\; z)|$$

die Determinante der Matrix ist, deren Spalten die Vektoren x, y und z sind.

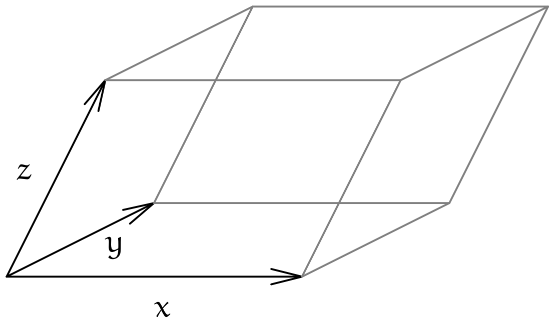


ABBILDUNG 2. Das Parallelotop $P(x,y,z)$ im $\mathbb{R}^3$

Elementargeometrisch berechnet sich das Volumen von $P(x,y,z)$ als Grundfläche mal Höhe, d.h. als Fläche A des von x und y aufgespannten Parallelogramms multipliziert mit der Höhe h des Parallelogramms. Dabei berechnet sich A als Länge von x mal der Höhe h' des Parallelogramms.

Wenden wir uns zunächst letzterer Berechnung zu. Es sei $U = \mathrm{Lin}\,(x)$ und π_U sei die orthogonale Projektion auf U. Dann ist die Höhe h' des Parallelogramms genau die Länge des Vektors $y - \pi_U(y)$ (siehe Abbildung 3).

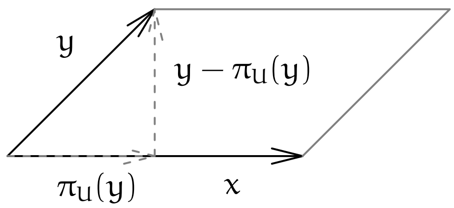


ABBILDUNG 3. Die Höhe im Parallelogramm zu x und y.

Für die Fläche A des Parallelogramms gilt deshalb

$$A = ||x|| \cdot ||y - \pi_U(y)||.$$

Auf ähnliche Weise kann man die Höhe h des Parallelotops $P(x,y,z)$ bestimmen. Hierzu betrachten wir den Unterraum $W = \mathrm{Lin}\,(x,y)$ und die orthogonale Projektion π_W auf W. Dann ist h die Länge des Vektors $z - \pi_W(z)$ (siehe Abbildung 4).

Für das Volumen von $P(x,y,z)$ erhalten wir deshalb

$$\text{Volumen}\big(P(x,y,z)\big) = A \cdot h = ||x|| \cdot ||y - \pi_U(y)|| \cdot ||z - \pi_W(z)||.$$

Wegen $\pi_U(y) \in U = \mathrm{Lin}\,(x)$ und $\pi_W(z) \in W = \mathrm{Lin}\,(x,y)$, gibt es $\lambda, \mu, \nu \in \mathbb{R}$ mit

$$\pi_U(y) = \lambda x \quad \text{und} \quad \pi_W(z) = \mu x + \nu y.$$

Dann gilt aber

$$\det(x \;\; y \;\; z) = \det(x \;\; y - \lambda x \;\; z - \mu x - \nu y) = \det(x \;\; y - \pi_U(y) \;\; z - \pi_W(z)),$$

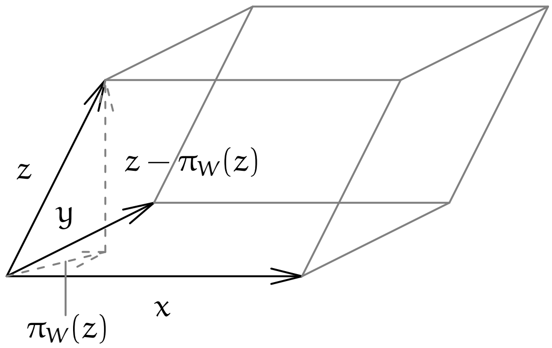


ABBILDUNG 4. Die Höhe in $P(x, y, z)$

da sich die Determinante einer Matrix nicht ändert, wenn wir Vielfache einer Spalte zu einer anderen addieren. Nun beachten wir, daß nach Konstruktion die Spalten der rechten Matrix orthogonal zueinander sind (siehe Abbildung 3 und 4). Normieren wir sie, so bilden sie eine ONB von $\mathbb{R}^3$ und die Matrix wird orthogonal. Da die Determinante einer orthogonalen Matrix Betrag 1 hat, erhalten wir also

$$\begin{aligned} |\det(x\ \ y\ \ z)| &= ||x|| \cdot ||y - \pi_U(y)|| \cdot ||z - \pi_W(z)|| \cdot \left|\det\begin{pmatrix} \frac{x}{||x||} & \frac{y-\pi_U(y)}{||y-\pi_U(y)||} & \frac{z-\pi_W(z)}{||z-\pi_W(z)||}\end{pmatrix}\right| \\ &= ||x|| \cdot ||y - \pi_U(y)|| \cdot ||z - \pi_W(z)||. \end{aligned}$$

Dies beweist die Aussage und begründet den Begriff *Volumenform* im Zusammenhang mit Determinanten. Man beachte auch, daß die entsprechende Aussage für Parallelogramme analog gezeigt werden kann.

Aufgaben

Aufgabe 12.34
Es sei $V = \mathcal{C}([0, 1], \mathbb{R})$ der Vektorraum der auf $[0, 1]$ stetigen Funktionen. Zeige, daß durch

$$\langle f, g\rangle := \int_0^1 f(x) \cdot g(x)\ dx$$

ein Skalarprodukt auf V definiert wird.

Aufgabe 12.35 (Orthogonale Projektion)
Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $U \leq V$.

a. Zeige, $\pi_U \in \mathrm{End}_{\mathbb{K}}(V)$ ist eine Projektion mit $\mathrm{Ker}(\pi_U) = U^\perp$ und $\mathrm{Im}(\pi_U) = U$.
b. Zeige, ist $\pi \in \mathrm{End}_{\mathbb{K}}(V)$ eine Projektion mit $\mathrm{Ker}(\pi) = U^\perp$ und $\mathrm{Im}(\pi) = U$, dann ist $\pi = \pi_U$.
c. Ist $(x_1, \ldots, x_r)$ eine ONB von U und $x \in V$, dann gilt

$$\pi_U(x) = \sum_{i=1}^{r} \langle x_i, x\rangle \cdot x_i.$$

Aufgabe 12.36
Es sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $U \leq V$. Dann gilt $(U^\perp)^\perp = U$.

Aufgabe 12.37
Zeige, durch $\big\langle (x_1,x_2,x_3)^t, (y_1,y_2,y_3)^t \big\rangle := x_1y_1 + x_1y_2 + x_2y_1 + 2x_2y_2 + x_2y_3 + x_3y_2 + 2x_3y_3$ für $(x_1,x_2,x_3)^t, (y_1,y_2,y_3)^t \in \mathbb{R}^3$ wird ein Skalarprodukt auf $\mathbb{R}^3$ definiert und bestimme eine Orthonormalbasis des $\mathbb{R}^3$ bezüglich dieses Skalarproduktes.

Aufgabe 12.38
Bestimme eine Orthonormalbasis des $\mathbb{R}$-Vektorraums $U = \{(v,w,x,y,z)^t \in \mathbb{R}^5 \mid v + w + x + y + z = 0\}$ bezüglich des Standardskalarproduktes auf $\mathbb{R}^5$.

Aufgabe 12.39 (Legendre-Polynome)
Betrachte den Vektorraum $U = \big\{ f : [0,1] \to \mathbb{R} : x \mapsto ax^2 + bx + c \mid a,b,c \in \mathbb{R} \big\}$ und bestimme eine ONB bezüglich des Skalarproduktes aus Aufgabe 12.34.

Aufgabe 12.40 (Tschebyscheff-Polynome)

a. Zeige, daß auf $V = \mathcal{C}\big([-1,1], \mathbb{R}\big)$ durch
$$V \times V \longrightarrow \mathbb{R} : (f,g) \mapsto \langle f, g \rangle := \int_{-1}^{1} f(x) \cdot g(x) \cdot \frac{1}{\sqrt{1-x^2}}\, dx$$
ein Skalarprodukt definiert wird.

b. Berechne für den Unterraum $U\big\{ f : [-1,1] \to \mathbb{R} : x \mapsto ax^2 + bx + c \mid a,b,c \in \mathbb{R} \big\}$ von V eine Orthonormalbasis bezüglich des Skalarproduktes aus Teil a..

Hinweis, in Teil a. substitutiere man $x = \cos(t)$, um die Konvergenz des uneigentlichen Integrals zu zeigen.

Aufgabe 12.41
Für $V = \mathrm{Mat}_n(\mathbb{R})$ definieren wir $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R} : (A,B) \mapsto \mathrm{Spur}\big(A^t \circ B\big)$.

a. Zeige, $\langle \cdot, \cdot \rangle$ ist ein Skalarprodukt auf V.

b. Zeige, für $U = \{A \in V \mid A^t = A\}$ gilt $U^\perp = \{A \in V \mid A^t = -A\}$.

Aufgabe 12.42
Sei V ein euklidischer oder unitärer Raum und $||\cdot|| : V \to \mathbb{R} : x \mapsto \sqrt{\langle x, x \rangle}$ die durch das Skalarprodukt definierte Norm. Zeige, für $x, y \in V$ gelten:

a. Die Parallelogrammgleichung: $||x+y||^2 + ||x-y||^2 = 2 \cdot \big(||x||^2 + ||y||^2\big)$.

b. Der Satz des Pythagoras': $x \perp y \implies ||x||^2 + ||y||^2 = ||x+y||^2$.

Aufgabe 12.43
Zeige, für $a, b \in \mathbb{R}$ gibt es genau dann zwei normierte Vektoren $x = (u,v,a)^t \in \mathbb{R}^3$ und $y = (r,s,b)^t \in \mathbb{R}^3$ die bezüglich des Standardskalarproduktes orthogonal zueinander sind, wenn $a^2 + b^2 \leq 1$.

Aufgabe 12.44 (Der p-adische Betrag)
Sei p eine Primzahl. Für $0 \neq a \in \mathbb{Z}$ bezeichne $\nu_p(a)$ die höchste Potenz von p, die a

teilt, und für $0 \neq q = \frac{b}{c} \in \mathbb{Q}$ setzen wir $\nu_p(q) = \nu_p(b) - \nu_p(c)$. Zeige, die Abbildung

$$|\cdot|_p : \mathbb{Q} \longrightarrow \mathbb{R} : q \mapsto \begin{cases} p^{-\nu_p(q)}, & \text{wenn } q \neq 0, \\ 0, & \text{wenn } q = 0. \end{cases}$$

ist positiv definit, multiplikativ und genügt der Dreiecksungleichung.

§ 13 Spektralsatz und Hauptachsentransformation

In diesem Abschnitt sei V ein euklidischer oder unitärer Raum der Dimension $1 \leq n < \infty$ mit Skalarprodukt $\langle\cdot,\cdot\rangle$ und euklidischer Norm $\|\cdot\|$.

A) Die adjungierte Abbildung

Satz 13.1 (Die adjungierte Abbildung)
Zu jedem Endomorphismus $f \in \mathrm{End}_{\mathbb{K}}(V)$ *gibt es genau ein* $f^* \in \mathrm{End}_{\mathbb{K}}(V)$*, so daß*

$$\langle f(x), y\rangle = \langle x, f^*(y)\rangle \tag{46}$$

für alle $x, y \in V$ *gilt. Ist* $B = (x_1, \ldots, x_n)$ *eine ONB von* V*, so gilt für* $y \in V$

$$f^*(y) = \sum_{i=1}^{n} \big\langle f(x_i), y\big\rangle \cdot x_i. \tag{47}$$

Die Abbildung f^* *heißt die* adjungierte Abbildung *zu* f*.*

Beweis: Wir wollen zunächst zeigen, daß es einen Endomorphismus $f^* \in \mathrm{End}_{\mathbb{K}}(V)$ mit der Eigenschaft (46) gibt. Dazu wählen wir eine ONB $B = (x_1, \ldots, x_n)$ von V und definieren

$$f^* : V \longrightarrow V : y \mapsto \sum_{i=1}^{n} \big\langle f(x_i), y\big\rangle \cdot x_i,$$

d.h. wir definieren $f^*(y)$ durch die Formel in (47). Da das Skalarprodukt in der zweiten Komponente linear ist, ist f^* in der Tat eine lineare Abbildung, also ein Endomorphismus von V.

Seien nun $x, y \in V$ gegeben. Unter Anwendung der Parsevalschen Gleichung 12.17 gilt dann

$$\begin{aligned}
\langle f(x), y\rangle &\overset{12.17}{=} \left\langle f\left(\sum_{i=1}^{n}\langle x_i, x\rangle \cdot x_i\right), y\right\rangle = \left\langle \sum_{i=1}^{n}\langle x_i, x\rangle \cdot f(x_i), y\right\rangle \\
&= \sum_{i=1}^{n} \overline{\langle x_i, x\rangle} \cdot \langle f(x_i), y\rangle = \sum_{i=1}^{n}\langle x, x_i\rangle \cdot \langle f(x_i), y\rangle \\
&= \sum_{i=1}^{n} \big\langle x, \langle f(x_i), y\rangle \cdot x_i\big\rangle = \left\langle x, \sum_{i=1}^{n}\langle f(x_i), y\rangle \cdot x_i\right\rangle = \langle x, f^*(y)\rangle.
\end{aligned}$$

Damit ist gezeigt, daß der durch (47) definierte Endomorphismus die Gleichung (46) erfüllt.

Es bleibt noch die Eindeutigkeit zu zeigen. Sei dazu $h \in \mathrm{End}_{\mathbb{K}}(V)$ mit

$$\langle f(x), y\rangle = \langle x, h(y)\rangle \tag{48}$$

für alle $x, y \in V$. Wegen der Parsevalschen Gleichung 12.17 gilt für $y \in V$ dann

$$h(y) \overset{12.17}{=} \sum_{i=1}^{n} \langle x_i, h(y) \rangle \cdot x_i \overset{(48)}{=} \sum_{i=1}^{n} \langle f(x_i), y \rangle \cdot x_i \overset{(47)}{=} f^*(y).$$

Mithin stimmen f^* und h überein, so daß f^* eindeutig bestimmt ist. □

Korollar 13.2 ($f^{**} = f$)
Ist $f \in \mathrm{End}_{\mathbb{K}}(V)$, *so gilt* $f^{**} = f$, *d.h.* $\langle f^*(x), y \rangle = \langle x, f(y) \rangle$ *für* $x, y \in V$.

Beweis: Für $x, y \in V$ gilt

$$\langle f^*(x), y \rangle = \overline{\langle y, f^*(x) \rangle} \overset{(46)}{=} \overline{\langle f(y), x \rangle} = \langle x, f(y) \rangle.$$

Damit erfüllt f die Bedingung, durch die die Abbildung f^{**} eindeutig festgelegt ist. Also muß $f = f^{**}$ gelten. □

Korollar 13.3 (Matrixdarstellung der adjungierten Abbildung)
Sei B *eine ONB von* V *und* $f \in \mathrm{End}_{\mathbb{K}}(V)$, *so gilt*

$$M_B^B(f^*) = M_B^B(f)^*,$$

d.h. die Matrixdarstellung der adjungierten Abbildung ist die Adjungierte der Matrixdarstellung.

Beweis: Seien $M_B^B(f) = (a_{ij})_{i,j}$ und $M_B^B(f^*) = (b_{ji})_{j,i}$. Unter Berücksichtigung der Parsevalschen Gleichung 12.17 gilt

$$\sum_{i=1}^{n} a_{ij} \cdot x_i = f(x_j) \overset{12.17}{=} \sum_{i=1}^{n} \langle x_i, f(x_j) \rangle \cdot x_i \tag{49}$$

und

$$\sum_{j=1}^{n} b_{ji} \cdot x_j = f^*(x_i) \overset{12.17}{=} \sum_{j=1}^{n} \langle x_j, f^*(x_i) \rangle \cdot x_j. \tag{50}$$

Da die Darstellung als Linearkombination einer Basis eindeutig ist, erhalten wir

$$\overline{a_{ij}} \overset{(49)}{=} \overline{\langle x_i, f(x_j) \rangle} = \langle f(x_j), x_i \rangle \overset{(46)}{=} \langle x_j, f^*(x_i) \rangle \overset{(50)}{=} b_{ji}.$$

Daraus folgt

$$M_B^B(f^*) = (b_{ji})_{j,i} = \left(\overline{a_{ij}}\right)_{j,i} = \left(\overline{a_{ij}}\right)_{i,j}^t = M_B^B(f)^*.$$

□

Beispiel 13.4 (Adjungierte)
Wir betrachten $V = \mathbb{C}^2$ als unitären Raum mit dem kanonischen Skalarprodukt sowie die Abbildung

$$f : \mathbb{C}^2 \longrightarrow \mathbb{C}^2 : (x, y)^t \mapsto (2x + 4y, 2y - 4x)^t.$$

Bezüglich der kanonischen Basis E hat f die Matrixdarstellung

$$M_E^E(f) = \begin{pmatrix} 2 & 4 \\ -4 & 2 \end{pmatrix}$$

und somit gilt

$$M^E_E(f)^* = \overline{\begin{pmatrix} 2 & 4 \\ -4 & 2 \end{pmatrix}}^t = \begin{pmatrix} 2 & 4 \\ -4 & 2 \end{pmatrix}^t = \begin{pmatrix} 2 & -4 \\ 4 & 2 \end{pmatrix}.$$

Da E eine ONB bezüglich des kanonischen Skalarproduktes ist, ist somit

$$f^* : \mathbb{C}^2 \longrightarrow \mathbb{C}^2 : (x, y)^t \mapsto (2x - 4y, 2y + 4x)^t$$

nach Korollar 13.3 die Adjungierte von f.

B) Der Spektralsatz für normale Endomorphismen

Definition 13.5
Es sei $f \in \mathrm{End}_{\mathbb{K}}(V)$ und $A \in \mathrm{Mat}_n(\mathbb{K})$.

a. f heißt *normal*, falls $f^* \circ f = f \circ f^*$.
b. A heißt *normal*, falls $A^* \circ A = A \circ A^*$.

Bemerkung 13.6

a. Jede symmetrische Matrix $A \in \mathrm{Mat}_n(\mathbb{R})$ und jede hermitesche Matrix $A \in \mathrm{Mat}_n(\mathbb{C})$ ist normal, denn wegen $A = A^*$ gilt auch

$$A \circ A^* = A \circ A = A^* \circ A.$$

b. Jede orthogonale Matrix $A \in O(n)$ und jede unitäre Matrix $A \in U(n)$ ist normal, denn wegen $A^* = A^{-1}$ gilt auch

$$A \circ A^* = A \circ A^{-1} = \mathbb{1}_n = A^{-1} \circ A = A^* \circ A.$$

Lemma 13.7 (Matrixdarstellung normaler Endomorphismen)
Sei $f \in \mathrm{End}_{\mathbb{K}}(V)$ *und* B *eine ONB von* V.
Genau dann ist f *normal, wenn* $M^B_B(f)$ *normal ist.*

Beweis: Ist f normal, so gilt

$$\begin{aligned} M^B_B(f)^* \circ M^B_B(f) &\overset{13.3}{=} M^B_B(f^*) \circ M^B_B(f) = M^B_B(f^* \circ f) \\ &= M^B_B(f \circ f^*) = M^B_B(f) \circ M^B_B(f^*) \overset{13.3}{=} M^B_B(f) \circ M^B_B(f)^* \end{aligned}$$

und somit ist $M^B_B(f)$ normal. Ist umgekehrt $M^B_B(f)$ normal, so gilt

$$\begin{aligned} M^B_B(f^* \circ f) &= M^B_B(f^*) \circ M^B_B(f) \overset{13.3}{=} M^B_B(f)^* \circ M^B_B(f) \\ &= M^B_B(f) \circ M^B_B(f)^* \overset{13.3}{=} M^B_B(f) \circ M^B_B(f^*) = M^B_B(f \circ f^*). \end{aligned}$$

Dann stimmen aber die Abbildungen $f^* \circ f$ und $f \circ f^*$ überein, und somit ist f normal. $\square$

Beispiel 13.8 (Normale Abbildung)
In Beispiel 13.4 gilt

$$\begin{aligned} M_E^E(f) \circ M_E^E(f)^* &= \begin{pmatrix} 2 & 4 \\ -4 & 2 \end{pmatrix} \circ \begin{pmatrix} 2 & -4 \\ 4 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 20 & 0 \\ 0 & 20 \end{pmatrix} = \begin{pmatrix} 2 & -4 \\ 4 & 2 \end{pmatrix} \circ \begin{pmatrix} 2 & 4 \\ -4 & 2 \end{pmatrix} = M_E^E(f)^* \circ M_E^E(f). \end{aligned}$$

Somit ist $M_E^E(f)$ normal und da E eine ONB ist, ist dann auch f normal.

Lemma 13.9 (Normale Abbildungen)
Sei $f \in \mathrm{End}_{\mathbb{K}}(V)$ *normal. Genau dann gilt* $x \in \mathrm{Eig}(f, \lambda)$, *wenn* $x \in \mathrm{Eig}\left(f^*, \overline{\lambda}\right)$.

Beweis: Für ein beliebiges $x \in V$ gilt

$$\begin{aligned} \langle f^*(x) - \overline{\lambda}x, f^*(x) - \overline{\lambda}x \rangle &= \langle f^*(x), f^*(x) \rangle - \langle f^*(x), \overline{\lambda}x \rangle - \langle \overline{\lambda}x, f^*(x) \rangle + \langle \overline{\lambda}x, \overline{\lambda}x \rangle \\ &= \langle f \circ f^*(x), x \rangle - \overline{\lambda}\langle f^*(x), x \rangle - \lambda\langle x, f^*(x) \rangle + \lambda\overline{\lambda}\langle x, x \rangle \\ &= \langle f^* \circ f(x), x \rangle - \overline{\lambda}\langle x, f(x) \rangle - \lambda\langle f(x), x \rangle + \lambda\overline{\lambda}\langle x, x \rangle \\ &= \langle f(x), f(x) \rangle - \langle \lambda x, f(x) \rangle - \langle f(x), \lambda x \rangle + \langle \lambda x, \lambda x \rangle \\ &= \langle f(x) - \lambda x, f(x) - \lambda x \rangle. \end{aligned}$$

Da das Skalarprodukt positiv definit ist, gilt somit

$$\begin{aligned} x \in \mathrm{Eig}(f, \lambda) &\iff f(x) - \lambda x = 0 \iff \langle f(x) - \lambda x, f(x) - \lambda x \rangle = 0 \\ &\iff \langle f^*(x) - \overline{\lambda}x, f^*(x) - \overline{\lambda}x \rangle = 0 \iff f^*(x) - \overline{\lambda}x = 0 \\ &\iff x \in \mathrm{Eig}\left(f^*, \overline{\lambda}\right). \end{aligned}$$

□

Satz 13.10 (Spektralsatz für normale Endomorphismen)
Für $f \in \mathrm{End}_{\mathbb{K}}(V)$ *sind die folgenden beiden Aussagen gleichwertig:*

a. f *ist normal und* χ_f *zerfällt über* $\mathbb{K}$ *in Linearfaktoren.*

b. V *besitzt eine ONB aus Eigenvektoren von* f.

Insbesondere ist f *dann bezüglich einer ONB diagonalisierbar.*

Beweis:
b. $\Longrightarrow$ a.: Besitzt V eine ONB B aus Eigenvektoren, so zerfällt χ_f nach Satz 10.20 über $\mathbb{K}$ in Linearfaktoren. Zudem ist dann $M_B^B(f)$ eine Diagonalmatrix, und wegen Korollar 13.3 ist dann auch

$$M_B^B(f)^* = \overline{M_B^B(f)}^t$$

eine Diagonalmatrix. Da zwei Diagonalmatrizen stets kommutieren, gilt also

$$M_B^B(f)^* \circ M_B^B(f) = M_B^B(f) \circ M_B^B(f)^*,$$

d.h. $M_B^B(f)$ ist normal. Nach Lemma 13.7 ist dann aber auch f normal.

a. $\Longrightarrow$ b.: Wir führen den Beweis mit Induktion nach $n = \dim_{\mathbb{K}}(V)$, wobei für $n = 1$ der Endomorphismus f für jede ONB $B = (x_1)$ diagonalisierbar ist und zudem x_1 ein Eigenvektor von f ist. Sei also $n > 1$.

Da χ_f über $\mathbb{K}$ in Linearfaktoren zerfällt, besitzt f einen Eigenwert $\lambda \in \mathbb{K}$ und einen zugehörigen Eigenvektor $0 \neq x \in V$. Dann ist $U := \mathrm{Lin}\,(x)$ ein f-invarianter Unterraum der Dimension 1 und es gilt

$$V = U \perp U^{\perp} = U \oplus U^{\perp}$$

nach Proposition 12.31.

Wir zeigen nun zunächst, daß auch $U^{\perp}$ ein f-invarianter Unterraum ist, der dann die Dimension $n - 1$ hat. Sei dazu $y \in U^{\perp}$, dann gilt

$$\langle f(y), x\rangle = \langle y, f^*(x)\rangle \overset{13.9}{=} \langle y, \overline{\lambda}x\rangle = \overline{\lambda} \cdot \langle y, x\rangle = 0,$$

da $y \perp x$. Damit gilt dann aber auch $f(y) \perp x$, und somit $f(y) \in U^{\perp}$.

Als nächstes wollen wir zeigen, daß $f_{U^{\perp}}$ normal ist. Dazu beachten wir zunächst, daß $U^{\perp}$ auch f^*-invariant ist, da für $y \in U^{\perp}$ wie oben

$$\langle f^*(y), x\rangle = \langle y, f(x)\rangle = \langle y, \lambda x\rangle = \lambda \cdot \langle y, x\rangle = 0$$

und somit $f^*(y) \perp x$ und $f^*(y) \in U^{\perp}$ gilt. Aus der definierenden Eigenschaft der adjungierten Abbildung folgt dann aber, daß die adjungierte Abbildung $(f_{U^{\perp}})^*$ der Einschränkung $f_{U^{\perp}}$ genau die Einschränkung $(f^*)_{U^{\perp}}$ der adjungierten Abbildung f^* auf $U^{\perp}$ ist. Die Normalität von f überträgt sich also direkt auf $f_{U^{\perp}}$ durch

$$f_{U^{\perp}} \circ (f_{U^{\perp}})^* = (f \circ f^*)_{U^{\perp}} = (f^* \circ f)_{U^{\perp}} = (f_{U^{\perp}})^* \circ f_{U^{\perp}}.$$

Außerdem gilt

$$\chi_f = \chi_{f_U} \cdot \chi_{f_{U^{\perp}}},$$

da V die direkte Summe der beiden f-invarianten Unterräume U und $U^{\perp}$ ist, und deshalb zerfällt $\chi_{f_{U^{\perp}}}$ über $\mathbb{K}$ in Linearfaktoren. Nach Induktion besitzt $U^{\perp}$ deshalb eine ONB $(x_2, \ldots, x_n)$ aus Eigenvektoren von $f_{U^{\perp}}$. Dann ist aber $B = (x_1, x_2, \ldots, x_n)$ mit $x_1 = \frac{x}{||x||}$ eine ONB aus Eigenvektoren von f. $\square$

Korollar 13.11 (Spektralsatz für normale Matrizen)
Für eine Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ sind die folgenden beiden Aussagen gleichwertig:

a. A *ist normal und* χ_A *zerfällt über* $\mathbb{K}$ *in Linearfaktoren.*
b. *Es gibt ein* T *in* $\mathrm{O}(n)$ *bzw.* $\mathrm{U}(n)$, *so daß* $T^{-1} \circ A \circ T$ *eine Diagonalmatrix ist.*

Beweis: Wenden wir den Spektralsatz 13.10 auf f_A und $\mathbb{K}^n$ mit dem kanonischen Skalarprodukt an, so enthält die Basistransformationsmatrix $T = T^B_E$ genau die Vektoren der ONB B als Spalten und ist nach Proposition 12.26 daher orthogonal bzw. unitär. $\square$

Der Beweis ist konstruktiv, sofern man die Eigenwerte von A exakt kennt. Man leitet daraus folgenden prinzipiellen Algorithmus zur Bestimmung von T her.

Algorithmus 13.12

INPUT: $A \in \mathrm{Mat}_n(\mathbb{K})$ normal mit χ_A zerfällt über $\mathbb{K}$.

OUTPUT: T in $\mathrm{O}(n)$ bzw. $\mathrm{U}(n)$, so daß $T^{-1} \circ A \circ T$ Diagonalgestalt hat.

1. **Schritt:** Bestimme die Eigenwerte von A.
2. **Schritt:** Bestimme für jeden Eigenwert von A eine Basis des zugehörigen Eigenraumes.
3. **Schritt:** Orthonormalisiere die Basen der Eigenräume mit dem Orthonormalisierungsverfahren von Gram-Schmidt und schreibe die Vektoren als Spalten in eine Matrix T.
4. **Schritt:** Gib schließlich T zurück.

Beispiel 13.13 (Diagonalisierung einer normalen Abbildung)
Die Abbildung f in Beispiel 13.4 ist nach Beispiel 13.8 normal.

$$\chi_f = \det\big(t\cdot\mathbb{1}_2 - M_E^E(f)\big) = \begin{vmatrix} t-2 & -4 \\ 4 & t-2 \end{vmatrix} = t^2-4t+20 = \big(t-(2+4i)\big)\cdot\big(t-(2-4i)\big)$$

zerfällt über $\mathbb{C}$ in Linearfaktoren. Mithin gibt es wegen des Spektralsatzes eine ONB aus Eigenvektoren von f, so daß

$$M_B^B(f) = \begin{pmatrix} 2+4i & 0 \\ 0 & 2-4i \end{pmatrix}.$$

Um B zu berechnen, berechnen wir zunächst die beiden Eigenräume $\mathrm{Eig}(f, 2+4i)$ und $\mathrm{Eig}(f, 2-4i)$ ausgehend von der Matrixdarstellung $M_E^E(f)$ in Beispiel 13.4.

Für $\mathrm{Eig}(f, 2+4i) = \text{Lös}\big(M_E^E(f) - (2+4i)\cdot\mathbb{1}_2, 0\big)$ liefert unser Algorithmus

$$\begin{pmatrix} -4i & 4 \\ -4 & -4i \end{pmatrix} \overset{\mathrm{II}\mapsto\mathrm{II}+i\cdot\mathrm{I}}{\underset{\mathrm{I}\mapsto\frac{1}{-4i}\cdot\mathrm{I}}{\longmapsto}} \begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix} \overset{-1\text{'en}}{\underset{\text{ergänzen}}{\longmapsto}} \begin{pmatrix} 1 & i \\ 0 & -1 \end{pmatrix},$$

so daß $x_1 = (i, -1)^t$ eine Basis von $\mathrm{Eig}(f, 2+4i)$ ist. Analog erhalten wir für $\mathrm{Eig}(f, 2-4i) = \text{Lös}\big(M_E^E(f) - (2-4i)\cdot\mathbb{1}_2, 0\big)$

$$\begin{pmatrix} 4i & 4 \\ -4 & 4i \end{pmatrix} \overset{\mathrm{II}\mapsto\mathrm{II}-i\cdot\mathrm{I}}{\underset{\mathrm{I}\mapsto\frac{1}{4i}\cdot\mathrm{I}}{\longmapsto}} \begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix} \overset{-1\text{'en}}{\underset{\text{ergänzen}}{\longmapsto}} \begin{pmatrix} 1 & -i \\ 0 & -1 \end{pmatrix},$$

so daß $x_2 = (-i, -1)^t$ eine Basis von $\mathrm{Eig}(f, 2-4i)$ ist. Die Vektoren x_1 und x_2 sind bereits orthogonal zueinander, da

$$\langle x_1, x_2\rangle = \overline{i}\cdot(-i) + \overline{-1}\cdot(-1) = -1+1 = 0.$$

Mithin reicht es, sie zu normieren, und wir erhalten die gewünschte ONB

$$B = \left(\tfrac{1}{\|x_1\|}\cdot x_1, \tfrac{1}{\|x_2\|}\cdot x_2\right) = \left(\left(\tfrac{i}{\sqrt{2}}, \tfrac{-1}{\sqrt{2}}\right)^t, \left(\tfrac{-i}{\sqrt{2}}, \tfrac{-1}{\sqrt{2}}\right)^t\right).$$

Es ist übrigens kein Zufall, daß die beiden Eigenvektoren im letzten Beispiel orthogonal zueinander standen.

Lemma 13.14
Ist $f \in \mathrm{End}_{\mathbb{K}}(V)$ *normal,* $x \in \mathrm{Eig}(f,\lambda)$, $y \in \mathrm{Eig}(f,\mu)$ *und* $\lambda \neq \mu$, *dann gilt* $x \perp y$.

Beweis: Nach Voraussetzung gilt

$$\lambda \cdot \langle x, y\rangle = \langle \overline{\lambda} x, y\rangle \overset{13.9}{=} \langle f^*(x), y\rangle \overset{13.2}{=} \langle x, f(y)\rangle = \langle x, \mu y\rangle = \mu \cdot \langle x, y\rangle.$$

Für die Differenz der beiden Seiten erhalten wir dann $(\lambda - \mu) \cdot \langle x, y\rangle = 0$, wobei nach Voraussetzung $\lambda - \mu \neq 0$ gilt. Also ist $\langle x, y\rangle = 0$ und somit $x \perp y$. □

Satz 13.15 (Spektralzerlegung für normale Endomorphismen)
Sei $f \in \mathrm{End}_{\mathbb{K}}(V)$ *normal mit zerfallendem charakteristischem Polynom und seien* $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ *die paarweise verschiedenen Eigenwerte von* f. *Ferner bezeichne*

$$\pi_i : V \longrightarrow V$$

die orthogonale Projektion von V *auf* $\mathrm{Eig}(f, \lambda_i)$.

Dann ist

$$V = \mathrm{Eig}(f, \lambda_1) \perp \ldots \perp \mathrm{Eig}(f, \lambda_r)$$

die orthogonale Summe *der Eigenräume von* f *und es gilt*

$$f = \lambda_1 \cdot \pi_1 + \ldots + \lambda_r \cdot \pi_r.$$

Man nennt dies die Spektralzerlegung *von* f.

Beweis: Nach dem Spektralsatz für normale Abbildungen ist f diagonalisierbar. Deshalb folgt aus Satz 10.20, daß V die direkte Summe der Eigenräume von f ist. Wegen Lemma 13.14 ist diese Summe dann eine orthogonale Summe.

Ist nun $x = x_1 + \ldots + x_r \in V$ mit $x_i \in \mathrm{Eig}(f, \lambda_i)$ gegeben, so gilt

$$\pi_j(x_i) = \delta_{ij} \cdot x_i,$$

da $x_j \in \mathrm{Eig}(f, \lambda_j)$ und $x_i \perp \mathrm{Eig}(f, \lambda_j)$ für $i \neq j$, und somit

$$\pi_j(x) = \pi_j(x_1) + \ldots + \pi_j(x_r) = x_j.$$

Damit erhalten wir dann

$$\begin{aligned}(\lambda_1 \cdot \pi_1 + \ldots + \lambda_r \cdot \pi_r)(x) &= \lambda_1 \cdot \pi_1(x) + \ldots + \lambda_r \cdot \pi_r(x)\\ &= \lambda_1 \cdot x_1 + \ldots + \lambda_r \cdot x_r = f(x_1) + \ldots + f(x_r) = f(x).\end{aligned}$$

□

C) Orthogonale und unitäre Abbildungen

Wir kommen jetzt zu den strukturerhaltenden Abbildungen, d. h. zu solchen, die mit dem Skalarprodukt verträglich sind. Diese haben einen speziellen Namen.

Definition 13.16 (Orthogonale / unitäre Abbildungen)

a. Ein Endomorphismus $f \in \operatorname{End}_{\mathbb{R}}(V)$ heißt *orthogonal*, falls $f^* \circ f = \mathrm{id}_V$ gilt. $O(V) := \{f \in \operatorname{End}_{\mathbb{R}}(V) \mid f \text{ ist orthogonal}\}$ heißt die *orthogonale Gruppe* von V.

b. Ein Endomorphismus $f \in \operatorname{End}_{\mathbb{C}}(V)$ heißt *unitär*, falls $f^* \circ f = \mathrm{id}_V$ gilt. Wir nennen $U(V) := \{f \in \operatorname{End}_{\mathbb{C}}(V) \mid f \text{ ist unitär}\}$ die *unitäre Gruppe* von V.

Proposition 13.17 (Matrixdarstellung orthogonaler / unitärer Abbildungen)
Es sei $f \in \operatorname{End}_{\mathbb{K}}(V)$ und B eine ONB von V.
Genau dann ist f orthogonal bzw. unitär, wenn $M_B^B(f)$ orthogonal bzw. unitär ist.

Beweis: Es gilt:

$$\begin{aligned} f \text{ ist orthogonal bzw. unitär} &\iff f^* \circ f = \mathrm{id}_V \\ &\iff M_B^B(f^*) \circ M_B^B(f) = M_B^B(f^* \circ f) = \mathbb{1}_n \\ &\overset{13.3}{\iff} M_B^B(f)^* \circ M_B^B(f) = \mathbb{1}_n \\ &\iff M_B^B(f) \text{ ist orthogonal bzw. unitär.} \end{aligned}$$

□

Korollar 13.18 (Orthogonal / unitär $\Longrightarrow$ normal)
Orthogonale und unitäre Abbildungen sind normal.

Beweis: Ist $f \in \operatorname{End}_{\mathbb{K}}(V)$ orthogonal oder unitär und B eine ONB von V, so ist $M_B^B(f)$ nach Proposition 13.17 orthogonal oder unitär. Nach Bemerkung 13.6 ist dann $M_B^B(f)$ auch normal, und mit Lemma 13.7 ist f deshalb normal. □

Proposition 13.19 (Charakterisierung orthogonaler / unitärer Endomorphismen)
Ein Endomorphismus $f \in \operatorname{End}_{\mathbb{K}}(V)$ ist genau dann orthogonal bzw. unitär, wenn

$$\langle f(x), f(y)\rangle = \langle x, y\rangle$$

für alle $x, y \in V$ gilt.

Beweis: Ist f orthogonal bzw. unitär und sind $x, y \in V$, so gilt

$$\langle x, y\rangle = \langle x, \mathrm{id}_V(y)\rangle = \langle x, f^* \circ f(y)\rangle = \langle f(x), f(y)\rangle.$$

Ist umgekehrt f mit dem Skalarprodukt verträglich, so gilt

$$\langle x, y\rangle = \langle f(x), f(y)\rangle = \langle (f^* \circ f)(x), y\rangle$$

und somit

$$\langle (f^* \circ f)(x) - \mathrm{id}_V(x), y\rangle = \langle (f^* \circ f)(x), y\rangle - \langle x, y\rangle = 0$$

für alle $x, y \in V$. Für $x \in V$ beliebig setzen wir $y = (f^* \circ f)(x) - \mathrm{id}_V(x)$ und erhalten

$$\langle (f^* \circ f)(x) - \mathrm{id}_V(x), (f^* \circ f)(x) - \mathrm{id}_V(x) \rangle = 0,$$

was $(f^* \circ f)(x) - \mathrm{id}_V(x) = 0$ und $(f^* \circ f)(x) = \mathrm{id}_V(x)$ zur Folge hat. Mithin ist $f^* \circ f = \mathrm{id}_V$ und f ist orthogonal bzw. unitär. □

Proposition 13.20 (Eigenschaften orthogonaler / unitärer Abbildungen)
Es seien $f, g \in \mathrm{End}_{\mathbb{K}}(V)$ *orthogonal bzw. unitär und* $x, y \in V$

a. $||f(x)|| = ||x||$.

b. $x \perp y$ *genau dann, wenn* $f(x) \perp f(y)$.

c. *Jeder Eigenwert von* f *hat Betrag* 1.

d. f *ist bijektiv.*

e. f^{-1} *und* $f \circ g$ *sind orthogonal bzw. unitär, d.h.* $\mathrm{O}(V)$ *und* $\mathrm{U}(V)$ *sind Gruppen.*

Insbesondere, orthogonale und unitäre Abbildungen erhalten Längen und Abstände.

Beweis:

a. $||f(x)||^2 = \langle f(x), f(x) \rangle = \langle x, x \rangle = ||x||^2$.

b. $x \perp y \iff \langle f(x), f(y) \rangle = \langle x, y \rangle = 0 \iff f(x) \perp f(y)$.

c. Ist $0 \neq x \in V$ ein Eigenvektor zum Eigenwert $\lambda \in \mathbb{K}$, so gilt nach a.

$$||x|| = ||f(x)|| = ||\lambda x|| = |\lambda| \cdot ||x||,$$

also $|\lambda| = 1$, da $x \neq 0$.

d. Ist $x \in \mathrm{Ker}(f)$, so gilt nach a. $0 = ||f(x)|| = ||x||$, und somit $x = 0$. Also ist f injektiv, und da V endlich-dimensional ist, ist f somit auch bijektiv.

e. f^{-1} und $f \circ g$ sind orthogonal bzw. unitär, da für $x, y \in V$ gelten

$$\left\langle f^{-1}(x), f^{-1}(y) \right\rangle = \left\langle f\big(f^{-1}(x)\big), f\big(f^{-1}(y)\big) \right\rangle = \langle x, y \rangle$$

und

$$\big\langle (f \circ g)(x), (f \circ g)(y) \big\rangle = \left\langle f\big(g(x)\big), f\big(g(y)\big) \right\rangle = \big\langle g(x), g(y) \big\rangle = \langle x, y \rangle.$$

□

Bemerkung 13.21 (Orthogonale Abbildungen sind winkeltreu.)
Orthogonale Abbildungen erhalten Winkel, d.h. ist $f \in \mathrm{O}(V)$ und sind $0 \neq x, y \in V$, so gilt

$$\measuredangle\big(f(x), f(y)\big) = \measuredangle(x, y),$$

da f das Skalarprodukt und die euklidische Norm erhält. Es gilt nämlich

$$\measuredangle\big(f(x), f(y)\big) = \arccos\left(\frac{\langle f(x), f(y) \rangle}{||f(x)|| \cdot ||f(y)||}\right) = \arccos\left(\frac{\langle x, y \rangle}{||x|| \cdot ||y||}\right) = \measuredangle(x, y).$$

Satz 13.22 (Spektralsatz für unitäre Abbildungen)
Ein Endomorphismus $f \in \mathrm{End}_{\mathbb{C}}(V)$ *ist genau dann unitär, wenn* V *eine ONB aus Eigenvektoren von* f *besitzt und alle Eigenwerte von* f *Betrag* 1 *haben.*

Beweis: Ist f unitär, so ist f nach Korollar 13.18 normal und χ_f zerfällt nach dem Fundamentalsatz der Algebra über $\mathbb{C}$ in Linearfaktoren. Aufgrund des Spektralsatzes für normale Abbildungen 13.10 besitzt V dann eine ONB aus Eigenvektoren von f. Zudem haben die Eigenwerte nach Proposition 13.20 alle Betrag 1.

Besitzt umgekehrt V eine ONB $B = (x_1, \ldots, x_n)$ aus Eigenvektoren von f zu Eigenwerten $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ vom Betrag 1, so gilt

$$
\begin{aligned}
M_B^B(f)^* \circ M_B^B(f) &= \begin{pmatrix} \overline{\lambda_1} & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \overline{\lambda_n} \end{pmatrix} \circ \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix} \\
&= \begin{pmatrix} \overline{\lambda_1}\lambda_1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \overline{\lambda_n}\lambda_n \end{pmatrix} = \begin{pmatrix} |\lambda_1|^2 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & |\lambda_n|^2 \end{pmatrix} = \mathbb{1}_n.
\end{aligned}
$$

Mithin ist $M_B^B(f)$ unitär, und damit ist auch f unitär nach Proposition 13.17. □

Korollar 13.23 (Spektralsatz für unitäre Matrizen)
Ist $A \in \mathrm{U}(n)$, *dann gibt es ein* $T \in \mathrm{U}(n)$ *mit*

$$
T^{-1} \circ A \circ T = T^* \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & \lambda_n \end{pmatrix}
$$

wobei $\lambda_i \in \mathbb{C}$ *mit* $|\lambda_i| = 1$, $i = 1, \ldots, n$, *die Eigenwerte von* A *sind. Insbesondere ist jede unitäre Matrix diagonalisierbar.*

Beweis: Ist A unitär, dann ist f_A unitär und wir finden eine ONB von $\mathbb{C}^n$ aus Eigenvektoren von f_A, also von A, und alle Eigenwerte haben Betrag 1. Schreiben wir die Eigenvektoren als Spalten in eine Matrix T, so ist $T \in \mathrm{U}(n)$ und T transformiert A in eine Diagonalmatrix. □

Beispiel 13.24 (Unitäre Matrix)
Betrachten wir $\mathbb{C}^3$ mit dem kanonischen Skalarprodukt sowie die Matrix

$$
A = \frac{1}{9} \cdot \begin{pmatrix} 1 & 8 & -4 \\ -4 & 4 & 7 \\ 8 & 1 & 4 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{C}).
$$

Man rechnet sofort nach, daß $A^* \circ A = \mathbb{1}_3$, daß A also orthogonal bzw. unitär ist, mit charakteristischem Polynom

$$\chi_A = t^3 - t^2 + t - 1 = (t-1)\cdot(t-i)\cdot(t+i).$$

Es gibt also eine unitäre Matrix $T \in \mathrm{U}(3)$ mit

$$T^* \circ A \circ T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix}.$$

Um T zu bestimmen, berechnen wir zunächst den Eigenraum $\mathrm{Eig}(A,1)$ und finden $\left(\frac{1}{3},\frac{2}{3},\frac{2}{3}\right)^t$ als ONB. Durch Einsetzen in das Gleichungssystem überzeugt man sich, daß $(4,-1+3i,-1-3i)^t$ eine Lösung von $(A-i\mathbb{1}_3)x = 0$ ist, und durch Normierung erhalten wir dann $\left(\frac{2}{3},\frac{-1+3i}{6},\frac{-1-3i}{6}\right)^t$ als ONB von $\mathrm{Eig}(A,i)$. Da A eine reelle Matrix ist, muß somit $-i$ gerade den konjugiert komplexen Vektor als Eigenvektor haben, d. h. $\left(\frac{2}{3},\frac{-1-3i}{6},\frac{-1+3i}{6}\right)^t$ ist eine ONB von $\mathrm{Eig}(A,-i)$.

Wir erhalten also

$$T = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{-1+3i}{6} & \frac{-1-3i}{6} \\ \frac{2}{3} & \frac{-1-3i}{6} & \frac{-1+3i}{6} \end{pmatrix} \in \mathrm{U}(3)$$

als Transformationsmatrix mit

$$T^{-1} \circ A \circ T = T^* \circ A \circ T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix}.$$

Bemerkung 13.25 (Spektralsatz für orthogonale Abbildungen)
Mit dem gleichen Beweis wie in Satz 13.22 kann man zeigen, daß eine orthogonale Abbildung $f \in \mathrm{O}(V)$, deren charakteristisches Polynom über $\mathbb{R}$ in Linearfaktoren zerfällt, ebenfalls bezüglich einer ONB diagonalisierbar ist.

Orthogonale Abbildungen lassen sich im allgemeinen aber nicht diagonalisieren, insbesondere nicht durch eine ONB. Wir haben in Beispiel 9.15 gesehen, daß die Matrix

$$T(\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix},$$

die eine Drehung um den Ursprung um den Winkel α beschreibt, i.a. nicht diagonalisierbar ist. Man kann zeigen, daß dieses Beispiel im wesentlichen auch das einzige ist. Es gilt nämlich:

Ist $f \in \mathrm{End}_{\mathbb{R}}(V)$ orthogonal, dann gibt es eindeutig bestimmte Zahlen $r,s,t \in \mathbb{N}$ sowie Winkel $\alpha_1,\ldots,\alpha_t \in (0,2\pi)\setminus\{\pi\}$ und eine ONB B von V, so daß

$$M_B^B(f) = \mathbb{1}_r \oplus -\mathbb{1}_s \oplus T(\alpha_1) \oplus \ldots \oplus T(\alpha_t).$$

D) Spektralsatz für selbstadjungierte Endomorphismen

Definition 13.26 (Selbstadjungierter Endomorphismus)
$f \in \mathrm{End}_{\mathbb{K}}(V)$ heißt *selbstadjungiert* oder *hermitesch*, wenn $f = f^*$ gilt.

Proposition 13.27 (Matrixdarstellung selbstadjungierter Endomorphismen)
Es sei $f \in \mathrm{End}_{\mathbb{K}}(V)$ und B eine ONB von V.
Genau dann ist f selbstadjungiert, wenn $M_B^B(f)$ symmetrisch bzw. hermitesch ist.

Beweis: Aus Korollar 13.3 wissen wir, daß $M_B^B(f^*) = M_B^B(f)^*$ gilt. Deshalb gilt

$$\begin{aligned} f \text{ selbstadjungiert } &\iff f = f^* \iff M_B^B(f) = M_B^B(f^*) = M_B^B(f)^* \\ &\iff M_B^B(f) \text{ symmetrisch bzw. hermitesch.} \end{aligned}$$

□

Korollar 13.28 (Selbstadjungiert impliziert normal.)
Ist $f \in \mathrm{End}_{\mathbb{K}}(V)$ selbstadjungiert, so ist f normal.

Beweis: Ist f selbstadjungiert, so gilt $f \circ f^* = f \circ f = f^* \circ f$, also ist f normal. □

Lemma 13.29 (Selbstadjungierte Endomorphismen haben nur reelle Eigenwerte.)
Ist $f \in \mathrm{End}_{\mathbb{K}}(V)$ selbstadjungiert, dann ist $\chi_f \in \mathbb{R}[t]$ und χ_f zerfällt über $\mathbb{R}$.
Insbesondere gilt, ist $\lambda \in \sigma(f)$ ein Eigenwert von f, dann ist $\lambda \in \mathbb{R}$.

Beweis: Ist B eine ONB, dann ist $A = M_B^B(f)$ symmetrisch bzw. hermitesch und es reicht zu zeigen, daß χ_A in $\mathbb{R}[t]$ liegt und über $\mathbb{R}$ zerfällt.

Hierfür machen wir uns zunutze, daß wir A auf alle Fälle als eine Matrix in $\mathrm{Mat}_n(\mathbb{C})$ auffassen können und daß $A = A^*$ gilt. Über $\mathbb{C}$ zerfällt das charakteristische Polynom von A, d. h. es gibt $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ mit

$$\chi_A = (t - \lambda_1) \cdots (t - \lambda_n).$$

Es reicht, zu zeigen, daß $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ gilt. Nun gibt es zu jedem λ_i aber einen Vektor $0 \neq x_i \in \mathbb{C}^n$ mit $Ax_i = \lambda_i x_i$. Wegen $A = A^* = \overline{A}^t$ gilt für diesen Vektor

$$\begin{aligned} \lambda_i \cdot \left(\overline{x_i}^t \circ x_i\right) &= \overline{x_i}^t \circ (\lambda_i x_i) = \overline{x_i}^t \circ (A x_i) = \overline{x_i}^t \circ A \circ x_i \\ &= \overline{x_i}^t \circ \overline{A}^t \circ x_i = \overline{A x_i}^t \circ x_i = \overline{\lambda_i x_i}^t \circ x_i = \overline{\lambda_i} \cdot \left(\overline{x_i}^t \circ x_i\right). \end{aligned}$$

Aus $\overline{x_i}^t \circ x_i \neq 0$ folgt dann $\lambda_i = \overline{\lambda_i}$, d. h. $\lambda_i \in \mathbb{R}$. □

Satz 13.30 (Spektralsatz für selbstadjungierte Endomorphismen)
Ein Endomorphismus $f \in \mathrm{End}_{\mathbb{K}}(V)$ ist genau dann selbstadjungiert, wenn V eine ONB aus Eigenvektoren von f besitzt und alle Eigenwerte von f reell sind.

Beweis: Ist f selbstadjungiert, so ist f nach Korollar 13.28 normal und nach Lemma 13.29 zerfällt χ_f über $\mathbb{R}$ in Linearfaktoren. Aus dem Spektralsatz für normale

Abbildungen folgt dann, daß V eine ONB aus Eigenvektoren von V besitzt und die Eigenwerte sind alle reell.

Besitzt umgekehrt V eine ONB $B = (x_1, \ldots, x_n)$ aus Eigenvektoren von V mit reellen Eigenwerten $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, so gilt

$$M_B^B(f) = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix} = \begin{pmatrix} \overline{\lambda_1} & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \overline{\lambda_n} \end{pmatrix} = M_B^B(f)^*.$$

Mithin ist $M_B^B(f)$ symmetrisch bzw. hermitesch, und somit ist f selbstadjungiert. □

Korollar 13.31 (Spektralsatz für symmetrische und hermitesche Matrizen)
Zu jeder symmetrischen bzw. hermiteschen Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ *gibt es eine Matrix* $T \in \mathrm{O}(n)$ *bzw.* $T \in \mathrm{U}(n)$ *mit*

$$T^{-1} \circ A \circ T = T^* \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & \lambda_n \end{pmatrix}$$

und $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. *Insbesondere ist jede symmetrische bzw. hermitesche Matrix diagonalisierbar und hat nur reelle Eigenwerte.*

Beweis: Die Aussage folgt aus dem Spektralsatz für selbstadjungierte Endomorphismen über $\mathbb{K}^n$ mit kanonischem Skalarprodukt angewendet auf f_A, wenn wir $T = T_E^B$ für die dortige ONB B wählen. □

Dies ist eine wichtige Ergänzung des Satzes über die Jordansche Normalform.

Beispiel 13.32
Wir betrachten $\mathbb{C}^n$ mit dem kanonischen Skalarprodukt sowie die Matrix

$$A = \begin{pmatrix} 0 & -1 & i \\ -1 & 0 & -i \\ -i & i & 0 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{C}).$$

Da $A = A^*$ gilt, ist A hermitesch, und wir wollen mit Hilfe des Algorithmus 13.12 eine unitäre Transformationsmatrix berechnen, die A auf Diagonalgestalt transformiert.

Dazu bestimmen wir zunächst das charakteristische Polynom von A als

$$\chi_A = \begin{vmatrix} t & 1 & -i \\ 1 & t & i \\ i & -i & t \end{vmatrix} = t^3 - 3t - 2 = (t-2) \cdot (t+1)^2.$$

Da A diagonalisierbar ist, wissen wir nun schon, daß

$$\dim_{\mathbb{C}} \mathrm{Eig}(A, 2) = \mathrm{mult}(\chi_A, 2) = 1$$

und

$$\dim_{\mathbb{C}} \mathrm{Eig}(A, -1) = \mathrm{mult}(\chi_A, -1) = 2$$

gilt. Mit Hilfe des Gauß-Algorithmus können wir dann Basen der Eigenräume von A zu den Eigenwerten 2 und -1 berechnen. Die Rechnung wollen wir hier nicht vorführen, sondern nur das Ergebnis angeben:

$$B' = \big((-i, i, -1)^t\big)$$

ist eine Basis von $\mathrm{Eig}(A, 2)$ und

$$B'' = \big((1, 1, 0)^t, (0, i, 1)^t\big)$$

ist eine Basis von $\mathrm{Eig}(A, -1)$.

Dann müssen wir B' und B'' mittels des Gram-Schmidtschen-Orthonormalisierungsverfahrens in Orthonormalbasen der jeweiligen Eigenräume überführen. Bei B' ist das sehr einfach, da wir den einzigen Vektor in B' nur normieren müssen. Wir erhalten als einzigen Vektor in der ONB von $\mathrm{Eig}(A, 2)$ deshalb

$$z_1 = \left(\tfrac{-i}{\sqrt{3}}, \tfrac{i}{\sqrt{3}}, \tfrac{-1}{\sqrt{3}}\right)^t.$$

Für $B'' = (x_2, x_3)$ ist es etwas mehr Aufwand. Wir setzen zunächst

$$z_2 = \frac{1}{||x_2||} \cdot x_2 = \left(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0\right)^t.$$

Als nächstes setzen wir

$$y_3 = x_3 - \langle z_2, x_3\rangle \cdot z_2 = (0, i, 1)^t - \tfrac{i}{\sqrt{2}} \cdot \left(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0\right)^t = \left(\tfrac{-i}{2}, \tfrac{i}{2}, 1\right)^t$$

und normieren diesen Vektor anschließend zu

$$z_3 = \frac{1}{||y_3||} \cdot y_3 = \left(\tfrac{-i}{\sqrt{6}}, \tfrac{i}{\sqrt{6}}, \tfrac{2}{\sqrt{6}}\right)^t.$$

Die Vektoren z_2 und z_3 bilden eine ONB von $\mathrm{Eig}(A, -1)$, und $B = (z_1, z_2, z_3)$ ist somit eine ONB von $\mathbb{C}^3$. Schreiben wir die Vektoren als Spalten in die Matrix T, so erhalten wir die gesuchte Transformationsmatrix

$$T = \begin{pmatrix} \frac{-i}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{6}} \\ \frac{i}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{6}} \\ \frac{-1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix} \in \mathrm{U}(3).$$

Man rechnet folgendes leicht nach:

$$T^* \circ A \circ T = \begin{pmatrix} \frac{i}{\sqrt{3}} & \frac{-i}{\sqrt{3}} & \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{i}{\sqrt{6}} & \frac{-i}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{pmatrix} \circ \begin{pmatrix} 0 & -1 & i \\ -1 & 0 & -i \\ -i & i & 0 \end{pmatrix} \circ \begin{pmatrix} \frac{-i}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{6}} \\ \frac{i}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{6}} \\ \frac{-1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

E) Positiv definite symmetrische und hermitesche Matrizen

Eine symmetrische Bilinearform auf $\mathbb{R}^n$ bzw. eine hermitesche Sesquilinearform auf $\mathbb{C}^n$ ist genau dann ein Skalarprodukt, wenn sie positiv definit ist. Es ist daher außerordentlich nützlich, Kriterien für die positive Definitheit zur Hand zu haben.

Definition 13.33 (Definite Matrizen)

a. Eine symmetrische oder hermitesche Matrix $A \in \mathrm{Mat}_n(\mathbb{R})$ heißt *positiv definit* bzw. *negativ definit*, wenn für alle $0 \neq x \in \mathbb{K}^n$ der Wert $x^* \circ A \circ x$ stets positiv bzw. stets negativ ist.
   Sie heißt *indefinit*, wenn es Vektoren $x, y \in \mathbb{K}^n$ gibt mit
   $$x^* \circ A \circ x > 0 > y^* \circ A \circ y$$
b. Sei $A \in \mathrm{Mat}_n(\mathbb{K})$ und entsteht die $k \times k$-Untermatrix $A(k)$ von A durch Streichen der letzten $n-k$ Zeilen und Spalten, so nennen wir $A(k)$ die k-te *Hauptmatrix* von A und $\det\big(A(k)\big)$ den k-ten *Hauptminor* von A.

Bemerkung 13.34 (Negativ definite Matrizen)
Man beachte, daß eine symmetrische bzw. hermitesche Matrix A genau dann positiv definit ist, wenn die zugehörige Bilinearform bzw. die zugehörige Sesquilinearform positiv definit ist. Die analoge Aussage für negative Definitheit und Indefinitheit gilt auch.

Zudem folgt aus der Definition unmittelbar daß A genau dann negativ definit ist, wenn $-A$ positiv definit ist. Es reicht deshalb, ein Kriterium für positive Definitheit zu finden, um zugleich ein Kriterium für negative Definitheit zu erhalten, indem man A durch $-A$ ersetzt.

Lemma 13.35 (Positiv definite Matrizen)
Sei $A \in \mathrm{Mat}_n(\mathbb{K})$ symmetrisch bzw. hermitesch und $T \in \mathrm{Gl}_n(\mathbb{K})$ invertierbar. Genau dann ist A positiv definit, wenn $T^* \circ A \circ T$ positiv definit ist.

Beweis: Wir beachten, daß
$$\mathbb{K}^n \setminus \{0\} = \{T \circ x \mid 0 \neq x \in \mathbb{K}^n\} \tag{51}$$
gilt, da T invertierbar ist, und daß die Matrix $T^* \circ A \circ T$ symmetrisch bzw. hermitesch ist, wegen
$$(T^* \circ A \circ T)^* = T^* \circ A^* \circ T^{**} = T^* \circ A \circ T.$$
Wir erhalten deshalb
$$\begin{aligned} A \text{ ist positiv definit } &\Longleftrightarrow \overline{x}^t \circ A \circ x > 0 \quad \forall\, x \in \mathbb{K}^n \\ &\overset{(51)}{\Longleftrightarrow} \overline{T \circ x}^t \circ A \circ (T \circ x) > 0 \quad \forall\, 0 \neq x \in \mathbb{K}^n \\ &\Longleftrightarrow \overline{x}^t \circ T^* \circ A \circ T \circ x > 0 \quad \forall\, 0 \neq x \in \mathbb{K}^n \\ &\Longleftrightarrow T^* \circ A \circ T \text{ ist positiv definit} \end{aligned}$$

□

Satz 13.36 (Hurwitz-Kriterium für positive Definitheit)
Für eine symmetrische bzw. hermitesche Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ *sind die folgenden Aussagen gleichwertig:*

a. A *ist positiv definit.*

b. *Alle Eigenwerte von* A *sind positiv.*

c. *Alle Hauptminoren von* A *sind positiv.*

Beweis:
a. $\Longleftrightarrow$ b.: Wegen des Spektralsatzes für symmetrische und hermitesche Matrizen 13.31 gibt es eine orthogonale bzw. unitäre Matrix $T \in \mathrm{Mat}_n(\mathbb{K})$, so daß

$$T^* \circ A \circ T = T^{-1} \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \tag{52}$$

gilt, wobei $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ genau die Eigenwerte von A sind.

Ist A positiv definit und ist x_i die i-te Spalte von T, so folgt aus (52)

$$\lambda_i = \overline{x_i}^t \circ A \circ x_i > 0.$$

Seien nun umgekehrt alle Eigenwerte positiv und sei $0 \neq x \in \mathbb{K}^n$ gegeben. Die Spaltenvektoren $x_1, \ldots, x_n$ von T sind eine ONB $B = (x_1, \ldots, x_n)$ von $\mathbb{K}^n$, da T orthogonal bzw. unitär ist. Ist $M_B(x) = (\mu_1, \ldots, \mu_n)^t$ der Koordinatenvektor von x bezüglich B, so gilt

$$T \circ M_B(x) = \sum_{i=1}^{n} \mu_i \cdot x_i = x$$

und

$$\begin{aligned} \overline{x}^t \circ A \circ x &= \overline{T \circ M_B(x)}^t \circ A \circ (T \circ M_B(x)) \\ &= \overline{M_B(x)}^t \circ (T^* \circ A \circ T) \circ M_B(x) \\ &\overset{(52)}{=} \sum_{i=1}^{n} \overline{\mu_i} \cdot \mu_i \cdot \lambda_i = \sum_{i=1}^{n} |\mu_i|^2 \cdot \lambda_i > 0, \end{aligned}$$

da nicht alle μ_i null sind. Damit ist die Äquivalenz von a. und b. gezeigt.

a. $\Longrightarrow$ c.: Da wir die Äquivalenz von a. und b. bereits gezeigt haben, können wir hier beide Bedingungen voraussetzen. Wegen des Spektralsatzes für symmetrische und hermitesche Matrizen 13.31 gibt es eine Matrix $T \in \mathrm{Gl}_n(\mathbb{K})$, so daß (52) erfüllt ist, und deshalb gilt

$$\det(A) = \det(T^{-1} \circ A \circ T) = \lambda_1 \cdot \ldots \cdot \lambda_n > 0.$$

Jede positiv definite symmetrische oder hermitesche Matrix hat also eine positive Determinante.

Die k-te Hauptmatrix $A(k)$ beschreibt die Einschränkung der durch A definierten Bilinearform b_A bzw. Sesquilinearform b_A^s auf den Unterraum $\mathrm{Lin}\,(e_1,\ldots,e_k)$ von $\mathbb{K}^n$. Da die Einschränkung einer positiv definiten Bilinearform bzw. Sesquilinearform offenbar wieder positiv definit ist, ist mithin auch die definierende Matrix $A(k)$ positiv definit. Dann ist aber ihre Determinante, der k-te Hauptminor von A, positiv.

<u>**c. $\Longrightarrow$ a.:**</u> Wir führen den Beweis durch Induktion über n unter Ausnutzung der bereits gezeigten Äquivalenzen, wobei für $n=1$ die Determinante $\det(A)>0$ der einzige Eigenwert ist.

Sei also $n>1$. Wegen des Spektralsatzes für symmetrische und hermitesche Matrizen 13.31 existiert für die symmetrische bzw. hermitesche Matrix $A(n-1)$ eine orthogonale bzw. unitäre Matrix $S\in\mathrm{Gl}_{n-1}(\mathbb{K})$, die $A(n-1)$ auf Diagonalgestalt transformiert:

$$S^{-1}\circ A(n-1)\circ S=S^*\circ A(n-1)\circ S=\bigoplus_{i=1}^{n-1}\lambda_i\mathbb{1}_1=:D.$$

Da $A(n-1)$ die Induktionsvoraussetzung erfüllt, muß $A(n-1)$ dann positiv definit sein und somit sind die Eigenwerte $\lambda_1,\ldots,\lambda_{n-1}$ von $A(n-1)$ positiv.

Wir setzen nun $T=S\oplus\mathbb{1}_1\in\mathrm{Gl}_n(\mathbb{K})$. Dann gilt

$$T^*\circ A\circ T=\left(\begin{array}{ccc|c} & & & a_1\\ & D & & \vdots\\ & & & a_{n-1}\\ \hline \overline{a_1} & \ldots & \overline{a_{n-1}} & a_n\end{array}\right)=:B$$

für geeignete $a_1,\ldots,a_n\in\mathbb{K}$. Setzen wir ferner $c_j=-\frac{a_j}{\lambda_j}$, $j=1,\ldots,n-1$, und

$$C=\left(\begin{array}{ccc|c} & & & c_1\\ & \mathbb{1}_{n-1} & & \vdots\\ & & & c_{n-1}\\ \hline 0 & \ldots & 0 & 1\end{array}\right)\in\mathrm{Gl}_n(\mathbb{K}),$$

dann folgt

$$E:=(T\circ C)^*\circ A\circ(T\circ C)=C^*\circ T^*\circ A\circ T\circ C=C^*\circ B\circ C=\bigoplus_{i=1}^{n}\lambda_i\mathbb{1}_1,$$

wobei $\lambda_n\in\mathbb{R}$ geeignet ist. Man beachte, daß damit $\lambda_1,\ldots,\lambda_n$ die Eigenwerte von E sind, und daß

$$\lambda_1\cdot\ldots\cdot\lambda_n=\det\big((T\circ C)^*\circ A\circ(T\circ C)\big)=\det(A)\cdot|\det(C\circ T)|^2>0,$$

da $\det(A)>0$ der n-te Hauptminor von A ist. Da aber $\lambda_1,\ldots,\lambda_{n-1}$ nach Voraussetzung positiv waren, ist dann auch λ_n positiv. E hat somit nur positive Eigenwerte und ist deshalb positiv definit. Aber mit Lemma 13.35 ist dann auch A positiv definit. □

Bemerkung 13.37 (Negativ definite und indefinite Matrizen)
Wie im Beweis von “a. $\Longleftrightarrow$ b.” im Beweis von Satz 13.36 sieht man:

$$A \text{ ist } \textit{negativ definit} \quad\Longleftrightarrow\quad A \text{ hat nur negative Eigenwerte}$$

und

$$A \text{ ist } \textit{indefinit} \quad\Longleftrightarrow\quad A \text{ hat einen positiven und einen negativen Eigenwert.}$$

F) Hauptachsentransformationssatz für quadratische Formen

Definition 13.38 (Bilinearformen)
Es sei V ein $\mathbb{R}$-Vektorraum. Eine Abbildung

$$b : V \times V \to \mathbb{R},$$

die linear in beiden Argumenten ist, nennen wir *bilinear* oder eine *Bilinearform*, d. h. für $x, y, z \in V$ und $\lambda, \mu \in \mathbb{R}$ gilt (vgl. Definition 8.8):

$$b(\lambda x + \mu y, z) = \lambda b(x, z) + \mu b(y, z)$$

und

$$b(z, \lambda x + \mu y) = \lambda b(z, x) + \mu b(z, y).$$

Gilt zudem

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

für alle $x, y \in V$, so heißt b eine *symmetrische* Bilinearform, und wir nennen die Abbildung

$$q_b : V \longrightarrow \mathbb{R} : x \mapsto b(x, x)$$

die zu b gehörende *quadratische Form.*

Bemerkung 13.39 (Polarisierung einer Bilinearform)
Ist $b : V \times V \longrightarrow \mathbb{R}$ eine symmetrische Bilinearform, dann gilt

$$b(x, y) = \tfrac{1}{2}\big(q_b(x + y) - q_b(x) - q_b(y)\big)$$

für $x, y \in V$, wie man durch Einsetzen der Definition leicht nachrechnet. Die Bilinearform ist durch die zugehörige quadratische Form also eindeutig bestimmt. Man nennt die Gleichung die *Polarisierung* der Bilinearform.

Beispiel 13.40
Jedes Skalarprodukt auf einem $\mathbb{R}$-Vektorraum ist eine symmetrische Bilinearform. Zudem definiert jede symmetrische Matrix $A \in \mathrm{Mat}_n(\mathbb{R})$ durch

$$b_A : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R} : (x, y) \mapsto x^t \circ A \circ y$$

eine symmetrische Bilinearform. Die zu b_A gehörende quadratische Form bezeichnen wir mit

$$q_A : \mathbb{R}^n \longrightarrow \mathbb{R} : x \mapsto x^t \circ A \circ x.$$

Satz 13.41 (Hauptachsentransformationssatz für quadratische Formen)
Ist $A \in \mathrm{Mat}_n(\mathbb{R})$ *eine symmetrische Matrix, so gibt es eine ONB* $B = (x_1, \ldots, x_n)$ *von* $\mathbb{R}^n$ *zum kanonischen Skalarprodukt mit*

$$A \circ x_i = \lambda_i \cdot x_i$$

für $i = 1, \ldots, n$ *und*

$$q_A(x) = \sum_{i=1}^{n} \lambda_i \cdot \langle x_i, x \rangle^2.$$

Die x_i *werden die* Hauptachsen *der quadratischen Form genannt.*

Beweis: Aus dem Spektralsatz für selbstadjungierte Endomorphismen 13.30 wissen wir, daß es eine ONB $B = (x_1, \ldots, x_n)$ von $\mathbb{R}^n$ aus Eigenvektoren von f_A gibt. Diese wählen wir, so daß

$$A \circ x_i = f_A(x_i) = \lambda_i \cdot x_i$$

für geeignete $\lambda_i \in \mathbb{R}$ gilt. Sei $x \in V$, so folgt aus der Parsevalschen Gleichung 12.17

$$x = \sum_{i=1}^{n} \langle x_i, x \rangle \cdot x_i.$$

Für die quadratische Form q_A ausgewertet in x ergibt sich dann

$$\begin{aligned} q_A(x) =& x^t \circ A \circ x = x^t \circ A \circ \sum_{i=1}^{n} \langle x_i, x \rangle \cdot x_i = \sum_{i=1}^{n} \langle x_i, x \rangle \cdot x^t \circ A \circ x_i \\ =& \sum_{i=1}^{n} \langle x_i, x \rangle \cdot x^t \circ \lambda_i \cdot x_i = \sum_{i=1}^{n} \langle x_i, x \rangle \cdot \lambda_i \cdot \langle x, x_i \rangle = \sum_{i=1}^{n} \lambda_i \cdot \langle x, x_i \rangle^2. \end{aligned}$$

□

Bemerkung 13.42 (Geometrische Interpretation der Hauptachsentransformation)
Ist $A \in \mathrm{Mat}_n(\mathbb{R})$ eine positiv definite symmetrische Matrix mit quadratischer Form q_A, so interessieren wir uns für die Einheitssphäre zu q_A

$$S_{q_A} = \{y \in \mathbb{R}^n \mid q_A(y) = y^t \circ A \circ y = 1\}.$$

Man beachte, daß die Bilinearform b_A in diesem Fall ein Skalarprodukt ist und daß S_{q_A} die Menge der Vektoren in $\mathbb{R}^n$ ist, die bezüglich der zu diesem Skalarprodukt gehörenden Norm die Länge 1 haben.

Der Spektralsatz für selbstadjungiert Endomorphismen liefert die Existenz einer orthogonalen Matrix $T \in \mathrm{O}(n)$, so daß

$$T^t \circ A \circ T = \lambda_1 \cdot \mathbb{1}_1 \oplus \ldots \oplus \lambda_n \cdot \mathbb{1}_1$$

eine Diagonalmatrix ist, bei der die $\lambda_i > 0$ die Eigenwerte von A sind. Die Spaltenvektoren von T sind dann ein neues orthonormales Koordinatensystem, in dem die quadratische Form die Gestalt

$$q = \lambda_1 \cdot y_1^2 + \ldots + \lambda_n \cdot y_n^2$$

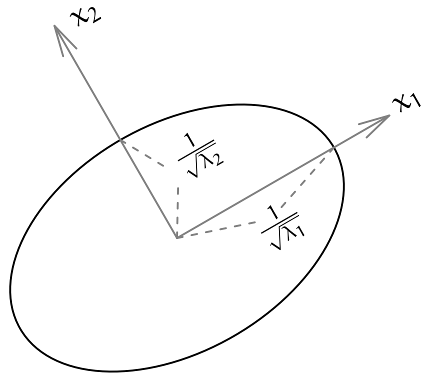


Abbildung 5. Ellipse S_{q_A} mit Hauptachsen $x_1 = Te_1$ und $x_2 = Te_2$

hat. Die Einheitssphäre zu q ist dann ein n-dimensionales Ellipsoid

$$S_q = \{y \in \mathbb{R}^n \mid q(y) = \lambda_1 \cdot y_1^2 + \ldots + \lambda_n \cdot y_n^2 = 1\}.$$

Man kann sich S_{q_A} deshalb als eine Einheitskugel vorstellen, die in Richtung $x_i = Te_i$ um den Faktor $\frac{1}{\sqrt{\lambda_i}}$ gestreckt wurde. Die Koordinatenvektoren x_i sind dann die *Hauptachsen* des Ellipsoids (siehe Abbildung 5). Im Fall $n = 2$ besagt der Satz der Hauptachsentransformation dann, daß wir allein durch Drehen und Spiegeln die Ellipse S_{q_A} so bewegen können, daß ihre Hauptachsen mit den Koordinatenachsen zusammenfallen. Daher rührt der Begriff der *Hauptachsentransformation*.

Aufgaben

Aufgabe 13.43
Zeige, wenn für $f \in \mathrm{End}_{\mathbb{K}}(V)$ stets $\|f(x)\| = \|x\|$ gilt, so ist f orthogonal bzw. unitär.

Aufgabe 13.44 (Lineare Funktionale)
Es sei V ein endlich-dimensionaler euklidischer oder unitärer Raum.
Dann gibt es für jedes $g \in \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$ genau ein $y \in V$, so daß für alle $x \in V$ gilt

$$g(x) = \langle y, x \rangle.$$

Aufgabe 13.45 (Die adjungierte Abbildung)
Seien V und W zwei endlich-dimensionale euklidische oder unitäre Räume mit Skalarprodukten $\langle \cdot, \cdot \rangle_V$ und $\langle \cdot, \cdot \rangle_W$. Dann gibt es zu jeder linearen Abbildung $f : V \longrightarrow W$ genau eine lineare Abbildung $f^* : W \longrightarrow V$, so daß

$$\langle f(x), y \rangle_W = \langle x, f^*(y) \rangle_V \tag{53}$$

für alle $x \in V$ und $y \in W$. Die Abbildung f^* heißt die *adjungierte Abbildung* zu f.

Aufgabe 13.46 (Matrixdarstellung der adjungierten Abbildung)
Seien V und W zwei endlich-dimensionale euklidische oder unitäre Räume mit Orthonormalbasen B bzw. D. Dann gilt für jede $\mathbb{K}$-lineare Abbildung $f : V \longrightarrow W$

$$M_B^D(f^*) = \left(M_D^B(f)\right)^*,$$

d.h. die Matrixdarstellung der adjungierten Abbildung ist die Adjungierte der Matrixdarstellung.

Aufgabe 13.47
Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum.
Zeige, ist $f \in \mathrm{End}_{\mathbb{K}}(V)$ normal, so gelten

$$\mathrm{Ker}(f) = \mathrm{Ker}(f^*)$$

und

$$V = \mathrm{Ker}(f) \perp \mathrm{Im}(f).$$

Aufgabe 13.48
Sei V ein endlich-dimensionaler unitärer Raum und $f \in \mathrm{End}_{\mathbb{C}}(V)$ normal.
Zeige, es gibt ein Polynom $p \in \mathbb{C}[t]$ mit $f^* = p(f)$.

Aufgabe 13.49
Es sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $f \in \mathrm{End}_{\mathbb{K}}(V)$ bijektiv. Zeige, die folgenden Aussagen sind äquivalent:

a. Für $x, y \in V$ mit $x \perp y$ gilt $f(x) \perp f(y)$.
b. Für $x, y \in V$ mit $||x|| = ||y||$ gilt $||f(x)|| = ||f(y)||$.
c. Es gibt ein $\lambda \in \mathbb{R}_{>0}$ und ein $g \in \mathrm{O}(V)$ bzw. $g \in \mathrm{U}(V)$ mit $f = \lambda g$.

Aufgabe 13.50
Es sei $V \neq 0$ ein endlich-dimensionaler unitärer Raum und $f \in \mathrm{End}_{\mathbb{C}}(V)$. Zeige, die folgenden Aussagen sind gleichwertig:

a. $f^* = -f$.
b. Für alle $x \in V$ gilt: $\langle f(x), x\rangle \in i\mathbb{R}$.
c. Es gibt eine Orthonormalbasis von V aus Eigenvektoren von f und der Realteil aller Eigenwerte ist Null.

Aufgabe 13.51
Überprüfe, ob die folgende symmetrische Matrix $A \in \mathrm{Mat}_3(\mathbb{R})$ positiv definit ist:

$$A = \begin{pmatrix} 9 & 1 & 2 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}.$$

Aufgabe 13.52
Bestimme eine orthogonale Matrix $T \in \mathrm{O}(3)$, die die folgende symmetrische Matrix

A diagonalisiert:

$$A = \begin{pmatrix} 5 & 2 & 2 \\ 2 & 2 & -4 \\ 2 & -4 & 2 \end{pmatrix}.$$

Ist die Matrix A positiv definit?

ANHANG A

Grundlegende Begriffe

§ A1 Etwas Logik

Wie alle Wissenschaftler versuchen auch die Mathematiker *Aussagen* über die Objekte ihrer Forschungsarbeit aufzustellen und *als wahr nachzuweisen.* Anders aber als etwa in den Naturwissenschaften werden die zu untersuchenden Objekte nicht von außen an die Mathematiker herangetragen, vielmehr schaffen sie sie sich selbst durch die Vorgabe sogenannter *Axiome.* Wie hat man dies zu verstehen? Was ist ein Axiom? Was heißt es, eine Aussage als wahr nachzuweisen? Und was eigentlich ist eine Aussage?

Nun, sobald wir uns auf eine Sprache geeinigt haben, in der wir uns verständigen wollen, sind wir in der Lage, Sätze zu bilden, Sätze, wie etwa (in unserer Alltagssprache)

"Dieser Satz enthält fünf Worte."

oder

"Löse die folgende Aufgabe."

Ein solcher Satz stellt eine *Aussage* in unserem Sinne dar, wenn wir entscheiden können, ob er wahr oder falsch ist. Gemäß dieser Konvention ist der erste der obigen Sätze eine – wahre – Aussage, während beim zweiten Satz, einer Aufforderung, die Frage nach wahr oder falsch wenig Sinn ergibt. Er ist mithin keine Aussage. Wir halten fest:

> Aussagen erkennen wir daran, daß ihnen ein Wahrheitswert zugeordnet ist, **w** für *wahr* oder **f** für *falsch.*

Im folgenden werden wir als Platzhalter für Aussagen meist Großbuchstaben verwenden: $\mathsf{A}, \mathsf{B}, \mathsf{C}, \ldots$.

Eine Aussage als *wahr nachzuweisen*, soll bedeuten, daß wir sie durch logische Schlüsse auf andere, uns als wahr bekannte Aussagen zurückführen. Nehmen wir etwa den folgenden Satz:

A : Der Bundespräsident ist stets mindestens vierzig Jahre alt.

Wir stellen zunächst einmal fest, daß es sich um eine Aussage handelt – und zwar um eine *wahre* Aussage, wie wir aus Artikel 54 des Grundgesetzes ableiten. Dort nämlich finden wir zur Wahl des Bundespräsidenten folgende Aussage:

B : Wählbar ist jeder Deutsche, der das Wahlrecht zum Bundestage besitzt und das vierzigste Lebensjahr vollendet hat.

Weil nun das Grundgesetz gültig ist, ist Aussage A wahr. Wir haben Aussage A also auf eine uns bekannte wahre Aussage zurückgeführt.

Daß die von uns aus dem Grundgesetz zitierte Aussage B ihrerseits wahr ist, läßt sich nicht weiter auf andere Aussagen zurückführen. Vielmehr handelt es sich hierbei um eine Festlegung des Gesetzgebers, der das Gesetz erlassen und damit diese Aussage für wahr erklärt hat.

> Eine Aussage, der der Wahrheitswert **w** schlicht durch Festlegung zugewiesen wurde, nennen wir ein *Axiom*.

Man kann in diesem Sinne das Grundgesetz als eine Sammlung von Axiomen, oder ein Axiomensystem, auffassen – auch wenn der Vergleich in mancher Hinsicht hinken mag.

Eingangs haben wir erklärt, daß die Mathematiker sich die Welt, die sie untersuchen, und ihre Objekte selbst erschaffen. Sie tun dies, indem sie sich einige wenige Aussagen als Axiome vorgeben und sodann studieren, was sich aus diesen durch logisch korrekte Schlüsse ableiten läßt. Freilich, so wie der Gesetzgeber seine Gesetze nicht willkürlich erläßt, so wählen auch die Mathematiker die Axiome, die sie sich vorgeben, mit Bedacht, das heißt, mit dem Ziel, interessante Strukturen zu gewinnen – und die vielfältigen Anwendungen zeigen, daß die Mathematiker bei diesem Vorhaben nicht nur sehr kreativ, sondern auch sehr erfolgreich gewesen sind. Immer wieder haben sie sich von Fragestellungen der Alltagswelt inspirieren lassen, haben die Probleme auf wenige Kernpunkte reduziert und in ein (mathematisches) *Modell* übersetzt. Dabei bedeutet letzteres nichts anderes, als daß man die zu benutzende Sprache und die geltenden Axiome festlegt und daß man die Fragen in dieser neuen Sprache formuliert. Die Stärke dieser *Modellbildung* besteht nun darin, daß man innerhalb des Modells exakt und ohne Wenn und Aber feststellen kann, ob eine Aussage wahr ist oder nicht. Wahr ist sie stets dann, wenn sie durch eine ganze Reihe logisch korrekter Schlüsse aus den vorgegebenen Axiomen hervorgeht. Wann aber ist denn eine Aussage aus einer anderen durch einen *logisch korrekten Schluß* hervorgegangen?

Bevor wir uns dieser Frage erneut zuwenden, wollen wir klären, wie man aus gegebenen Aussagen überhaupt neue Aussagen gewinnen und so das Arsenal an Aussagen erweitern kann.

Eine ganz natürliche Möglichkeit ist die Verneinung oder *Negation* einer Aussage, etwa

$\neg\mathsf{A}$: Der Bundespräsident ist *nicht* stets vierzig Jahre alt.

Wir wollen generell die Negation einer Aussage X mit dem Symbol $\neg\mathsf{X}$ bezeichnen, und es sollte gelten, wenn X wahr ist, so ist $\neg\mathsf{X}$ falsch, und umgekehrt. Das heißt insbesondere, der Wahrheitswert von $\neg\mathsf{X}$ hängt nur vom Wahrheitswert von X ab. Dies erlaubt es uns, den Wahrheitswert von $\neg\mathsf{X}$ in Abhängigkeit des Wahrheitswertes von X in einer Tabelle festzuhalten:

X	$\neg\mathsf{X}$
w	**f**
f	**w**

Aus unserer Alltagssprache sind wir es gewohnt, mehrere Aussagen in auflistender Weise durch das Wort "und" miteinander zu verbinden. Betrachten wir etwa die folgenden Aussagen

C : Wählbar sind nur Deutsche, die das Wahlrecht zum Bundestag besitzen.

sowie

D : Wählbar sind nur Deutsche, die das vierzigste Lebensjahr vollendet haben.

Man erkennt unschwer, daß die Verknüpfung der Aussagen C und D durch "und" inhaltlich mit unserer obigen Aussage B übereinstimmt, und man spricht von der *Konjunktion* von C und D. Auch hier wollen wir wieder eine symbolische Schreibweise einführen. Sind X und Y zwei Aussagen, so schreiben wir für "X *und* Y" auch $\mathsf{X}\wedge\mathsf{Y}$. Wenn nun $\mathsf{X}\wedge\mathsf{Y}$ wieder eine Aussage ist, so muß ihr auch ein Wahrheitswert zugeordnet sein. Dabei sollte wohl $\mathsf{X}\wedge\mathsf{Y}$ nur dann wahr sein, wenn sowohl X als auch Y wahr sind. Wir können den Wahrheitswert von $\mathsf{X}\wedge\mathsf{Y}$ also wieder in Abhängigkeit von den Wahrheitswerten von X und Y in einer Tabelle, auch *Wahrheitstafel* genannt, festhalten.

X	Y	$\mathsf{X}\wedge\mathsf{Y}$
w	**w**	**w**
w	**f**	**f**
f	**w**	**f**
f	**f**	**f**

Ebenso ist uns aus unserem alltäglichen Gebrauch ein weiteres Bindewort bekannt, "oder", welches wir hier instrumentalisieren wollen. Sind X und Y wieder Aussagen, so werden wir gewöhnlich $\mathsf{X}\vee\mathsf{Y}$ statt "X *oder* Y" schreiben. Die so entstandene neue Aussage nennt man die *Disjunktion* von X und Y, und damit sie wahr ist, soll es uns reichen, daß eine der Aussagen X und Y wahr ist. Dies führt zur folgenden

Wahrheitstafel:

X	Y	$\mathsf{X} \vee \mathsf{Y}$
w	**w**	**w**
w	**f**	**w**
f	**w**	**w**
f	**f**	**f**

Man beachte, daß *oder* hier nicht das ausschließende *entweder oder* ist!

Die Aussage etwa, daß die Kinder unserer Bundestagsabgeordneten stets die deutsche *oder* eine andere Staatsangehörigkeit haben, ist wahr, weil sie nicht ausschließt, daß sie die deutsche und eine andere Staatsangehörigkeit haben.

Im Absatz zur Konjunktion heißt es, daß die Aussage B mit der Konjunktion der Aussagen C und D inhaltlich übereinstimme. Sprachlich sind beide Aussagen aber deutlich verschieden. Anstatt sie *gleich* zu nennen, wollen wir deshalb nur davon sprechen, daß B und $\mathsf{C} \wedge \mathsf{D}$ *gleichwertig* oder *äquivalent* sind. Dies soll zum Ausdruck bringen, daß sie den gleichen Wahrheitswert besitzen. Gehen wir einen Schritt weiter, so können wir eine neue Verknüpfung zweier Aussagen X und Y einführen, die *Äquivalenz* von X und Y, in Symbolen $\mathsf{X} \Leftrightarrow \mathsf{Y}$. Sie soll genau dann wahr sein, wenn X und Y den gleichen Wahrheitswert besitzen. Dies führt zu folgender Wahrheitstafel:

X	Y	$\mathsf{X} \Leftrightarrow \mathsf{Y}$
w	**w**	**w**
w	**f**	**f**
f	**w**	**f**
f	**f**	**w**

Ein kurzer Blick auf die bislang eingeführten Operationen zur Gewinnung neuer Aussagen aus gegebenen zeigt, daß die Wahrheitswerte der neuen Aussagen stets allein von den Wahrheitswerten der gegebenen Aussagen abhängen, und nicht von deren konkretem Inhalt.

Wir erlauben uns deshalb, eine letzte Verknüpfung von Aussagen, die *Implikation*, dadurch einzuführen, daß wir bei gegebenen Aussagen X und Y den Wahrheitswert der Aussage "X impliziert Y" oder "wenn X, dann Y", in Zeichen $\mathsf{X} \Rightarrow \mathsf{Y}$, festlegen:

X	Y	$\mathsf{X} \Rightarrow \mathsf{Y}$
w	**w**	**w**
w	**f**	**f**
f	**w**	**w**
f	**f**	**w**

(54)

Die Wortwahl legt nahe, daß die Aussage $\mathsf{X} \Rightarrow \mathsf{Y}$ es erlaubt, aus der Wahrheit von X Rückschlüsse auf die Wahrheit von Y zu ziehen. Dies kommt auch in den ersten beiden Zeilen der Wahrheitstafel zum Ausdruck, wird aber noch deutlicher, wenn wir zeigen, daß die Aussagen $\mathsf{X} \Rightarrow \mathsf{Y}$ und $\neg\mathsf{X} \vee \mathsf{Y}$ zueinander äquivalent sind. Ist dann nämlich X wahr, so ist $\neg X$ falsch. Damit $\neg X \vee Y$ wahr sein kann, muß mithin Y wahr sein. Dies läßt sich so interpretieren, daß sich bei wahrer Aussage X und korrekter Implikation $X \Rightarrow Y$ für Y nur die Möglichkeit ergibt, ebenfalls wahr zu sein.

In dieser Weise werden wir die Implikation immer wieder anwenden. Wir werden mit einer wahren Aussage starten und mittels einer logisch korrekten Argumentationskette Y aus X ableiten – sprich wir werden $X \Rightarrow Y$ als wahr erweisen. Damit haben wir dann zugleich die Wahrheit von Y bewiesen.

Die Gültigkeit der behaupteten Äquivalenz leiten wir durch eine Betrachtung der Wahrheitstafeln her. Es reicht, festzustellen, daß die Werte in den Spalten von $X \Rightarrow Y$ und von $\neg X \vee Y$ übereinstimmen:

X	Y	$\neg X$	$\neg X \vee Y$	$X \Rightarrow Y$
w	**w**	**f**	**w**	**w**
w	**f**	**f**	**f**	**f**
f	**w**	**w**	**w**	**w**
f	**f**	**w**	**w**	**w**

Die bisherigen Betrachtungen erläutern die ersten beiden Zeilen der Wahrheitstafel der Implikation. Mysteriöser sind auf den ersten Blick zweifellos die beiden letzten, erlauben sie es doch, aus einer falschen Aussage eine beliebige andere Aussage herzuleiten und den vorgenommenen Schluß als korrekt anzusehen. Widerstrebt uns das nicht zutiefst? Wir möchten an einem Beispiel, das auf ein wenig Schulwissen in Mathematik zurückgreift, verdeutlichen, daß die obige Festlegung sehr wohl Sinn macht. Will man etwa die Lösungen der Gleichung

$$x^2 - 2x = -1$$

finden, so wird man auf beiden Seiten der Gleichung zunächst die Zahl 1 addieren, um so auf der linken Seite den Ausdruck $(x-1)^2$ zu erhalten, ein Verfahren, welches als *quadratische Ergänzung* bekannt ist. Man leitet aus der Aussage $x^2-2x=-1$ die Aussage $x^2-2x+1=0$ her. Dieser Schluß läßt sich formulieren als die Implikation

$$\left(x^2 - 2x = -1\right) \quad \Rightarrow \quad \left(x^2 - 2x + 1 = 0\right).$$

Der Schluß, daß die Addition einer Zahl auf beiden Seiten einer Gleichung, die Gleichheit nicht zerstört, ist uns wohl vertraut und wir sehen ihn als korrekt an, unabhängig davon, was auf beiden Seiten der Gleichung steht. Wenden wir diesen Schluß nun auf eine andere Gleichung an, etwa auf die Gleichung $0 = 1$, so erhalten wir die Implikation

$$(0 = 1) \quad \Rightarrow \quad (0 + 1 = 1 + 1).$$

Die beiden Aussagen links und rechts des Implikationspfeiles sind offenbar falsch, der Schluß an sich ist jedoch nach dem eben Gesagten zulässig. Mithin sollte die Implikation den Wahrheitswert **w** tragen.

Ein Beispiel dafür, daß sich aus einer falschen Aussage durch einen korrekten Schluß auch eine wahre Aussage herleiten läßt, erhalten wir in analoger Weise, wenn wir uns vergegenwärtigen, daß die Gleichheit auch durch Multiplikation mit einer Zahl nicht zerstört wird. Dies führt dann zu der wahren Implikation

$$(0 = 1) \quad \Rightarrow \quad (0 \cdot 0 = 1 \cdot 0),$$

bei der die Aussage auf der linken Seite des Implikationspfeiles falsch ist, während die auf der rechten Seite wahr ist.

Wir halten fest:

> Der Wahrheitswert der Implikation $X \Rightarrow Y$ bewertet nur die Korrektheit des Schließens, nicht jedoch die Wahrheit der Aussagen X und Y.

Es sei deshalb jedem ans Herz gelegt, die Voraussetzungen, auf die er seine Aussagen gründet, genauestens auf ihren Wahrheitsgehalt zu prüfen! Sonst nützt auch noch so sauberes Schließen gar nichts.

Wir wollen den eingeführten Begriffsapparat nun an zwei Beispielen testen, die uns einige wichtige Erkenntnisse liefern werden.

Beispiel A1.1
Es seien X und Y zwei Aussagen.

a. Wir haben bereits bei der Definition der Äquivalenz davon gesprochen, daß $X \Leftrightarrow Y$ bedeuten solle, daß "X *genau dann* wahr ist, *wenn* Y wahr ist". Dies wollte verkürzt ausdrücken, "wenn X, dann Y" und "wenn Y, dann X". Wir behaupten deshalb, daß die Aussagen "$X \Leftrightarrow Y$" und "$(X \Rightarrow Y) \wedge (Y \Rightarrow X)$" äquivalent sind, mit anderen Worten, die Aussagen X und Y sind genau dann äquivalent, wenn Y aus X folgt und umgekehrt.
Diese Tatsache werden wir immer wieder verwenden, wenn wir die Äquivalenz zweier Aussagen beweisen wollen. Ihre Gültigkeit leiten wir wieder durch eine Betrachtung der Wahrheitstafeln her.

X	Y	$X \Rightarrow Y$	$Y \Rightarrow X$	$(X \Rightarrow Y) \wedge (Y \Rightarrow X)$	$X \Leftrightarrow Y$
w	w	w	w	w	w
w	f	f	w	f	f
f	w	w	f	f	f
f	f	w	w	w	w

b. Die Aussagen “$X \Rightarrow Y$” und “$\neg Y \Rightarrow \neg X$” sind ebenfalls äquivalent, wie die folgende Tabelle zeigt:

X	Y	$\neg X$	$\neg Y$	$X \Rightarrow Y$	$\neg Y \Rightarrow \neg X$
w	**w**	**f**	**f**	**w**	**w**
w	**f**	**f**	**w**	**f**	**f**
f	**w**	**w**	**f**	**w**	**w**
f	**f**	**w**	**w**	**w**	**w**

Man nennt diese Äquivalenz auch *Kontraposition*. Will man also zeigen, daß eine Aussage X eine Aussage Y impliziert, so kann man statt dessen beide Aussagen verneinen und zeigen, daß aus $\neg Y$ die Aussage $\neg X$ folgt.

□

Kehren wir nun zu der Frage zurück, wann eine Aussage Y aus einer Aussage X durch einen *logisch korrekten Schluß* hervorgegangen ist. Bedeutet dies nur, daß $X \Rightarrow Y$ den Wahrheitswert **w** besitzt? Ja ... und nein! Ist X wahr und hat die Implikation $X \Rightarrow Y$ den Wahrheitswert **w**, so folgt unmittelbar, daß Y wahr ist. In diesem Sinne gilt die Antwort *ja*. Aber damit haben wir das Problem nur verlagert, da die Frage bleibt, wie wir prüfen, ob $X \Rightarrow Y$ denn wahr ist, ohne den Wahrheitswert von Y zu kennen. Wir haben bereits weiter oben – sehr vage – angedeutet, daß wir hierzu meist eine Kette von logisch korrekten und in sich schlüssigen Argumenten verwenden, und viel deutlicher wollen wir hier auch nicht werden. Im Verlauf der folgenden Kapitel werden wir viele Beispiele dafür sehen, wie eine Implikation durch eine Reihe von Argumenten bewiesen – oder besser untermauert – wird; und es wird sicher immer wieder vorkommen, daß Euch diese auf den ersten Blick *nicht* wirklich schlüssig vorkommen, daß es eines genaueren Hinsehens und vielleicht auch der Ergänzung einiger Argumente bedarf, bis Ihr der Kette das Prädikat *logisch korrekt und in sich schlüssig* verleihen wollt. Und das ist eine wichtige Erkenntnis, ob ein Schluß als logisch korrekt erkannt wird, hängt vom Betrachter ab. Und deshalb ist die Frage, ob ein Schluß logisch korrekt ist, weit mehr als nur die Frage, ob $X \Rightarrow Y$ wahr ist.

Beispiel A1.2

Hier nun einige mathematische Aussagen.

A. Jede gerade Zahl ist Summe zweier ungerader Zahlen.

B. Es gibt unendlich viele Primzahlen.

C. Jede gerade Zahl größer zwei ist Summe zweier Primzahlen.

D. Zu jedem Kreis läßt sich, nur mit Zirkel und Lineal, ein Quadrat konstruieren, das den gleichen Flächeninhalt hat.

E. Die Gleichung $x^n + y^n = z^n$ besitzt für $n > 2$ keine Lösung mit positiven ganzen Zahlen x, y, z.

F. Gegeben sei eine Ansammlung nicht-leerer Mengen. Dann läßt sich aus jeder der Mengen ein Element auswählen.

Die Aussage A ist offensichtlich wahr, und auch die Aussage B ist richtig, allerdings ist dies keine triviale Aussage. Sie muß bewiesen werden. Die Aussage C ist die bekannte *Goldbachsche Vermutung* aus dem Jahre 1742. Sie ist bis heute weder bewiesen noch widerlegt.

Die Aussage D ist unter dem Begriff *Quadratur des Kreises* bekannt. Sie ist falsch, was sich daraus ableiten läßt, daß die Kreiszahl π transzendent ist (Lindemann 1882). Umgangssprachlich sollte man also die Quadratur des Kreises nicht als Synonym für etwas extrem Schwieriges verwenden, sondern für etwas Unmögliches.

Die Aussage E hat jahrhundertelang als *Fermatsche Vermutung* die Mathematiker beschäftigt. Sie wurde erst 1995 von dem englischen Mathematiker Andrew Wiles als wahr nachgewiesen. Für den Beweis wurden modernste und tiefste mathematische Methoden verwendet.

Die Aussage F, möchte man meinen, ist offensichtlich wahr, eher noch als Aussage A. In gewissem Sinne ist diese Aussage jedoch weder beweisbar noch widerlegbar. Sie ist im Axiomensystem der Mengenlehre von Zermelo und Fraenkel unabhängig von den anderen Axiomen. In der Tat kann man die Aussage F, die als *Auswahlaxiom* bezeichnet wird, als Axiom der Mengenlehre zulassen (was wir, wie die überwiegende Zahl der Mathematiker, tun wollen) oder auch nicht. Da das Auswahlaxiom, wenn überhaupt, so nur für sogenannte überabzählbare Ansammlungen strittig ist, sind Zustimmung oder Ablehnung in dieser Vorlesung kaum von praktischer Relevanz. □

Wir wollen nun der besseren Übersichtlichkeit halber in einer Bemerkung zusammenfassen, was wir bisher gelernt haben.

Bemerkung A1.3

a. Eine *Aussage* ist eine Äußerung, der eindeutig ein Wahrheitswert wahr (**w**) oder falsch (**f**) zugeordnet ist.

b. Aus Aussagen X und Y können wir durch Anwenden *logischer Operatoren* neue Aussagen bilden:

Symbol	Bedeutung	Bezeichnung	Alternative Beschreibung
$\neg X$	nicht X	*Negation*	
$X \vee Y$	X oder Y	*Disjunktion*	
$X \wedge Y$	X und Y	*Konjunktion*	
$X \Rightarrow Y$	aus X folgt Y	*Implikation*	$(\neg X) \vee Y$
$X \Leftrightarrow Y$	genau dann X, wenn Y	*Äquivalenz*	$(X \Rightarrow Y) \wedge (Y \Rightarrow X)$

Neben Aussagen, die wahr oder falsch sein können, sind *Aussageformen* oder *Prädikate* wichtig.

Eine *Aussageform* ist eine Äußerung, die eine oder mehrere Variablen enthält und zu einer Aussage (d.h. wahr oder falsch) wird, wenn man zulässige Werte für diese Variablen einsetzt.

So ist etwa

$$a > b$$

eine Aussageform, die von den Variablen a und b abhängt, für die wir die ganzen Zahlen als zulässige Werte ansehen wollen. Setzen wir konkrete Werte ein, so entsteht eine Aussage, die wahr sein kann (z.B. für $a = 42$ und $b = 37$) oder falsch (z.B. für $a = 2$ und $b = 4$).

Aussageformen werden in der Praxis häufig mit *Quantoren* gebraucht:

$\forall$: "für alle".

$\exists$: "es existiert ein".

$\exists_1$: "es existiert genau ein".

$\nexists$: "es existiert kein".

Ist P eine Aussageform, die von einer Variablen x abhängt, so bedeutet:

$\forall\, x\, :\, P(x)$: "für alle x gilt $P(x)$",

$\exists\, x\, :\, P(x)$: "es gibt ein x, so daß $P(x)$ gilt".

Mit Hilfe der Quantoren haben wir aus den Aussageformen neue Aussagen gebildet.

Beispiel A1.4

$$\forall\, x, \forall\, y, \forall\, z, \forall\, n\, :\, n > 2 \Rightarrow x^n + y^n \neq z^n.$$

Dies ist für positive natürliche Zahlen x, y, z und n die in Beispiel A1.2 formulierte Fermatsche Vermutung. □

Wichtig ist das richtige Verneinen einer Aussage.

$$\neg\big(\forall\, x\, :\, P(x)\big) \;\Leftrightarrow\; \exists\, x\, :\, \big(\neg P(x)\big).$$

Die Verneinung der Aussage "für alle x gilt die Aussage $P(x)$" ist gleichbedeutend mit "es gibt ein x, für das die Aussage $P(x)$ nicht gilt".

$$\neg\big(\exists\, x\, :\, P(x)\big) \;\Leftrightarrow\; \forall\, x\, :\, \big(\neg P(x)\big).$$

Die Verneinung der Aussage "es gibt ein x, für das die Aussage $P(x)$ gilt" ist gleichbedeutend mit "für alle x gilt die Aussage $P(x)$ nicht" bzw. mit "für kein x gilt die Aussage $P(x)$".

$$(A \Rightarrow B) \;\Leftrightarrow\; (\neg B \Rightarrow \neg A).$$

Die Aussage "aus A folgt B" ist gleichbedeutend mit "aus nicht B folgt nicht A". Letzteres bezeichnet man auch als *Kontraposition* von ersterem.

Proposition A1.5
Es seien X, Y *und* Z *Aussagen.*

a. Assoziativgesetze
   - $(\mathsf{X} \vee \mathsf{Y}) \vee \mathsf{Z} \iff \mathsf{X} \vee (\mathsf{Y} \vee \mathsf{Z})$.
   - $(\mathsf{X} \wedge \mathsf{Y}) \wedge \mathsf{Z} \iff \mathsf{X} \wedge (\mathsf{Y} \wedge \mathsf{Z})$.

b. Kommutativgesetze
   - $\mathsf{X} \vee \mathsf{Y} \iff \mathsf{Y} \vee \mathsf{X}$.
   - $\mathsf{X} \wedge \mathsf{Y} \iff \mathsf{Y} \wedge \mathsf{X}$.

c. Distributivgesetze
   - $\mathsf{X} \wedge (\mathsf{Y} \vee \mathsf{Z}) \iff (\mathsf{X} \wedge \mathsf{Y}) \vee (\mathsf{X} \wedge \mathsf{Z})$.
   - $\mathsf{X} \vee (\mathsf{Y} \wedge \mathsf{Z}) \iff (\mathsf{X} \vee \mathsf{Y}) \wedge (\mathsf{X} \vee \mathsf{Z})$.

Beweis: Den Nachweis der Äquivalenzen überlassen wir dem Leser als Übungsaufgabe. □

Bemerkung A1.6 (Griechisches Alphabet)
Es hat sich in der Mathematik eingebürgert, neben den lateinischen auch griechische Buchstaben zu verwenden, um Objekte und Variablen zu bezeichnen, und das werden wir immer wieder mal tun. Deshalb füge ich hier das griechische Alphabet an:

A α Alpha	B β Beta	Γ γ Gamma	Δ δ Delta	E ϵ ε Epsilon	Z ζ Zeta	H η Eta	Θ θ ϑ Theta
I ι Iota	K κ Kappa	Λ λ Lambda	M μ My	N ν Ny	Ξ ξ Xi	O o Omikron	Π π Pi
P ρ Rho	Σ σ Sigma	T τ Tau	Υ υ Ypsilon	Φ ϕ φ Phi	X χ Chi	Ψ ψ Psi	Ω ω Omega

Aufgaben

Aufgabe A1.7

a. Negiere die folgenden Aussagen:
   (i) Jedes Auto, das am Samstag um 9:00 auf dem Parkplatz parkte, war rot.
   (ii) Mindestens ein Auto, das am Samstag um 9:00 auf dem Parkplatz parkte, war rot.
   (iii) Am Samstag um 9:00 parkten rote Autos auf dem Parkplatz.
   (iv) Es gibt keine größte ganze Zahl.
   (v) Keine Regel ohne Ausnahme.

   Warum ist das Sprichwort „Keine Regel ohne Ausnahme“ in sich widersprüchlich?

b. Beweise oder widerlege Aussage (iv).

Aufgabe A1.8
Es seien X und Y Aussagen. Zeige die folgenden Äquivalenzen:

a. *De Morgansche Regeln*
   - $\neg(X \vee Y) \Longleftrightarrow \neg X \wedge \neg Y$.
   - $\neg(X \wedge Y) \Longleftrightarrow \neg X \vee \neg Y$.

b. $(\neg X \Longrightarrow f) \Longleftrightarrow X$.

Aufgabe A1.9

a. Drücke die folgenden Aussagen in Worten aus und, falls eine Aussage falsch sein sollte, ersetze sie dabei durch ihre Negation.
   (i) $\forall\, m \in \mathbb{N},\ \exists\, n \in \mathbb{N} \,:\, m = n + n$,
   (ii) $\exists\, m \in \mathbb{N},\ \exists\, n \in \mathbb{N} \,:\, (m \neq n) \wedge (m^n = n^m)$.

b. Drücke die folgenden Aussagen in Symbolen aus:
   (i) Zwischen je zwei verschiedenen reellen Zahlen gibt es eine weitere reelle Zahl.
   (ii) Es gibt keine größte Primzahl in den natürlichen Zahlen.

Aufgabe A1.10
Welche der folgenden Schlußfolgerungen ist korrekt?

a. Falls es anfängt zu regnen, wird die Straße naß. Aber, da die Straße nicht naß werden wird, wird es auch nicht regnen.

b. Einige Politiker sind ehrlich. Einige Frauen sind Politiker. Also sind einige weibliche Politiker ehrlich.

Aufgabe A1.11
Drücke die folgende Aussage in Worten aus:

$$\forall\, m \in \mathbb{N},\ \forall\, n \in \mathbb{N} \,:\, m \geq n \implies \exists\, l \in \mathbb{N} \,:\, m = n + l.$$

Aufgabe A1.12 a. Negiere die folgenden Aussagen:
   (i) Zu jedem Vorschlag gibt es jemanden, der den Vorschlag kritisiert.
   (ii) In manchen Häusern haben nicht alle Wohnungen fließendes Wasser.

b. Beweise oder widerlege die folgenden Aussagen:
   (i) Jede ganze Zahl ist ein Vielfaches von drei.
   (ii) Die Summe von je zwei ungeraden Zahlen ist gerade.

§ A2 Mengen

Definitionsversuch A2.1 (Georg Cantor)
Eine *Menge* ist eine Zusammenfassung von bestimmten, wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen. Die in einer Menge zusammengefaßten Objekte nennen wir die *Elemente* der Menge.

Notation A2.2

a. *Mengen angeben durch Auflisten der Elemente*:

$$\text{z.B. } \{1, 2, 5, 3, 4, 0\}$$

b. *Mengen angeben durch Vorschreiben einer Eigenschaft*:

$$\text{z.B. } \{x \mid x \text{ ist eine natürliche Zahl kleiner als } 6\}$$

c. Sei M eine Menge.
   - $x \in M$ heißt "x *ist Element von* M"
   - $x \notin M$ heißt "x *ist nicht Element von* M"

d. $\{\}$ und $\emptyset$ bezeichnen die *leere Menge*, d.h. die Menge, die kein Element enthält.

Definition A2.3 (Inklusionsrelationen)
Für zwei Mengen M und N definieren wir:

1) $M \subseteq N \;:\Longleftrightarrow\; (x \in M \Rightarrow x \in N)$ "M ist *Teilmenge* von N"

2) $$\begin{aligned} M = N \;&:\Longleftrightarrow\; (M \subseteq N \;\wedge\; N \subseteq M) \\ &\Longleftrightarrow\; (x \in M \;\Leftrightarrow\; x \in N) \end{aligned}$$

3) $$\begin{aligned} M \neq N \;&:\Longleftrightarrow\; \neg(M = N) \\ &\Longleftrightarrow\; ((\exists\, x \in M \;:\; x \notin N) \;\vee\; (\exists\, x \in N \;:\; x \notin M)) \end{aligned}$$

4) $M \subsetneqq N \;:\Longleftrightarrow\; (M \subseteq N \;\wedge\; M \neq N)$ "M ist *echte Teilmenge* von N"

Beispiel A2.4

a. $\{1, 2, 5, 3, 4, 0\} = \{x \mid x \text{ ist eine natürliche Zahl kleiner als } 6\}$.

b. $\{1, 3\} \subsetneqq \{1, 2, 3\}$.

c. $\{1, 2, 1\} = \{1, 2\} = \{2, 1\}$.

d. $1 \notin \{2, 3\}$, $2 \in \{2, 3\}$.

Bemerkung A2.5 (Die Zahlbereiche)
Wir setzen die folgenden Mengen in unserer Vorlesung als bekannt voraus:

- $\mathbb{N} = \{0, 1, 2, 3, 4, \ldots\}$ die Menge der *natürlichen Zahlen*,
- $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ die Menge der *ganzen Zahlen*,
- $\mathbb{Q} = \left\{\frac{p}{q} \;\middle|\; p, q \in \mathbb{Z}, q \neq 0\right\}$ die Menge der *rationalen Zahlen*,
- $\mathbb{R}$, die Menge der *reellen Zahlen*, d.h. der Dezimalbrüche.

Beachte:

$$\mathbb{N} \subsetneqq \mathbb{Z} \subsetneqq \mathbb{Q} \subsetneqq \mathbb{R}.$$

Im Verlauf der Vorlesung werden wir viele bekannte Eigenschaften dieser Mengen nochmals ausführlich thematisieren.

Definition A2.6 (Operationen von Mengen)
Es seien M, N, P sowie M_i für $i \in I$ Mengen.

a. $M \cap N := \{x \mid x \in M \wedge x \in N\}$ heißt der *Durchschnitt* von M und N.

b. $M \cup N := \{x \mid x \in M \vee x \in N\}$ heißt die *Vereinigung* von M und N.

c. $M \setminus N := \{x \mid x \in M \wedge x \notin N\}$ heißt die *Differenzmenge* von M und N. Wir sagen auch M *ohne* N.

d. $M \times N := \{(x, y) \mid x \in M \wedge y \in N\}$ heißt das *kartesische Produkt* von M und N. Dabei ist (x, y) ein *geordnetes Paar*, und für zwei geordnete Paare $(x, y), (u, v) \in M \times N$ gilt
$$(x, y) = (u, v) \iff (x = u \wedge y = v).$$

e. M und N heißen genau dann *disjunkt*, wenn $M \cap N = \emptyset$, d.h. wenn sie kein Element gemeinsam besitzen.

f. $P = M \dot{\cup} N \quad :\iff \quad (P = M \cup N \quad \wedge \quad M \cap N = \emptyset)$.
Wir sagen dann, P ist die *disjunkte Vereinigung* von M und N.

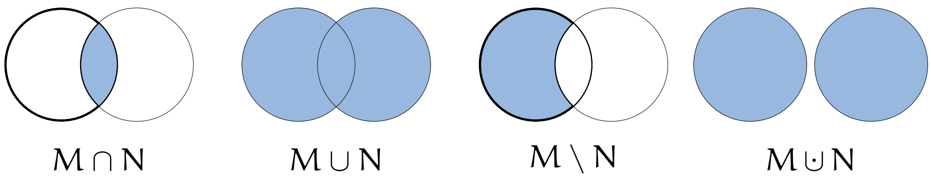


ABBILDUNG 1. Durchschnitt, Vereinigung, Differenzmenge, disjunkte Vereinigung

g. $\bigcap_{i \in I} M_i := \{x \mid x \in M_i \ \forall\, i \in I\}$ heißt der *Durchschnitt* der M_i.

h. $\bigcup_{i \in I} M_i := \{x \mid \exists\, i \in I \, : \, x \in M_i\}$ heißt die *Vereinigung* der M_i.

i. $P = \dot{\bigcup}_{i \in I} M_i \quad :\iff \quad (P = \bigcup_{i \in I} M_i \quad \wedge \quad M_i \cap M_j = \emptyset \ \forall\, i, j \in I$ mit $i \neq j)$.
Wir nennen die $(M_i)_{i \in I}$ dann auch eine *disjunkte Zerlegung* von P, und wir sagen, die M_i sind *paarweise disjunkt*.

Beispiel A2.7
Ist $M = \{1, 2\}$ und $N = \{e, \pi, i\}$, so ist
$$M \times N = \{(1, e), (1, \pi), (1, i), (2, e), (2, \pi), (2, i)\}.$$

Proposition A2.8 (Einfache Rechengesetze für Mengenoperationen)
Es seien M, N, P *Mengen.*

a. Assoziativgesetze
   - $(M \cup N) \cup P = M \cup (N \cup P)$.
   - $(M \cap N) \cap P = M \cap (N \cap P)$.

b. Kommutativgesetze
- $M \cup N = N \cup M$.
- $M \cap N = N \cap M$.

c. Distributivgesetze
- $M \cap (N \cup P) = (M \cap N) \cup (M \cap P)$.
- $M \cup (N \cap P) = (M \cup N) \cap (M \cup P)$.

d. Identitätsgesetze
- $M \cup \emptyset = M$.
- $M \subseteq N \implies M \cap N = M$.

e. Komplementgesetze
- $M \subseteq N \implies M \cup (N \setminus M) = N$.
- $M \subseteq N \implies M \cap (N \setminus M) = \emptyset$.

Beweis: a., d. und e. überlassen wir dem Leser als Übungsaufgabe.

b. Es gilt:

$$M \cup N = \{x \mid x \in M \vee x \in N\} \stackrel{\text{A1.5}}{=} \{x \mid x \in N \vee x \in M\} = N \cup M$$

und

$$M \cap N = \{x \mid x \in M \wedge x \in N\} \stackrel{\text{A1.5}}{=} \{x \mid x \in N \wedge x \in M\} = N \cap M.$$

c. Es gilt:

$$\begin{aligned} x \in M \cap (N \cup P) &\iff x \in M \ \wedge \ x \in N \cup P \\ &\iff x \in M \ \wedge \ (x \in N \ \vee \ x \in P) \\ &\stackrel{\text{A1.5}}{\iff} (x \in M \ \wedge \ x \in N) \ \vee \ (x \in M \ \wedge \ x \in P) \\ &\iff x \in M \cap N \ \vee \ x \in M \cap P \\ &\iff x \in (M \cap N) \cup (M \cap P) \end{aligned}$$

und

$$\begin{aligned} x \in M \cup (N \cap P) &\iff x \in M \ \vee \ x \in N \cap P \\ &\iff x \in M \ \vee \ (x \in N \ \wedge \ x \in P) \\ &\stackrel{\text{A1.5}}{\iff} (x \in M \ \vee \ x \in N) \ \wedge \ (x \in M \ \vee \ x \in P) \\ &\iff x \in M \cup N \ \wedge \ x \in M \cup P \\ &\iff x \in (M \cup N) \cap (M \cup P). \end{aligned}$$

□

Bemerkung A2.9 (Paradoxon von Russel)
Man muß bei der Definition von Mengen mittels Eigenschaften vorsichtig sein!

Betrachte die "*Menge*"

$$M = \{X \mid X \text{ ist Menge } \wedge X \notin X\}$$

aller Mengen, die sich nicht selbst als Element enthalten!

Angenommen, M wäre eine Menge. Dann sind zwei Fälle zu unterscheiden.

1. Fall: $M \notin M$**:** Dann ist M eine Menge, die sich nicht selbst als Element enthält. Mithin gilt $M \in M$ aufgrund der Definition von M. Dies ist ein Widerspruch.

2. Fall: $M \in M$**:** Dann ist M eine Menge, die sich selbst als Element enthält. Mithin gilt $M \notin M$ aufgrund der Definition von M. Dies ist ebenfalls ein Widerspruch.

Also kann keiner der beiden Fälle auftreten, und wir haben insgesamt einen Widerspruch hergeleitet.

Fazit: M ist *keine* Menge! Auch die *Menge aller Mengen* gibt es nicht!

Aufgaben

Aufgabe A2.10 (De Morgansche Regeln)
Es seien M und M_i, $i \in I$, Mengen. Zeige, die de Morganschen Regeln

$$M \setminus \bigcup_{i \in I} M_i = \bigcap_{i \in I} M \setminus M_i$$

und

$$M \setminus \bigcap_{i \in I} M_i = \bigcup_{i \in I} M \setminus M_i.$$

§ A3 Abbildungen

In diesem Abschnitt wollen wir den für die Mathematik zentralen Begriff der Abbildung einführen.

Definition A3.1 (Abbildungen)
Es seien M und N zwei Mengen. Eine *Abbildung* oder *Funktion* f von M nach N ist eine *eindeutige Zuordnung*, die *jedem* Element $x \in M$ *genau ein* Element $f(x) \in N$ zuweist. Wir werden den Begriff *Funktion* nur dann verwenden, wenn $N = \mathbb{R}$ ist.

Wir nennen M den *Definitionsbereich* von f und N den *Ziel-* oder *Wertebereich*.

Notation:

$$f : M \longrightarrow N : x \mapsto f(x).$$

Beachte, aufgrund der Definition einer Abbildung, gilt für zwei Abbildungen $f : M \longrightarrow N$ und $g : X \longrightarrow Y$:

$$f = g \quad \Longleftrightarrow \quad \big(M = X \ \wedge \ N = Y \ \wedge \ \forall\, x \in M \ : \ f(x) = g(x)\big).$$

Beispiel A3.2

a. Die folgenden Bilder sollen den Begriff der Abbildung graphisch veranschaulichen:

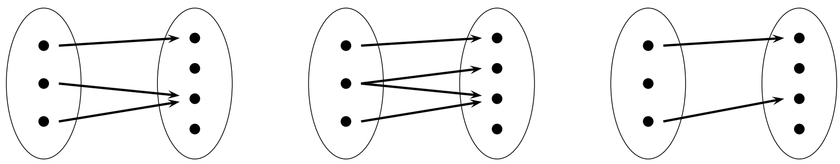


Abbildung — keine Abbildung — keine Abbildung

b. $f : \mathbb{N} \longrightarrow \mathbb{N} : x \mapsto x^2$ und $g : \mathbb{Z} \longrightarrow \mathbb{N} : x \mapsto x^2$. Beachte: $f \neq g$, da ihre Definitionsbereiche nicht übereinstimmen.

c. Sei $f : M \longrightarrow N$ eine Abbildung und $A \subseteq M$. Dann heißt die Abbildung

$$f_{|A} : A \longrightarrow N : x \mapsto f(x)$$

die *Einschränkung* von f auf A.

d. Sei M eine Menge. Dann heißt die Abbildung

$$\mathrm{id}_M : M \longrightarrow M : x \mapsto x$$

die *Identität* auf M.

Definition A3.3 (Bilder und Urbilder)
Es sei $f : M \longrightarrow N$ eine Abbildung, $A \subseteq M$ und $B \subseteq N$.

a. $\mathrm{Graph}(f) := \{(x, f(x)) \mid x \in M\} \subseteq M \times N$ heißt der *Graph* von f.

b. $f(A) := \{f(x) \mid x \in A\} \subseteq N$ heißt das *Bild* von A unter f.

c. $\mathrm{Im}(f) := f(M) \subseteq N$ heißt das *Bild* von f.

d. $f^{-1}(B) := \{x \in M \mid f(x) \in B\} \subseteq M$ heißt das *Urbild* von B unter f.

Beispiel A3.4

a. Wir betrachten die Abbildung $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^2$.

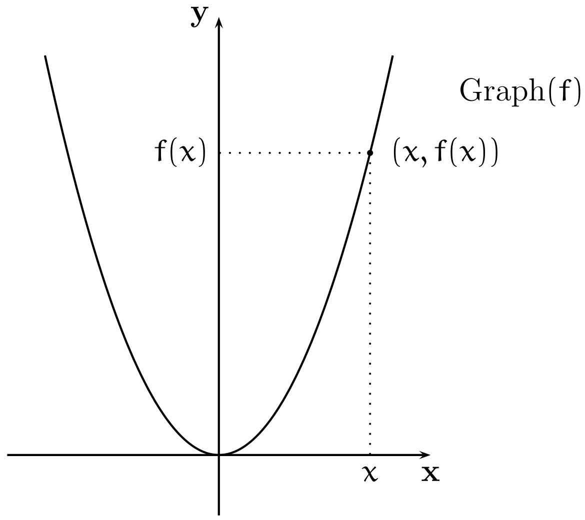


ABBILDUNG 2. Graph(f) für $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^2$

- Der Graph von f ist in Abbildung 2 zu sehen.
- Für $A = \{-1, 0, 1, 2\}$ ist $f(A) = \{0, 1, 4\}$.
- Für $B = \{0, 1\}$ ist $f^{-1}(B) = \{0, 1, -1\}$.
- Für $B' = \{-1\}$ ist $f^{-1}(B') = \emptyset$.
- $\mathrm{Im}(f) = \{x \in \mathbb{R} \mid x \geq 0\}$.

b. Die Abbildung $\mathrm{nf} : \mathbb{N} \longrightarrow \mathbb{N} : x \mapsto x + 1$ nennen wir die *Nachfolgerfunktion*. Es gelten

$$\mathrm{Im}(\mathrm{nf}) = \mathbb{N} \setminus \{0\}$$

und

$$\forall\, y \in \mathrm{Im}(f) \ : \ \mathrm{nf}^{-1}(\{y\}) = \{y - 1\}.$$

Bemerkung A3.5 (Abbildungen und ihre Graphen)

a. Für zwei Abbildungen $f : M \longrightarrow N$ und $g : P \longrightarrow N$ gilt:

$$f = g \quad \Longleftrightarrow \quad \mathrm{Graph}(f) = \mathrm{Graph}(g).$$

b. Ist $\Gamma \subseteq M \times N$ so, daß

$$\forall\, x \in M\ \exists_1\, y \in N \ : \ (x, y) \in \Gamma,$$

dann gibt es eine Abbildung $f : M \longrightarrow N$ mit $\Gamma = \mathrm{Graph}(f)$.

Fazit: Man hätte Abbildungen von M nach N auch als Teilmengen von $M \times N$ definieren können, die die Bedingung in b. erfüllen. So würde man vorgehen, wenn man die Mathematik ausgehend vom Begriff der Menge sukzessive aufbauen möchte.

Mit dieser Beschreibung sieht man übrigens sehr schön, daß es für jede Menge M *genau eine* Abbildung $f : \emptyset \longrightarrow M$ gibt, und daß es für eine nicht-leere Menge M keine Abbildung $f : M \longrightarrow \emptyset$ geben kann.

Definition A3.6 (Injektiv, surjektiv, bijektiv)
Es sei $f : M \longrightarrow N$ eine Abbildung.

a. f heißt genau dann *injektiv*, wenn

$$\forall\, x, x' \in M \;:\; f(x) = f(x') \implies x = x'.$$

b. f heißt genau dann *surjektiv*, wenn

$$\forall\, y \in N \;\exists\, x \in M \;:\; f(x) = y,$$

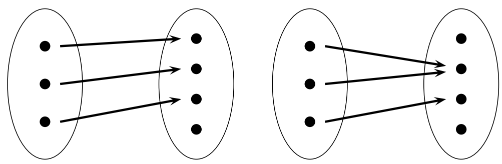
d.h. wenn $\mathrm{Im}(f) = N$.

c. f heißt genau dann *bijektiv*, wenn f injektiv und surjektiv ist.

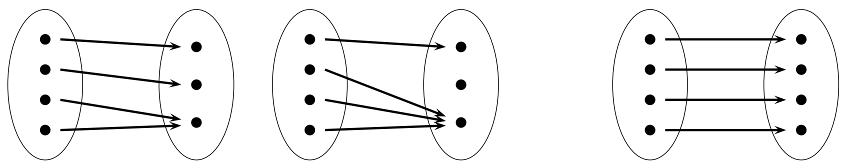
Bemerkung A3.7 (Injektiv, surjektiv, bijektiv)
Es sei $f : M \longrightarrow N$ eine Abbildung.

a. Ist $y \in N$ und $x \in M$ mit $f(x) = y$, so nennen wir x *ein Urbild* von y unter f.

b. Es gelten:
   - f ist injektiv $\iff$ jedes $y \in N$ hat *höchstens* ein Urbild.
   - f ist surjektiv $\iff$ jedes $y \in N$ hat *mindestens* ein Urbild.
   - f ist bijektiv $\iff$ jedes $y \in N$ hat *genau* ein Urbild.



injektiv nicht injektiv



surjektiv nicht surjektiv bijektiv

Beispiel A3.8

a. Die Nachfolgerfunktion $\mathrm{nf} : \mathbb{N} \longrightarrow \mathbb{N} : x \mapsto x + 1$ ist injektiv, aber nicht surjektiv. Denn, $x + 1 = \mathrm{nf}(x) = \mathrm{nf}(y) = y + 1$ für $x, y \in \mathbb{N}$ impliziert $x = y$, und $0 \notin \mathrm{Im}(f)$.

b. $g : \mathbb{Z} \longrightarrow \mathbb{N} : x \mapsto x^2$ ist nicht injektiv.
Denn, für $x = 1 \neq -1 = y$ gilt $g(x) = g(1) = 1 = g(-1) = g(y)$.

c. Die Abbildung id_M ist bijektiv für jede Menge M.
Denn, für $y \in M$ gilt $\mathrm{id}_M(y) = y$, so daß id_M surjektiv ist, und für $x, x' \in M$ mit $\mathrm{id}_M(x) = \mathrm{id}_M(x')$ gilt $x = x'$, so daß id_M injektiv ist.

d. Ist $f : M \longrightarrow N$ injektiv, so ist die Abbildung $M \longrightarrow \mathrm{Im}(f) : x \mapsto f(x)$ offenbar bijektiv.

Definition A3.9 (Komposition von Abbildungen)
Seien $f : M \longrightarrow N$ und $g : N \longrightarrow P$ zwei Abbildungen. Die Abbildung

$$g \circ f : M \longrightarrow P : x \mapsto g(f(x))$$

heißt die *Komposition* oder *Verkettung* von f und g.

Beispiel A3.10
Seien $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^2$ und $g : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x + 1$. Dann gilt

$$(g \circ f)(x) = g(f(x)) = g(x^2) = x^2 + 1$$

und

$$(f \circ g)(x) = f(g(x)) = f(x + 1) = (x + 1)^2 = x^2 + 2x + 1.$$

Man beachte, daß die Abbildungen $g \circ f$ und $f \circ g$ nicht gleich sind, da $(g \circ f)(1) = 2 \neq 4 = (f \circ g)(1)$.

Proposition A3.11 (Assoziativität der Komposition)
Seien $f : M \longrightarrow N$, $g : N \longrightarrow P$ *und* $h : P \longrightarrow Q$ *Abbildungen. Dann gilt*

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

Wir schreiben deshalb auch kurz $h \circ g \circ f$.

Beweis: Da die Definitions- und Zielbereiche der beiden Funktionen übereinstimmen, reicht es, die Abbildungsvorschrift zu überprüfen. Sei dazu $x \in M$. Dann gilt

$$\big((h \circ g) \circ f\big)(x) = (h \circ g)\big(f(x)\big) = h\Big(g\big(f(x)\big)\Big) = h\big((g \circ f)(x)\big) = \big(h \circ (g \circ f)\big)(x).$$

Dies zeigt die Behauptung. □

Satz A3.12 (Bijektivität = Existenz einer Umkehrabbildung)
Es sei $f : M \longrightarrow N$ *eine Abbildung.*

a. f *ist genau dann bijektiv, wenn eine Abbildung* $g : N \longrightarrow M$ *existiert, so daß* $g \circ f = \mathrm{id}_M$ *und* $f \circ g = \mathrm{id}_N$.

b. *Die Abbildung* g *in Teil a. ist dann eindeutig bestimmt und bijektiv. Wir nennen sie die* Inverse *oder* Umkehrabbildung *von* f *und bezeichnen sie mit* f^{-1}.

Beweis:

a. **"$\Longleftarrow$":** Wir wollen zunächst zeigen, daß f surjektiv ist. Sei dazu $y \in N$ gegeben. Setze $x := g(y) \in M$. Dann gilt

$$f(x) = f\big(g(y)\big) = (f \circ g)(y) = \mathrm{id}_N(y) = y.$$

Also ist f surjektiv.
Dann wollen wir zeigen, daß f injektiv ist. Seien dazu $x, x' \in M$ mit $f(x) = f(x')$ gegeben. Dann gilt

$$x = \mathrm{id}_M(x) = (g \circ f)(x) = g\big(f(x)\big) = g\big(f(x')\big) = (g \circ f)(x') = \mathrm{id}_M(x') = x'.$$

Also ist f injektiv.

<u>„$\Longrightarrow$":</u> Da f bijektiv ist, gibt es für jedes $y \in N$ genau ein Urbild $x_y \in M$ von y unter f, d.h. $f(x_y) = y$. Wir definieren nun eine Abbildung

$$g : N \longrightarrow M : y \mapsto x_y.$$

Dann gilt zunächst für $y \in N$

$$(f \circ g)(y) = f\big(g(y)\big) = f(x_y) = y = \mathrm{id}_N(y).$$

Also ist $f \circ g = \mathrm{id}_N$.
Zudem gilt für $x \in M$ und $y := f(x) \in N$

$$f(x_y) = y = f(x).$$

Da f injektiv ist, folgt daraus $x = x_y$, und wir erhalten

$$(g \circ f)(x) = g\big(f(x)\big) = g(y) = x_y = x = \mathrm{id}_M(x).$$

Damit ist auch $g \circ f = \mathrm{id}_M$ gezeigt.

b. Sei $h : N \longrightarrow M$ eine zweite Abbildung mit $h \circ f = \mathrm{id}_M$ und $f \circ h = \mathrm{id}_N$. Dann gilt für $y \in N$

$$f\big(g(y)\big) = (f \circ g)(y) = \mathrm{id}_N(y) = (f \circ h)(y) = f\big(h(y)\big).$$

Da f injektiv ist, folgt mithin $g(y) = h(y)$, und somit $g = h$. Die Eindeutigkeit von g ist also gezeigt. Außerdem ist g nach Teil a. auch bijektiv.

$\square$

Beispiel A3.13
Die Abbildung $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto 2x+1$ ist bijektiv mit $f^{-1} : \mathbb{R} \longrightarrow \mathbb{R} : y \mapsto \frac{1}{2} \cdot y - \frac{1}{2}$.
Denn für $y \in \mathbb{R}$ gilt

$$\big(f \circ f^{-1}\big)(y) = 2 \cdot \left(\frac{1}{2} \cdot y - \frac{1}{2}\right) + 1 = y = \mathrm{id}_{\mathbb{R}}(y)$$

und für $x \in \mathbb{R}$ gilt

$$\big(f^{-1} \circ f\big)(x) = \frac{1}{2} \cdot (2x + 1) - \frac{1}{2} = x = \mathrm{id}_{\mathbb{R}}(x).$$

Die Behauptung folgt also aus Satz A3.12.

Proposition A3.14 (Injektivität, Surjektivität, Bijektivität unter Komposition)
Seien $f : M \longrightarrow N$ *und* $g : N \longrightarrow P$ *zwei Abbildungen.*

a. *Sind* f *und* g *injektiv, so ist* $g \circ f$ *injektiv.*
b. *Sind* f *und* g *surjektiv, so ist* $g \circ f$ *surjektiv.*
c. *Sind* f *und* g *bijektiv, so ist* $g \circ f$ *bijektiv.*

Beweis: a. Seien $x, x' \in M$ mit $(g \circ f)(x) = (g \circ f)(x')$. Dann gilt

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

Da g injektiv ist, ist $f(x) = f(x')$, und da f injektiv ist, ist auch $x = x'$. Also ist $g \circ f$ injektiv.

b. Sei $z \in P$. Da g surjektiv ist, gibt es ein $y \in N$ mit $g(y) = z$, und da f surjektiv ist, gibt es ein $x \in M$ mit $f(x) = y$. Die Surjektivität von $g \circ f$ folgt dann aus

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

c. Wegen a. ist $g \circ f$ injektiv und wegen b. ist $g \circ f$ auch surjektiv, also bijektiv.

$\square$

Aufgaben

Aufgabe A3.15
Ist $f : M \longrightarrow N$ eine surjektive Abbildung und $y \in N$, so ist

$$g : M \setminus f^{-1}(\{y\}) \longrightarrow N \setminus \{y\} : x \mapsto f(x)$$

eine surjektive Abbildung.

Aufgabe A3.16
Es sei $f : M \longrightarrow N$ eine injektive Abbildung, $x' \in M$ und $y' = f(x') \in N$.

a. Dann ist $g : M \setminus \{x'\} \longrightarrow N \setminus \{y'\} : x \mapsto f(x)$ eine injektive Abbildung.
b. Ist f bijektiv, so ist g auch bijektiv.

Aufgabe A3.17
Für eine Abbildung $f : M \longrightarrow N$, $M \neq \emptyset$, beweise man die folgenden Aussagen:

a. f ist injektiv $\iff$ $\exists\, g : N \longrightarrow M$, so dass $g \circ f = \mathrm{id}_M$.
b. f ist surjektiv $\iff$ $\exists\, g : N \longrightarrow M$, so dass $f \circ g = \mathrm{id}_N$.

Aufgabe A3.18
Untersuche ob die folgenden Abbildungen injektiv, surjektiv oder bijektiv sind:

a. $f_1 : \mathbb{R} \longrightarrow \mathbb{R} :\ x \longmapsto 3x + 2$
b. $f_2 : \mathbb{Z} \longrightarrow \mathbb{Z} :\ x \longmapsto 3x + 2$
c. $f_3 : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R} \times \mathbb{R} :\ (x, y) \longmapsto (xy, x + 1)$
d. $f_4 : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R} \times \mathbb{R} :\ (x, y) \longmapsto (x - 2y, 2x + y)$

Aufgabe A3.19
Seien M, N zwei nicht-leere Mengen und $f : M \longrightarrow N$ eine Abbildung. Formuliere die folgende Aussage in Quantorenschreibweise und beweise sie:
f ist genau dann surjektiv, wenn für alle nicht-leeren Mengen X und für alle Abbildungen $g : N \longrightarrow X$ und $h : N \longrightarrow X$ aus $g \circ f = h \circ f$ schon $g = h$ folgt.

Aufgabe A3.20

Seien L, M, N Mengen und $f : L \longrightarrow M$, $g : M \longrightarrow N$ Abbildungen. Beweise oder widerlege - durch Gegenbeispiel - die folgenden Aussagen:

a. Ist $g \circ f$ injektiv, so ist g injektiv.

b. Ist $g \circ f$ injektiv, so ist f injektiv.

c. Ist $g \circ f$ surjektiv, so ist g surjektiv.

d. Ist $g \circ f$ surjektiv, so ist f surjektiv.

Aufgabe A3.21

Seien M, N Mengen, $A_1, A_2 \subseteq M$ und $B, B_1, B_2 \subseteq N$ Teilmengen und $f : M \longrightarrow N$ eine Abbildung. Beweise die folgenden Aussagen:

a. $f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$.

b. $f(f^{-1}(B)) \subseteq B$.

c. $f(A_1 \cup A_2) = f(A_1) \cup f(A_2)$.

d. $f(A_1 \cap A_2) \subseteq f(A_1) \cap f(A_2)$.

Gib außerdem konkrete Beispiele dafür an, dass in b. und d. keine Gleichheit gilt.

§ A4 Vollständige Induktion

Bemerkung A4.1 (Prinzip der vollständigen Induktion)
Die folgende Eigenschaft der natürlichen Zahlen ist uns wohl vertraut:

> Addiert man zur Zahl 0 sukzessive die Zahl 1, so erhält man nach und nach alle natürlichen Zahlen.

Man nennt sie das *Prinzip der vollständigen Induktion.*

Mit Hilfe der Nachfolgerfunktion $\mathrm{nf} : \mathbb{N} \longrightarrow \mathbb{N} : n \mapsto n + 1$ können wir die Eigenschaft auch wie folgt formulieren:

> Ist $M \subseteq \mathbb{N}$ mit $0 \in M$ und $\forall\, n \in M \ : \ n+1 = \mathrm{nf}(n) \in M$, so ist $M = \mathbb{N}$.

Daraus leitet sich das im folgenden Satz formulierte Beweisprinzip für Aussagen ab, die von einer natürlichen Zahl abhängen.

Satz A4.2 (Prinzip der vollständigen Induktion)
Sei $\mathcal{A}(n)$ eine Aussageform mit zulässigen Werten $n \in \mathbb{N}$.
Falls $\mathcal{A}(0)$ wahr ist und $\mathcal{A}(n) \Rightarrow \mathcal{A}(n+1)$ wahr ist, so ist $\mathcal{A}(n)$ wahr für alle $n \in \mathbb{N}$.

Beweis: Setze $M := \{n \in \mathbb{N} \mid \mathcal{A}(n) \text{ wahr}\}$. Nach Voraussetzung gilt dann $0 \in M$ und für $n \in M$ folgt $n + 1 \in M$. Aus dem Prinzip der Vollständigen Induktion in Bemerkung A4.1 folgt dann $M = \mathbb{N}$. Also ist $\mathcal{A}(n)$ wahr für alle $n \in \mathbb{N}$. □

Bemerkung A4.3
Man beachte, um den Schluß $\mathcal{A}(n) \Rightarrow \mathcal{A}(n + 1)$ als wahr zu erweisen, reicht es, den Fall zu betrachten, daß $\mathcal{A}(n)$ wahr ist, da andernfalls der Schluß ohnehin den Wahrheitswert wahr trägt.

Wir nennen:

- "$\mathcal{A}(0)$ wahr" den *Induktionsanfang*,
- "$\mathcal{A}(n)$ wird als wahr vorausgesetzt" die *Induktionsvoraussetzung* und
- "$\mathcal{A}(n) \Rightarrow \mathcal{A}(n + 1)$" den *Induktionsschluß*.

Beispiel A4.4
Die Zahl $n^3 - n$ ist für jedes $n \in \mathbb{N}$ durch 6 teilbar.

Beweis: Wir führen den Beweis durch vollständige Induktion und formulieren dazu zunächst unsere Aussageform $\mathcal{A}(n)$:

$$\mathcal{A}(n) : \text{Es gibt ein } k \in \mathbb{N} \text{ mit } n^3 - n = 6 \cdot k.$$

Induktionsanfang: $n = 0$: $0^3 - 0 = 0 = 6 \cdot 0$. Also ist $\mathcal{A}(0)$ wahr.
Induktionsvoraussetzung: Wir setzen voraus, daß $\mathcal{A}(n)$ wahr ist, d.h. es gibt ein $k \in \mathbb{N}$ mit $n^3 - n = 6 \cdot k$.

Induktionsschritt: $n \mapsto n+1$: Man beachte, daß eine der beiden Zahlen n oder $n+1$ gerade sein muß, und daß deshalb die Zahl $n \cdot (n+1)$ gerade ist. Es gibt also eine natürliche Zahl $l \in \mathbb{N}$ mit $n \cdot (n+1) = 2 \cdot l$. Damit erhalten wir dann

$$(n+1)^3 - (n+1) = (n^3 - n) + 3 \cdot n \cdot (n+1) = 6k + 6l = 6 \cdot (k+l).$$

Wir haben also gezeigt, daß $\mathcal{A}(n+1)$ wahr ist.

Also ist $\mathcal{A}(n)$ wahr für alle $n \in \mathbb{N}$, und das heißt, daß $n^3 - n$ stets durch 6 teilbar ist. □

Bemerkung A4.5 (Varianten der vollständigen Induktion)

a. *Alternativer Induktionsanfang*:
Statt $n = 0$ als Induktionsanfang zu wählen, kann eine beliebige ganze Zahl $n_0 \in \mathbb{Z}$ als Induktionsanfang dienen. Man erhält dann, daß $\mathcal{A}(n)$ wahr ist für alle $n \geq n_0$. Denn, man erhält alle ganzen Zahlen $n \geq n_0$, indem man zu n_0 sukzessive 1 addiert.

b. *Alternative Induktionsvoraussetzung*:
Im Induktionsschritt schließen wir von $\mathcal{A}(n)$ auf $\mathcal{A}(n+1)$, d.h. wir setzen nur $\mathcal{A}(n)$ als richtig voraus und schließen daraus die Korrektheit von $\mathcal{A}(n+1)$. Stattdessen können wir auch $\mathcal{A}(k)$ für $k = n_0, \ldots, n$ als richtig voraussetzen und auf $\mathcal{A}(n+1)$ schließen (wobei $\mathcal{A}(n_0)$ der Induktionsanfang sein soll). Das ist manchmal hilfreich.

Aufgaben

Aufgabe A4.6
Zeige, daß $3^{n+1} - 3$ für jede natürliche Zahl $n \in \mathbb{N}$ durch 6 teilbar ist.

Aufgabe A4.7
Es sei $a \in \mathbb{N}$ eine natürliche Zahl. Zeige, daß dann $a^{2n+1} - a$ für jede natürliche Zahl $n \in \mathbb{N}$ durch 6 teilbar ist.

§ A5 Mächtigkeit von Mengen

Definition A5.1 (Die Mächtigkeit von Mengen)

a. Wir nennen eine Menge M *endlich*, wenn sie nur endlich viele Elemente enthält. In diesem Fall bezeichnen wir mit $\#M = |M|$ die Anzahl an Elementen in M und nennen die Zahl die *Mächtigkeit* von M. Enthält M unendlich viele Elemente, so nennen wir M *unendlich* und setzen $\#M := |M| := \infty$.

b. Zwei Mengen M und N heißen *gleichmächtig*, wenn es eine bijektive Abbildung $f : M \longrightarrow N$ gibt.

c. Eine Menge heißt *abzählbar unendlich*, wenn sie gleichmächtig zu $\mathbb{N}$ ist.

d. Eine Menge heißt *überabzählbar*, wenn sie weder endlich noch abzählbar unendlich ist.

e. Für zwei ganze Zahlen $m, n \in \mathbb{Z}$ bezeichnen wir mit

$$\{m, \ldots, n\} := \{k \in \mathbb{Z} \mid m \leq k \leq n\}$$

die Menge der ganzen Zahlen zwischen m und n. Man beachte, daß $\{m, \ldots, n\} = \emptyset$, wenn $m > n$.

Bemerkung A5.2 (Einfache Eigenschaften der Mächtigkeit endlicher Mengen)

a. Ist eine Menge endlich und enthält genau n Elemente, so können wir die Elemente in M abzählen, etwa $x_1, x_2, x_3, \ldots, x_n$ und wir erhalten so eine bijektive Abbildung

$$f : \{1, \ldots, n\} \longrightarrow M : i \mapsto x_i.$$

Umgekehrt erlaubt eine solche Abbildung, die Elemente von M abzuzählen und wir erhalten $|M| = n$. Damit sehen wir, daß eine Menge genau dann endlich von Mächtigkeit n ist, wenn es eine Bijektion von $\{1, \ldots, n\}$ nach M gibt.

b. Ist M endlich und $A \subseteq M$, so ist auch A endlich und $|A| \leq |M|$.

c. Ist $M = A \cup B$ eine endliche Menge, so gilt $|M| = |A| + |B|$.

Wir wollen den in Bemerkung A5.2 angedeuteten Zusammenhang zwischen der Mächtigkeit einer endlichen Menge und der Existenz von Abbildungen mit bestimmten Eigenschaften im folgenden Satz vertiefen.

Satz A5.3

Es seien M und N zwei nicht-leere endliche Mengen.

a. *Genau dann gilt $|M| \leq |N|$, wenn es eine injektive Abbildung $f : M \to N$ gibt.*

b. *Genau dann gilt $|M| \geq |N|$, wenn es eine surjektive Abbildung $f : M \to N$ gibt.*

c. *Genau dann gilt $|M| = |N|$, wenn es eine bijektive Abbildung $f : M \to N$ gibt.*

Beweis: Es seien $M = \{x_1, \ldots, x_m\}$ und $N = \{y_1, \ldots, y_n\}$ mit paarweise verschiedenen Elementen $x_i \neq x_j$ für $i \neq j$ und $y_i \neq y_j$ für $i \neq j$. Es gilt $|M| = m > 0$ und $|N| = n > 0$.

a. Ist $m \leq n$, so definiere $f : M \to N$ durch $f(x_i) = y_i$ für $i = 1, \ldots, m$. Dann gilt für $i, j \in \{1, \ldots, m\}$ mit $i \neq j$

$$f(x_i) = y_i \neq y_j = f(x_j).$$

Mithin ist f injektiv.
Ist umgekehrt $f : M \to N$ eine injektive Abbildung, so gilt $f(M) = \{f(x_1), \ldots, f(x_m)\} \subseteq N$ eine Teilmenge von paarweise verschiedenen Elementen. Mithin enthält N mindestens m Elemente, und folglich gilt $m \leq n$.

b. Ist $m \geq n$, so definiere $f : M \to N$ durch $f(x_i) = y_i$ für $i = 1, \ldots, n$ und $f(x_i) = y_1$ für $i = n + 1, \ldots, m$. Dann gilt offenbar $f(M) = \{y_1, \ldots, y_n\} = N$ und f ist surjektiv.
Ist umgekehrt $f : M \to N$ eine surjektive Abbildung, so gilt $\{y_1, \ldots, y_n\} = N = f(M) = \{f(x_1), \ldots, f(x_m)\}$. Mithin enthält die Menge $\{f(x_1), \ldots, f(x_m)\}$ auch n verschiedene Elemente, und folglich ist $m \geq n$.

c. Die Aussage folgt unmittelbar aus den ersten beiden Teilen.

□

Aus diesem Satz leitet sich unmittelbar ab, daß für Selbstabbildungen endlicher Mengen die Begriffe injektiv, surjektiv und bijektiv zusammen fallen.

Korollar A5.4 (Injektiv = surjektiv = bijektiv für gleichmächtige endliche Mengen)
Es seien M und N endliche Mengen mit $|M| = |N|$. Dann sind die folgenden Aussagen für eine Abbildung $f : M \longrightarrow N$ äquivalent:

a. f *ist injektiv.*

b. f *ist surjektiv.*

c. f *ist bijektiv.*

Beweis:

a. $\Longrightarrow$ b.: Angenommen, f wäre nicht surjektiv, dann gibt es ein

$$y \in N \setminus \mathrm{Im}(f)$$

und mithin ist

$$\mathrm{Im}(f) \subseteq N \setminus \{y\}.$$

Da f injektiv ist, ist $g : M \longrightarrow \mathrm{Im}(f) : x \mapsto f(x)$ nach Beispiel A3.8 bijektiv, so daß mit Satz A5.3

$$|M| \overset{\text{A5.3}}{=} |\mathrm{Im}(f)| \leq |N| - 1 < |N| = |M|$$

folgt, was ein offensichtlicher Widerspruch ist. Mithin muß f surjektiv sein.

b. $\Longrightarrow$ c.: Wir müssen zeigen, daß f injektiv ist. Dazu nehmen wir an, f sei nicht injektiv. Dann gibt es $x, x' \in M$ mit $x \neq x'$ und $y := f(x) = f(x')$. Die Abbildung

$$h : M \setminus f^{-1}(\{y\}) \longrightarrow N \setminus \{y\} : z \mapsto f(z)$$

ist nach Aufgabe A3.15 surjektiv. Mithin gilt nach Satz A5.3

$$|M| - 1 \overset{\text{Vor.}}{=} |N| - 1 = |N \setminus \{y\}| \overset{\text{A5.3}}{\leq} |M \setminus f^{-1}(\{y\})| \leq |M \setminus \{x, x'\}| = |M| - 2,$$

was offenbar ein Widerspruch ist. Mithin muß f injektiv sein.

c. $\Longrightarrow$ a.: Jede bijektive Abbildung ist auch injektiv, also ist f injektiv.

Damit haben wir die Aussage durch einen Ringschluß gezeigt. □

Nachdem wir uns bislang im wesentlichen mit endlichen Mengen beschäftigt haben, wollen wir uns nun unendlichen Mengen zuwenden und dabei zeigen, daß es unterschiedliche Qualitäten der Unendlichkeit gibt.

Proposition A5.5 (Cantorsches Diagonalverfahren)
Die Menge $\mathbb{Q}$ der rationalen Zahlen ist abzählbar unendlich.

Beweis: Wir zeigen, wie man mit Hilfe des Cantorschen Diagonalverfahrens eine bijektive Abbildung von $\mathbb{N}$ nach $\mathbb{Q}$ konstruiert.

Dazu listen wir die rationalen Zahlen zunächst wie folgt auf

$$\begin{tikzcd}
0 \arrow[r] & 1 \arrow[d] & \frac{1}{2} \arrow[r] & \frac{1}{3} \arrow[dl] & \frac{1}{4} \arrow[r] & \frac{1}{5} \arrow[dl] & \cdots \\
& -1 \arrow[ur] & -\frac{1}{2} \arrow[dl] & -\frac{1}{3} \arrow[ur] & -\frac{1}{4} \arrow[dl] & -\frac{1}{5} & \cdots \\
& 2 \arrow[d] & \frac{2}{2} \arrow[ur] & \frac{2}{3} \arrow[dl] & \frac{2}{4} & \frac{2}{5} & \cdots \\
& -2 \arrow[ur] & -\frac{2}{2} & -\frac{2}{3} & -\frac{2}{4} & -\frac{2}{5} & \cdots \\
& \vdots & \vdots & \vdots & \vdots & \vdots &
\end{tikzcd}$$

und laufen sie dann wie angedeutet entlang der Pfeile ab. Dabei sammeln wir jede rationale Zahl, die mehrfach vorkommt, nur bei ihrem ersten Auftreten auf. Auf dem Weg erhalten wir eine bijektive Abbildung von $\mathbb{N}$ nach $\mathbb{Q}$. □

Proposition A5.6 ($\mathbb{R}$ ist überabzählbar.)
Die Menge $\mathbb{R}$ der reellen Zahlen ist überabzählbar.

Beweis: Auch dies zeigen wir mit Hilfe einer Variante des Cantorschen Diagonalverfahrens.

$\mathbb{R}$ ist sicherlich nicht endlich. Wäre $\mathbb{R}$ abzählbar unendlich, so gäbe es eine bijektive Abbildung von $\varphi : \mathbb{N} \to \mathbb{R}$, und wir schreiben dann $\varphi(i)$, $i \in \mathbb{N}$, in Dezimaldarstellung

$$\begin{array}{lcllclllll} \varphi(0) & = & a_{0,-p_0} & a_{0,-p_0+1} & \ldots & \underline{a_{0,0}}, & a_{01} & a_{02} & a_{03} & \cdots \\ \varphi(1) & = & a_{1,-p_1} & a_{1,-p_1+1} & \ldots & a_{1,0}, & \underline{a_{11}} & a_{12} & a_{13} & \cdots \\ \varphi(2) & = & a_{2,-p_2} & a_{2,-p_2+1} & \ldots & a_{2,0}, & a_{21} & \underline{a_{22}} & a_{23} & \cdots \\ \vdots & & & & & & & & \ddots & \end{array}$$

Dann setzen wir $a := a_{00}, a_{11}a_{22}a_{33}\cdots \in \mathbb{R}$, d. h. a ist diejenige Zahl, die in obiger Aufzählung durch die unterstrichenen Diagonalelemente gegeben ist. Nun ändern wir jede der Ziffern von a ab (etwa $b_{ii} = 1$, falls $a_{ii} = 0$ und $b_{ii} = 0$ sonst) und erhalten eine Zahl

$$b = b_{00}, b_{11}b_{22}b_{33}\cdots \in \mathbb{R},$$

mit $a_{ii} \neq b_{ii}$ für alle $i \in \mathbb{N}$. Da φ bijektiv ist, gibt es ein $i \in \mathbb{N}$ mit $\varphi(i) = b$, also $a_{ii} = b_{ii}$, im Widerspruch zur Konstruktion von b. (Wir müssen noch berücksichtigen, daß $0,9999\cdots = 1$, was aber die einzige Zweideutigkeit der Dezimaldarstellung ist, und dieser weichen wir durch unsere Wahl der b_{ii} aus.) Also ist $\mathbb{R}$ überabzählbar. □

Bemerkung A5.7 (Kontinuumshypothese)
Da $\mathbb{Q}$ und $\mathbb{R}$ nicht gleichmächtig sind und $\mathbb{Q}$ eine Teilmenge von $\mathbb{R}$ ist, stellt sich ganz natürlich die Frage, ob es eine Menge M mit $\mathbb{Q} \subsetneqq M \subsetneqq \mathbb{R}$ gibt, die weder zu $\mathbb{Q}$ noch zu $\mathbb{R}$ gleichmächtig ist. Es hat lange gedauert, bis man feststellen mußte, daß die Frage auf der Grundlage des allgemein anerkannten Axiomensystems der Mengenlehre von Zermelo-Fränkel nicht entscheidbar ist. Man hat nun also die Wahl, als neues Axiom hinzuzufügen, daß es eine solche Menge gibt, oder auch, daß es keine solche Menge gibt. Die lange bestehende Vermutung, daß man schon mit den übrigen Axiomen beweisen könnte, daß es keine solche Menge gibt, ist als *Kontinuumshypothese* bekannt.

Definition A5.8 (Potenzmenge)
Es sei M ein eine Menge. Wir nennen die Menge

$$\mathcal{P}(M) := \{A \mid A \subseteq M\}$$

aller Teilmengen von M die *Potenzmenge* von M.

Beispiel A5.9
$\mathcal{P}(\emptyset) = \{\emptyset\}$, $\mathcal{P}(\{1\}) = \{\emptyset, \{1\}\}$, $\mathcal{P}(\{1,2\}) = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}$.

Proposition A5.10 (Potenzmengen endlicher Mengen)
Sei M eine endliche Menge mit $n = |M|$, so ist $|\mathcal{P}(M)| = 2^n$.

Beweis: Wir führen den Beweis durch Induktion nach n.

Induktionsanfang: $n = 0$**:** Dann ist $M = \emptyset$ und $\mathcal{P}(M) = \{\emptyset\}$ hat genau $1 = 2^0$ Elemente.

Induktionsschritt: $n \mapsto n+1$**:** Sei also $|M| = n+1$. Wir wählen ein $y \in M$ und setzen $N = M \setminus \{y\}$, so daß $|N| = |M| - 1 = n$. Die Potenzmenge $\mathcal{P}(M)$ läßt sich nun wie folgt disjunkt aufspalten:

$$\mathcal{P}(M) = \{A \subseteq M \mid y \notin A\} \dot{\cup} \{A \subseteq M \mid y \in A\}.$$

Dabei ist

$$\{A \subseteq M \mid y \notin A\} = \{A \subseteq M \mid A \subseteq N\} = \mathcal{P}(N)$$

und

$$\{A \subseteq M \mid y \in A\} = \{B \cup \{y\} \subseteq M \mid B \subseteq N\} = \{B \cup \{y\} \subseteq M \mid B \in \mathcal{P}(N)\}.$$

Beide Mengen sind offenbar gleichmächtig zu $\mathcal{P}(N)$, und nach Induktionsvoraussetzung gilt $|\mathcal{P}(N)| = 2^n$. Insgesamt erhalten wir also

$$|\mathcal{P}(M)| = 2^n + 2^n = 2^{n+1}.$$

Damit folgt die Aussage mittels Induktion. □

Aufgaben

Aufgabe A5.11
Die Menge $\mathbb{Z}$ der ganzen Zahlen ist abzählbar unendlich.

§ A6 Äquivalenzrelationen

Äquivalenzrelationen stellen ein sehr wichtiges *Ordnungs-* und *Konstruktionsprinzip* innerhalb der Mathematik dar, auf das wir im Verlauf der Vorlesung an einigen zentralen Stellen zurückkommen werden, etwa im Zusammenhang mit Faktorräumen (siehe Bemerkung 2.29), der Äquivalenz von Matrizen (siehe Bemerkung 5.22) oder der Konjugation von Matrizen und der Jordanschen Normalform (siehe Bemerkung 9.3).

Definition A6.1 (Relation)
Seien M und N zwei Mengen, so nennen wir jede Teilmenge $R \subseteq M \times N$ eine *Relation* zwischen M und N.

Bemerkung A6.2
Ist R eine Relation zwischen M und N, $x \in M$ und $y \in N$, so wollen wir sagen x *steht in Relation zu* y *bezüglich* R, wenn $(x, y) \in R$. Die Menge R legt also fest, wann zwei Elemente in Relation zueinander stehen. Wir schreiben auch xRy statt $(x, y) \in R$.

Beispiel A6.3 (Abbildungen als Relationen)

a. Der Graph einer Abbildung $f : M \longrightarrow N$ ist ein Beispiel einer Relation, bei der jedes $x \in M$ zu genau einem $y \in N$ in Relation steht.

b. Ist M die Menge der Hörer der Vorlesung und N die Menge der in Kaiserslautern studierbaren Fächer, so ist

$$R = \{(x, y) \in M \times N \mid x \text{ studiert } y\}$$

eine Relation zwischen M und N, die ganz sicher nicht Graph einer Funktion ist.

Bemerkung A6.4 (Motivation des Begriffs Äquivalenzrelation)
Der folgende Begriff der *Äquivalenzrelation* bereitet den Studenten oft extreme Schwierigkeiten. Dabei liegt auch ihm ein ganz einfaches Prinzip zugrunde, das wir zunächst an einem Beispiel erläutern wollen.

Die Gesamtheit aller Schüler einer Schule werden von der Schulleitung zwecks sinnvoller Organisation des Unterrichts in Schulklassen eingeteilt. Dabei achtet die Schulleitung darauf, daß jeder Schüler zu einer Schulklasse gehört und auch nur zu dieser einen. Etwas mathematischer ausgedrückt, die Schulleitung teilt die *Menge* S der Schüler in *paarweise disjunkte Teilmengen* K_i, $i = 1, \ldots, k$, ein, so daß wir anschließend eine *disjunkte Zerlegung*

$$S = \dot{\bigcup_{i=1}^{k}} K_i$$

der Menge S in die Schulklassen $K_1, \ldots, K_k$ haben. Dabei kann man für die Zugehörigkeit der Schüler Alfred, Ben und Christoph zu einer Schulklasse folgendes feststellen:

1) Alfred gehört zu einer Schulklasse.
2) Wenn Alfred in der gleichen Schulklasse ist wie Ben, dann ist Ben auch in der gleichen Schulklasse wie Alfred.
3) Wenn Alfred in der gleichen Schulklasse ist wie Ben und wenn zugleich Ben in der gleichen Schulklasse ist wie Christoph, dann ist auch Alfred in der gleichen Schulklasse wie Christoph.

Diese Aussagen sind so offensichtlich, daß man kaum glauben mag, daß es einen tieferen Sinn hat, sie zu erwähnen. Aber nehmen wir für einen Augenblick an, die Schulleitung hat ihre Einteilung der Schüler vorgenommen und für jede Schulklasse eine Liste mit den Namen der Schüler erstellt, die zu dieser Schulklasse gehören sollen. Nehmen wir ferner an, die Schulleitung hat noch nicht überprüft, ob jeder Schüler in genau einer Schulklasse eingeteilt ist. Dann behaupte ich, wenn man in den drei Aussagen 1)-3) die Schüler Alfred, Ben und Christoph durch beliebige Schüler ersetzt und die Aussagen richtig sind für jede Kombination der Schülernamen, dann ist sichergestellt, daß auch jeder Schüler in genau einer Schulklasse eingeteilt ist.

Als Mathematiker suchen wir nach möglichst einfachen Regeln, denen die Einteilung der Schulklassen genügen muß, um sicherzustellen, daß sie wirklich eine disjunkte Zerlegung von S ist, d.h. daß wirklich jeder Schüler in genau einer Schulklasse ist, und die Regeln 1)-3) sind genau die Regeln, die wir dazu brauchen. Wenn wir nun die Zugehörigkeit zweier Schüler x und y zur gleichen Klasse verstehen als "x *steht in Relation zu* y", dann definieren uns die drei Regeln 1)-3) zudem eine Teilmenge von $S \times S$, nämlich die Relation

$$R = \{(x, y) \in S \times S \mid x \text{ ist in der gleichen Schulklasse wie } y\}.$$

Die Regeln 1)-3) lassen sich für Schüler $x, y, z \in S$ dann wie folgt formulieren:

- $(x, x) \in R$.
- Wenn $(x, y) \in R$, dann ist auch $(y, x) \in R$.
- Wenn $(x, y) \in R$ und $(y, z) \in R$, dann ist auch $(x, z) \in R$.

Eine solche Relation nennt man eine *Äquivalenzrelation*, man nennt Schüler der gleichen Schulklasse *äquivalent* und die Schulklassen nennt man dann auch *Äquivalenzklassen.*

Wir führen den Begriff der *Äquivalenzrelation* nun für beliebige Mengen ein.

Definition A6.5 (Äquivalenzrelation)
Es sei M eine Menge. Eine *Äquivalenzrelation* auf M ist eine Teilmenge $R \subseteq M \times M$, so daß für alle $x, y, z \in M$ gilt:

R1: $(x, x) \in R$, ("Reflexivität")
R2: $(x, y) \in R \;\Rightarrow\; (y, x) \in R$, ("Symmetrie")
R3: $(x, y), (y, z) \in R \;\Rightarrow\; (x, z) \in R$. ("Transitivität")

Bei Äquivalenzrelationen hat sich eine alternative Schreibweise zu $(x, y) \in R$ durchgesetzt, die auch wir im folgenden verwenden wollen.

Notation A6.6 (Schreibweise $\sim$ für Äquivalenzrelationen)
Sei M eine Menge und R eine Äquivalenzrelation auf M. Wir definieren für $x, y \in M$

$$x \sim y \quad :\Longleftrightarrow \quad (x, y) \in R,$$

und wir sprechen dann meist von der Äquivalenzrelation "$\sim$" statt R, sofern keine Mißverständnisse zu befürchten sind.

Mit dieser Schreibweise lassen sich die drei Axiome in Definition A6.5 wie folgt formulieren. Für $x, y, z \in M$ soll gelten:

R1: $x \sim x$, ("Reflexivität")
R2: $x \sim y \;\Rightarrow\; y \sim x$, ("Symmetrie")
R3: $x \sim y, y \sim z \;\Rightarrow\; x \sim z$. ("Transitivität")

Definition A6.7 (Äquivalenzklassen)
Es sei M eine Menge und $\sim$ eine Äquivalenzrelation auf M. Für $x \in M$ heißt die Menge

$$\overline{x} := \{y \in M \mid y \sim x\}$$

die *Äquivalenzklasse* von x. Jedes $y \in \overline{x}$ heißt ein *Repräsentant* der Klasse $\overline{x}$. Mit

$$M/\sim := \left\{\overline{x} \mid x \in M\right\}$$

bezeichnen wir die Menge der *Äquivalenzklassen modulo der Äquivalenzrelation* $\sim$.

Beispiel A6.8 (Der Abstand vom Ursprung als Äquivalenzrelation)
Wir betrachten die Menge $M = \mathbb{R}^2$ der Punkte in der reellen Zahlenebene und wir bezeichnen mit $|P|$ den Abstand von P zum Ursprung $(0, 0)$. Für zwei Punkte $P, Q \in M$ definieren wir

$$P \sim Q \quad \Longleftrightarrow \quad |P| = |Q|,$$

d.h. wir nennen die Punkte *äquivalent*, falls ihr Abstand zum Ursprung gleich ist. Dann ist $\sim$ eine Äquivalenzrelation.

R1: Sei $P \in M$, dann ist $|P| = |P|$, also $P \sim P$.
R2: Falls $P, Q \in M$ mit $P \sim Q$, dann ist $|P| = |Q|$ und somit auch $|Q| = |P|$. Damit gilt aber $Q \sim P$.
R3: Falls $P, Q, R \in M$ mit $P \sim Q$ und $Q \sim R$, dann gilt $|P| = |Q|$ und $|Q| = |R|$. Aber damit gilt auch $|P| = |R|$ und somit $P \sim R$.

Die Äquivalenzklasse

$$\overline{P} = \{Q \in M \mid |Q| = |P|\}$$

von $P \in M$ ist der Kreis um den Ursprung vom Radius $|P|$.

Wir haben anfangs behauptet, daß die drei Axiome einer Äquivalenzrelation sicherstellen, daß die zugehörigen Äquivalenzklassen eine disjunkte Zerlegung von M induzieren, und umgekehrt, daß jede disjunkte Zerlegung eine Äquivalenzrelation mit sich bringt. Dies wollen wir im Folgenden beweisen. Dazu sollten wir zunächst den Begriff disjunkt klären.

Proposition A6.9 (Die Äquivalenzrelation zu einer disjunkten Zerlegung)
Ist $(M_i)_{i\in I}$ eine disjunkte Zerlegung von M und definieren wir eine Relation auf M durch

$$x \sim y \quad \Longleftrightarrow \quad \exists\, i \in I\,:\, x, y \in M_i,$$

dann ist $\sim$ eine Äquivalenzrelation auf M.

Beweis: Ist $x \in M = \dot{\bigcup}_{i\in I} M_i$, so gibt es ein $i \in I$ mit $x \in M_i$ und somit gilt $x \sim x$. $\sim$ ist also reflexiv.

Sind $x, y \in M$ mit $x \sim y$, so gibt es ein $i \in I$ mit $x, y \in M_i$. Dann gilt aber auch $y \sim x$. Die Relation ist also symmetrisch.

Sind $x, y, z \in M$ mit $x \sim y$ und $y \sim z$, so gibt es $i, j \in I$ mit $x, y \in M_i$ und $y, z \in M_j$. Da die Zerlegung disjunkt ist und $y \in M_i \cap M_j$, folgt $M_i = M_j$. Also gilt $x, z \in M_i$ und somit $x \sim z$. $\sim$ ist also auch transitiv. □

Proposition A6.10 (Die disjunkte Zerlegung zu einer Äquivalenzrelation)
Es sei M eine Menge. Ist $\sim$ eine Äquivalenzrelation auf M, dann bilden die Äquivalenzklassen eine disjunkte Zerlegung von M, d. h. jedes $x \in M$ liegt in genau einer Äquivalenzklasse.

Insbesondere gilt für Äquivalenzklassen $\overline{x}$ und $\overline{y}$ entweder $\overline{x} = \overline{y}$ oder $\overline{x} \cap \overline{y} = \emptyset$.

Beweis: Sei $x \in M$ beliebig. Aus $x \sim x$ folgt $x \in \overline{x} \subseteq \bigcup_{\overline{y}\in M/\sim} \overline{y}$. Mithin gilt

$$M = \bigcup_{\overline{y}\in M/\sim} \overline{y}.$$

Es bleibt also zu zeigen, daß die Äquivalenzklassen paarweise disjunkt sind.

Seien $\overline{x}, \overline{y} \in M/\sim$ mit $\overline{x} \cap \overline{y} \neq \emptyset$. Dann gibt es ein $z \in \overline{x} \cap \overline{y}$, und es gilt $z \sim x$ und $z \sim y$. Wegen der Symmetrie gilt aber auch $x \sim z$ und mittels der Transitivität dann $x \sim y$. Sei nun $u \in \overline{x}$ beliebig, dann gilt $u \sim x$ und wieder wegen der Transitivität $u \sim y$. Also $u \in \overline{y}$ und damit $\overline{x} \subseteq \overline{y}$. Vertauschung der Rollen von x und y in der Argumentation liefert schließlich $\overline{x} = \overline{y}$. □

Korollar A6.11 (Äquivalenzrelationen auf endlichen Mengen)
Sei M eine endliche Menge, $\sim$ eine Äquivalenzrelation auf M und $M_1, \ldots, M_s$ seien die paarweise verschiedenen Äquivalenzklassen von $\sim$. Dann gilt:

$$|M| = \sum_{i=1}^{s} |M_i|.$$

Beweis: Mit M sind auch alle M_i endlich und die Behauptung folgt aus Proposition A6.10 und Bemerkung A5.2. □

Ein Beispiel aus dem Alltag für eine Äquivalenzrelation haben wir oben bereits gesehen. Ein weiteres wichtiges und wohlbekanntes Beispiel sind die rationalen Zahlen! Ein Bruch ist nichts weiter als die Äquivalenzklasse eines Tupels von ganzen Zahlen, und das Kürzen des Bruches, z.B. $\frac{1}{2} = \frac{2}{4}$, ist nur die Wahl eines möglichst einfachen Repräsentanten.

Beispiel A6.12 (Die rationalen Zahlen)
Man kann die rationalen Zahlen wie folgt als Äquivalenzklassen von Paaren ganzer Zahlen definieren. Für $(p, q), (p', q') \in M := \mathbb{Z} \times \big(\mathbb{Z} \setminus \{0\}\big)$ definiere

$$(p, q) \sim (p', q') \quad :\Longleftrightarrow \quad pq' = p'q.$$

Wir wollen nun zeigen, daß hierdurch wirklich eine Äquivalenzrelation auf M definiert wird. Seien dazu $x = (p, q), x' = (p', q'), x'' = (p'', q'') \in M$ gegeben:[1]

R1: Für die Reflexivität müssen wir $x \sim x$ zeigen. Nun gilt aber $pq = pq$, woraus $x = (p, q) \sim (p, q) = x$ folgt.

R2: Für die Symmetrie nehmen wir an, daß $x \sim x'$ gilt und müssen $x' \sim x$ folgern. Wegen $x \sim x'$ gilt aber nach Definition $pq' = p'q$, und folglich auch $p'q = pq'$. Letzteres bedeutet aber, daß $x' = (p', q') \sim (p, q) = x$.

R3: Für die Transitivität nehmen wir schließlich an, daß $x \sim x'$ und $x' \sim x''$ gilt, und müssen daraus schließen, daß $x \sim x''$. Wegen $x \sim x'$ gilt nun aber $pq' = p'q$, und wegen $x' \sim x''$ gilt $p'q'' = p''q'$. Multiplizieren wir die erste der Gleichungen mit q'' und die zweite mit q, so erhalten wir

$$pq'q'' = p'qq'' = p'q''q = p''q'q.$$

Da nach Voraussetzung $q' \neq 0$, können wir beide Seiten der Gleichung durch q' teilen und erhalten:

$$pq'' = p''q.$$

Das wiederum bedeutet, daß $x = (p, q) \sim (p'', q'') = x''$ gilt.

Die drei Axiome einer Äquivalenzrelation sind also erfüllt.

Wir setzen nun $\mathbb{Q} := M/\sim$ und für $(p, q) \in M$ setzen wir $\frac{p}{q} := \overline{(p, q)}$, d. h. die rationale Zahl $\frac{p}{q}$ ist die Äquivalenzklasse des Paares (p, q) unter der obigen Äquivalenzrelation. Dann bedeutet die Definition von $\sim$ soviel wie, daß $\frac{p}{q}$ und $\frac{p'}{q'}$ gleich

[1] Man sollte sich nicht dadurch verwirren lassen, daß die Elemente von M nun selbst schon Zahlenpaare sind! Wollte man die Relation als Teilmenge von $M \times M$ schreiben, so müßte man

$$R = \big\{\big((p, q), (p', q')\big) \in M \times M \;\big|\; pq' = p'q\big\}$$

betrachten. Das erläutert vielleicht auch, weshalb wir die *alternative* Schreibeweise bevorzugen – solche Paare von Paaren werden doch leicht unübersichtlich.

sind, wenn die kreuzweisen Produkte von Zähler und Nenner, pq' und $p'q$, übereinstimmen, oder in der vielleicht etwas bekannteren Formulierung, wenn die Brüche nach *Erweitern* mit q' bzw. mit q übereinstimmen: $\frac{p}{q} = \frac{pq'}{qq'} \stackrel{!}{=} \frac{p'q}{q'q} = \frac{p'}{q'}$.

Auch die Rechenregeln für rationale Zahlen lassen sich mit Hilfe der Äquivalenzklassen definieren. Für $(p, q), (r, s) \in M$ definiere:

$$\overline{(p,q)} + \overline{(r,s)} := \overline{(ps+qr, qs)},$$
$$\overline{(p,q)} \cdot \overline{(r,s)} := \overline{(pr, qs)}.$$

In Anlehnung an unser erstes Beispiel, der Einteilung der Schüler in Schulklassen, kann man das obige Rechenprinzip als "Rechnen mit Klassen" bezeichnen. Will man zwei Klassen addieren (bzw. multiplizieren), so nimmt man aus jeder der Klasse ein Element, addiert (bzw. multipliziert) diese Elemente und schaut, in welche Klasse das Resultat gehört. Diese Klasse ist dann die Summe (bzw. das Produkt) der beiden Klassen.

Was man sich bei diesem Vorgehen allerdings klar machen muß, ist, daß das Ergebnis nicht von der Wahl der Repräsentanten (d.h. der Elemente aus den Klassen) abhängt. Man spricht davon, daß die Operation *wohldefiniert* ist. Wir führen das für die Addition der rationalen Zahlen vor.

Sind $(p', q') \in \overline{(p,q)}$ und $(r', s') \in \overline{(r,s)}$ andere Repräsentanten, dann gilt $p'q = q'p$ und $r's = s'r$. Es ist zu zeigen, daß $(p's' + q'r', q's') \in \overline{(ps+qr, qs)}$ gilt. Ausmultiplizieren liefert

$$(p's' + q'r')(qs) = p'qs's + q'qr's = q'ps's + q'qs'r = (ps + qr)(q's'),$$

was zu zeigen war. □

Aufgaben

Aufgabe A6.13
Wir definieren für zwei Punkte $(x, y), (x', y') \in \mathbb{R}^2$

$$(x, y) \sim (x', y') \quad :\Longleftrightarrow \quad |x| + |y| = |x'| + |y'|.$$

Zeige, $\sim$ ist eine Äquivalenzrelation auf $\mathbb{R}^2$. Zeichne die Äquivalenzklassen zu $(1, 1)$ und zu $(-2, 3)$ in die Zahlenebene $\mathbb{R}^2$ ein.

Aufgabe A6.14 (Die ganzen Zahlen)
Es sei $M = \mathbb{N} \times \mathbb{N}$ und $m = (a, b) \in M$ und $m' = (a', b') \in M$ seien zwei Elemente in M. Wir definieren

$$m \sim m' \quad \Longleftrightarrow \quad a + b' = a' + b.$$

Zeige, daß $\sim$ eine Äquivalenzrelation ist und daß die folgende Abbildung bijektiv ist:

$$\Phi : \mathbb{Z} \longrightarrow M/\sim : z \mapsto \begin{cases} \overline{(z, 0)}, & \text{falls } z \geq 0, \\ \overline{(0, -z)}, & \text{falls } z < 0. \end{cases}$$

Aufgabe A6.15 (Die projektive Gerade)
Wir definieren für $v = (v_1, v_2), w = (w_1, w_2) \in \mathbb{R}^2 \setminus \{(0,0)\}$

$$v \sim w \quad \Longleftrightarrow \quad \exists\, \lambda \in \mathbb{R} \setminus \{0\} \,:\, v = \lambda \cdot w$$

wobei $\lambda \cdot w := (\lambda \cdot w_1, \lambda \cdot w_2)$.

a. Zeige, daß $\sim$ eine Äquivalenzrelation auf $M = \mathbb{R}^2 \setminus \{(0,0)\}$ ist. Es ist üblich die Äquivalenzklasse $\overline{(v_1, v_2)}$ von (v_1, v_2) mit $(v_1 : v_2)$ zu bezeichnen, und man nennt die Menge $M/\sim$ der Äquivalenzklassen die *projektive Gerade* über $\mathbb{R}$ und bezeichnet sie mit $\mathbb{P}^1_{\mathbb{R}}$.

b. Die Menge $S^1 = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ ist Kreis vom Radius Eins um den Mittelpunkt $(0,0)$. Zeige, daß die Abbilung

$$\Phi : S^1 \longrightarrow \mathbb{P}^1_{\mathbb{R}} : (x,y) \mapsto \overline{(x,y)}$$

surjektiv ist.

c. Wenn wir in der Definition von $\sim$ alle Elemente $v, w \in \mathbb{R}^2$ zulassen, definiert $\sim$ dann eine Äquivalenzrelation auf $\mathbb{R}^2$? Falls ja, was ist die Äquivalenzklasse von $(0,0)$?

Aufgabe A6.16 (Kongruenz modulo n)
Ist $n \in \mathbb{Z}_{>0}$ eine positive ganze Zahl, so definieren wir für $x, y \in \mathbb{Z}$

$$x \equiv y \;:\Longleftrightarrow\; x - y \text{ ist ein Vielfaches von } n.$$

Zeige, daß $\equiv$ eine Äquivalenzrelation ist mit genau den n paarweise verschiedenen Äquivalenzklassen $\overline{0}, \overline{1}, \ldots, \overline{n-1}$.

Man nennt zwei äquivalente Zahlen x und y dann auch *kongruent modulo* n.

ANHANG B

Grundlegende Algebraische Strukturen

In den folgenden Abschnitten werden einige grundlegende algebraische Strukuturen vorgestellt, die für die lineare Algebra wichtig sind. Die Abschnitte werden im Verlauf der Vorlesung jeweils dann eingefügt, wenn sie für das Verständnis und die Theorie wichtig sind.

§ B1 Gruppen und Homomorphismen

Es ist ein wichtiges Ziel der Mathematik, *Strukturen* auf Mengen zu studieren. Was eine *Struktur auf einer Menge* ganz konkret ist, hängt letztlich sehr vom Zweig der Mathematik und der betrachteten Fragestellung ab. In dieser Vorlesung wird die Struktur stets aus einer oder mehreren *zweistelligen Operationen* auf der Menge bestehen, die dann bestimmten *Gesetzmäßigkeiten*, sogenannten *Axiomen*, genügen sollen. Dabei ist eine *zweistellige Operation* auf einer Menge G eine Abbildung, die einem Paar (g, h) von Elementen aus G wieder ein Element in G zuordnet, also eine Abbildung $G \times G \to G$.

A) Gruppen

Die grundlegendste und wichtigste algebraische Struktur auf einer Menge ist die *Gruppenstruktur*. Im vorliegenden Abschnitt wollen wir einige wesentliche elementare Eigenschaften zusammenstellen, die auch in der *linearen* Algebra benötigt werden.

Definition B1.1 (Gruppen)

a. Eine *Gruppe* ist ein Paar $(G, *)$ bestehend aus einer *nicht-leeren* Menge G und einer zweistelligen Operation “$*$”, d. h. einer Abbildung

$$*: G \times G \to G : (g, h) \mapsto g * h,$$

so daß die folgenden *Gruppenaxiome* gelten:

G1: $(g * h) * k = g * (h * k) \quad \forall\, g, h, k \in G,$ (“Assoziativgesetz”)

G2: $\exists\, e \in G \;:\; \forall\, g \in G \;:\; e * g = g,$ (“Existenz eines Neutralen”)

G3: $\forall\, g \in G\; \exists\, g^{-1} \in G \;:\; g^{-1} * g = e.$ (“Existenz von Inversen”)

Ein Element mit der Eigenschaft von e nennt man *neutrales Element* der Gruppe G. Ein Element mit der Eigenschaft von g^{-1} nennt man ein *Inverses zu* g.

b. Eine Gruppe $(G, *)$ heißt *abelsch* oder *kommutativ*, wenn $(G, *)$ zudem noch dem folgenden Axiom genügt:

G4: $g * h = h * g \quad \forall\, g, h \in G$ (“Kommutativgesetz”)

Beispiel B1.2

a. $(\mathbb{Z},+)$, $(\mathbb{Q},+)$ und $(\mathbb{R},+)$ mit der üblichen Addition als Gruppenoperation sind abelsche Gruppen. Die Zahl Null erfüllt jeweils die Rolle eines neutralen Elements, und zu einer Zahl g existiert mit $-g$ ein inverses Element.

b. $(\mathbb{Q}\setminus\{0\},\cdot)$ und $(\mathbb{R}\setminus\{0\},\cdot)$ mit der üblichen Multiplikation als Gruppenoperation sind ebenfalls abelsche Gruppen. Die Zahl 1 ist jeweils ein neutrales Element, und zu einer Zahl g existiert als inverses Element die Zahl $\frac{1}{g}$.

c. Die einfachste Gruppe ist die *einelementige Gruppe* $G = \{e\}$, deren Gruppenoperation durch $e * e = e$ definiert ist.

d. Ist M eine Menge, so ist die Menge

$$\mathrm{Sym}(M) := \{f : M \longrightarrow M \mid f \text{ ist bijektiv}\}$$

mit der Komposition von Abbildungen als Gruppenoperation eine Gruppe. Die Assoziativität von "$\circ$" wird in Proposition A3.11 gezeigt, die Identität ist das neutrale Element und in Satz A3.12 wird gezeigt, daß jede bijektive Abbildung ein Inverses besitzt. Wir nennen $(\mathrm{Sym}(M),\circ)$ die *symmetrische Gruppe* auf M. Enthält M mehr als zwei Elemente, so ist $\mathrm{Sym}(M)$ nicht abelsch.

Bemerkung B1.3

Es sei $(G,*)$ eine Gruppe.

a. Das neutrale Element $e \in G$ ist eindeutig bestimmt und hat die Eigenschaft:

$$e * g = g * e = g \quad \forall\, g \in G.$$

b. Sei $g \in G$. Das inverse Element g^{-1} zu g ist eindeutig bestimmt und hat die Eigenschaft:

$$g^{-1} * g = g * g^{-1} = e.$$

c. Für $g, h \in G$ gelten $\left(g^{-1}\right)^{-1} = g$ und $(g * h)^{-1} = h^{-1} * g^{-1}$.

d. Für $g \in G$ und zwei ganze Zahlen $n, m \in \mathbb{Z}$ gelten die Potenzgesetze:

$$g^n * g^m = g^{n+m} \quad \text{und} \quad (g^m)^n = g^{m\cdot n}.$$

e. Wird die Gruppenoperation als Multiplikation und mit "$\cdot$" bezeichnet, so schreiben wir für das Neutrale Element meist 1 und für das Inverse zu g weiterhin g^{-1} oder $\frac{1}{g}$.
Wird die Gruppenoperation als Addition und mit "$+$" bezeichnet, so schreiben wir für das Neutrale Element meist 0 und für das Inverse zu g meist $-g$. Zudem schreiben wir statt $g + (-h)$ in aller Regel $g - h$.

f. In Ermangelung eines besseren Namens nennen wir auch "$*$" oft einfach die *Gruppenmultiplikation*.

Die Aussagen wollen wir in dieser Vorlesung nicht beweisen, aber wir fügen für den interessierten Leser einen Beweis ein.[1]

Beweis von Bemerkung B1.3: Da wir für das Paar $(G, *)$ die Axiome G1-G3 aus Definition B1.1 voraussetzen, gibt es ein neutrales Element $e \in G$, und zu beliebigem, aber fest gegebenem $g \in G$ gibt es ein Inverses $g^{-1} \in G$.

Wir wollen zunächst zeigen, daß für dieses e und dieses g^{-1} die in a. und b. geforderten zusätzlichen Eigenschaften gelten.

Da $(G, *)$ eine Gruppe ist, gibt es ein $(g^{-1})^{-1} \in G$ mit

$$(g^{-1})^{-1} * g^{-1} = e. \tag{55}$$

Also folgt:

$$\begin{aligned} g * g^{-1} &\overset{G2}{=} e * \left(g * g^{-1}\right) \overset{(55)}{=} \left((g^{-1})^{-1} * g^{-1}\right) * \left(g * g^{-1}\right) \overset{G1}{=} (g^{-1})^{-1} * \left(g^{-1} * (g * g^{-1})\right) \\ &\overset{G1}{=} (g^{-1})^{-1} * \left((g^{-1} * g) * g^{-1}\right) \overset{G3}{=} (g^{-1})^{-1} * \left(e * g^{-1}\right) \overset{G2}{=} (g^{-1})^{-1} * g^{-1} \overset{(55)}{=} e. \end{aligned} \tag{56}$$

Damit ist gezeigt, daß g^{-1} die zusätzliche Eigenschaft in b. erfüllt, und wir erhalten:

$$g * e \overset{G3}{=} g * \left(g^{-1} * g\right) \overset{G1}{=} \left(g * g^{-1}\right) * g \overset{(56)}{=} e * g \overset{G2}{=} g. \tag{57}$$

Nun war aber g ein beliebiges Element in G, so daß damit die zusätzliche Eigenschaft von e in a. gezeigt ist.

Sei nun $\tilde{e} \in G$ irgendein Element mit der Eigenschaft des Neutralen, d.h.

$$\tilde{e} * h = h \tag{58}$$

für alle $h \in G$. Wir müssen zeigen, daß $e = \tilde{e}$ gilt. Da wir bereits wissen, daß e die zusätzliche Eigenschaft in a. erfüllt, können wir diese, d.h. (57), mit $\tilde{e}$ in der Rolle von g anwenden, und anschließend (58) mit e in der Rolle von h:

$$\tilde{e} \overset{(57)}{=} \tilde{e} * e \overset{(58)}{=} e.$$

Schließlich müssen wir noch zeigen, wenn $\tilde{g}^{-1} \in G$ ein weiteres inverses Element zu g ist, d.h. wenn

$$\tilde{g}^{-1} * g = e \tag{59}$$

[1] Die Begriffe *Lemma*, *Proposition*, *Satz* und *Korollar* sind in der Mathematik übliche *Ordnungsstrukturen* (vergleichbar etwa einer Karteikarte), in denen bewiesene Aussagen festgehalten werden. Dabei werden die Aussagen, die als Propositionen formuliert werden, meist als wichtiger erachtet, als Aussagen in einem Lemma, und entsprechend sind Aussagen in einem Satz meist wesentlicher als Aussagen in einer Proposition. Das Korollar fällt etwas aus dem Rahmen, da es übersetzt *Folgerung* bedeutet und somit andeutet, daß es aus einer der unmittelbar zuvor getroffenen Aussagen abgeleitet werden kann. – Es kann vorkommen, daß der Beweis einer Aussage den Rahmen dieser Vorlesung sprengen würde oder wir aus anderen Gründen auf den Beweis verzichten wollen oder müssen. In einem solchen Fall werden wir die Aussage nur als *Bemerkung* formulieren und deutlich machen, weshalb wir auf einen Beweis verzichten.

gilt, dann ist schon $g^{-1} = \tilde{g}^{-1}$. Wenden wir das bislang Gezeigte an, so gilt:

$$\tilde{g}^{-1} \overset{(57)}{=} \tilde{g}^{-1} * e \overset{(56)}{=} \tilde{g}^{-1} * \left(g * g^{-1}\right) \overset{G1}{=} \left(\tilde{g}^{-1} * g\right) * g^{-1} \overset{(59)}{=} e * g^{-1} \overset{G2}{=} g^{-1}.$$

Damit sind die Aussagen in Teil a. und b. gezeigt und es bleibt noch, die Aussagen in Teil c. zu zeigen.

Um die erste Gleichheit zu zeigen, reicht es wegen der Eindeutigkeit des Inversen zu g^{-1} zu zeigen, daß g die Eigenschaft *des* Inversen zu g^{-1} besitzt. Beim Beweis können wir die Gruppenaxiome sowie die in a. und b. bewiesenen zusätzlichen Eigenschaften des Inversen anwenden:

$$g * g^{-1} \overset{b.}{=} e.$$

Also ist g ein Inverses zu g^{-1}, und damit gilt wie angedeutet wegen der Eindeutigkeit des Inversen zu g^{-1}:

$$\left(g^{-1}\right)^{-1} = g.$$

Analog ist nach Voraussetzung $(gh)^{-1}$ ein Inverses zu gh, und es reicht wegen der Eindeutigkeit des Inversen zu gh zu zeigen, daß $h^{-1}g^{-1}$ ebenfalls die Eigenschaft eines Inversen zu gh hat:

$$\begin{aligned}\left(h^{-1} * g^{-1}\right) * (g * h) &\overset{G1}{=} h^{-1} * \left(g^{-1} * (g * h)\right) \overset{G1}{=} h^{-1} * \left((g^{-1} * g) * h\right)\\ &\overset{G3}{=} h^{-1} * (e * h) \overset{G2}{=} h^{-1} * h \overset{G3}{=} e.\end{aligned}$$

Mithin ist $h^{-1} * g^{-1}$ ein Inverses zu gh, und somit

$$(g * h)^{-1} = h^{-1} * g^{-1}.$$

Die Potenzgesetze zeigt man dann mit Hilfe vollständiger Induktion. Wer den Beweis ausgeführt sehen möchte, sei auf [**Mar08**, Lemma 1.10] verwiesen. □

Lemma B1.4 (Kürzungsregeln)
Sei $(G, *)$ *eine Gruppe,* $g, a, b \in G$. *Dann gelten die* Kürzungsregeln*:*

a. $g * a = g * b \;\Rightarrow\; a = b$, *und*

b. $a * g = b * g \;\Rightarrow\; a = b$.

Beweis: Die erste Kürzungsregel folgt durch Multiplikation mit dem Inversen zu g von links:

$$\begin{aligned}a \overset{G2}{=} e * a \overset{G3}{=} \left(g^{-1} * g\right) * a &\overset{G1}{=} g^{-1} * (g * a)\\ &\overset{Vor.}{=} g^{-1} * (g * b) \overset{G1}{=} \left(g^{-1} * g\right) * b \overset{G3}{=} e * b \overset{G2}{=} b.\end{aligned}$$

Entsprechend folgt die zweite Kürzungsregel durch Multiplikation mit g^{-1} von rechts und unter Berücksichtigung der zusätzlichen Eigenschaft des Inversen in Bemerkung B1.3. Die Details überlassen wir dem Leser. □

B) Untergruppen

Ein wichtiges Prinzip in der Mathematik ist es, zu jeder betrachteten Struktur auch ihre *Unter-* oder *Teilstrukturen* zu betrachten. Für eine Menge sind das einfach ihre Teilmengen, für eine Gruppe werden es ihre *Untergruppen* sein – das sind Teilmengen, die die zusätzliche Struktur *respektieren.* Eine *Gruppe* besteht aus einer Menge G und zusätzlich einer zweistelligen Operation $\cdot : G \times G \to G$, die gewissen Axiomen genügt. Ist $U \subseteq G$ eine Teilmenge von G, so kann man die Abbildung "$\cdot$" auf $U \times U$ einschränken und erhält eine Abbildung

$$U \times U \longrightarrow G : (u, v) \mapsto u \cdot v,$$

wobei der Ausdruck $u \cdot v$ ein Element aus G ist, in aller Regel aber nicht in U liegt. Letzteres bedeutet, daß die Einschränkung von "$\cdot$" auf $U \times U$ in aller Regel keine zweistellige Operation auf U definiert! Das ist aber sicher eine Minimalforderung an U um zu sagen, daß U die Gruppenoperation "$\cdot$" respektiert. Nehmen wir nun an, daß wider Erwarten die Einschränkung von "$\cdot$" tatsächlich eine zweistellige Operation auf U definiert, dann stellt sich die Frage, ob das Paar bestehend aus U und der Einschränkung von "$\cdot$" den Gruppenaxiomen G1-G3 genügt, sprich selbst eine Gruppe ist – und erst in letzterem Fall kann man wirklich guten Gewissens sagen, die Teilmenge U respektiere die zusätzliche Struktur. Diese Überlegungen führen zum Begriff der *Untergruppe*, auch wenn dies aus unserer Definition nicht unmittelbar ersichtlich ist. Sie lassen sich im Übrigen auf alle weiteren von uns betrachteten algebraischen Strukturen übertragen.

Definition B1.5
Sei $(G, \cdot)$ eine Gruppe. Eine nicht-leere Teilmenge $U \subseteq G$ heißt *Untergruppe* von G, wenn

$$u \cdot v \in U \quad \text{und} \quad u^{-1} \in U$$

für alle $u, v \in U$. Ist U eine Untergruppe von G, so schreiben wir dafür $U \leq G$.

Bemerkung B1.6
Genau dann ist eine Teilmenge U von G mit der Einschränkung der Gruppenoperation selbst eine Gruppe, wenn U eine Untergruppe von G ist.

Beweis von Bemerkung B1.6: Sei zunächst U mit der Einschränkung der Gruppenoperation selbst eine Gruppe. Dies bedeutet, daß das Bild von $U \times U$ unter der Abbildung "$\cdot$" in U liegt, d. h. für $u, v \in U$ gilt $uv \in U$. Außerdem gelten in U die Gruppenaxiome. Sei also $e_U \in U$ das Neutrale in U und $e_G \in G$ das Neutrale in G. Ferner bezeichne zu $u \in U \subseteq G$ u_G^{-1} das Inverse zu u in G und u_U^{-1} das Inverse zu u in U, d. h. $u_G^{-1}u = uu_G^{-1} = e_G$ und $u_U^{-1}u = uu_U^{-1} = e_U$. In der folgenden Gleichung benötigen wir das Inverse von e_U in der Gruppe G, was in unserer Notation zu dem etwas unübersichtlichen $(e_U)_G^{-1}$ wird. Mit dieser Schreibweise gilt nun:

$$e_U \overset{\text{G2 in G}}{=} e_G e_U \overset{\text{G3 in G}}{=} \big((e_U)_G^{-1} e_U\big) e_U \overset{\text{G1 in G}}{=} (e_U)_G^{-1}(e_U e_U) \overset{\text{G2 in U}}{=} (e_U)_G^{-1} e_U \overset{\text{G3 in G}}{=} e_G. \tag{60}$$

Zudem gilt aber

$$u_U^{-1} u \overset{\text{G3 in U}}{=} e_U \overset{(60)}{=} e_G,$$

also ist $u_U^{-1} = u_G^{-1}$ wegen der Eindeutigkeit des Inversen in G, und damit $u_G^{-1} \in U$.

Sei nun umgekehrt vorausgesetzt, daß U eine Untergruppe von G ist. Da $uv \in U$ für alle $u, v \in U$, ist das Bild von $U \times U$ unter der Abbildung "$\cdot$" in der Tat in U enthalten. Es bleibt also, die Axiome G1-G3 nachzuprüfen. Dabei überträgt sich G1 von der größeren Menge G auf die Teilmenge U. Da $U \neq \emptyset$, existiert ein $u \in U$. Nach Voraussetzung gilt dann aber $u_G^{-1} \in U$ und damit $e_G = u_G^{-1} u \in U$. Da aber $e_G u = u$ für alle $u \in U$, ist auch G2 erfüllt und es gilt $e_U = e_G$. Ferner haben wir bereits bemerkt, daß für $u \in U$ auch $u_G^{-1} \in U$, und es gilt

$$u_G^{-1} \cdot u = e_G = e_U.$$

Somit ist auch G3 erfüllt und die Inversen von u in U bzw. in G stimmen überein. □

Man interessiert sich auch deshalb für die Untergruppen einer Gruppe, weil die Kenntnis dieser wichtige Informationen über die Struktur der Gruppe selbst liefert.

Beispiel B1.7

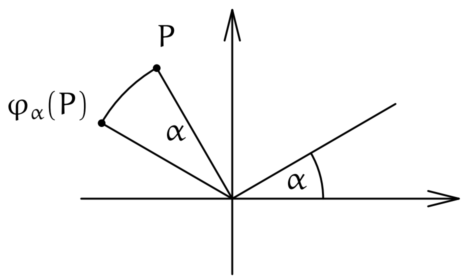
a. Ist $(G, \cdot)$ eine Gruppe mit neutralem Element e_G, so sind die beiden Teilmengen $\{e_G\}$ und G von G stets Untergruppen. Man nennt sie deshalb auch die *trivialen Untergruppen*. Sie geben keine zusätzliche Information über die Struktur der Gruppe selbst.

b. $(\{-1, 1\}, \cdot)$ ist eine Untergruppe von $(\mathbb{Q} \setminus \{0\}, \cdot)$.

c. Für $\alpha \in \mathbb{R}$ bezeichnet

$$\varphi_\alpha : \mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \to \big(\cos(\alpha) \cdot x - \sin(\alpha) \cdot y, \sin(\alpha) \cdot x + \cos(\alpha) \cdot y\big)$$

die Drehung der Ebene $\mathbb{R}^2$ um den Nullpunkt um den Winkel α im Bogenmaß.

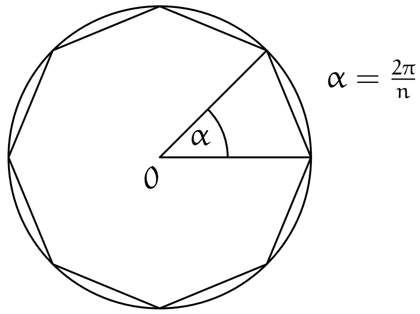


Offensichtlich gilt $\varphi_\alpha \circ \varphi_\beta = \varphi_{\alpha+\beta}$ für $\alpha, \beta \in \mathbb{R}$, und für $\alpha \in \mathbb{R}$ ist somit $\varphi_{-\alpha} = (\varphi_\alpha)^{-1}$, da $\varphi_0 = \mathrm{id}_{\mathbb{R}^2}$. Insbesondere ist φ_α also bijektiv für jedes $\alpha \in \mathbb{R}$. Also ist

$$SO(2) := \{\varphi_\alpha : \mathbb{R}^2 \to \mathbb{R}^2 \mid \alpha \in \mathbb{R}\}$$

eine Untergruppe von $\mathrm{Sym}\big(\mathbb{R}^2\big)$.

d. Sei $E_n \subset \mathbb{R}^2$ das *reguläre* n*-Eck*.



Die Menge

$$U := \{\varphi_\alpha \in SO(2) \mid \varphi_\alpha(E_n) = E_n\}$$

ist dann eine Untergruppe von $(SO(2), \circ)$.

Denn für $\varphi_\alpha, \varphi_\beta \in U$ gilt

$$(\varphi_\alpha \circ \varphi_\beta)(E_n) = \varphi_\alpha(\varphi_\beta(E_n)) = \varphi_\alpha(E_n) = E_n$$

und

$$\varphi_\alpha^{-1}(E_n) = \varphi_\alpha^{-1}(\varphi_\alpha(E_n)) = (\varphi_\alpha^{-1} \circ \varphi_\alpha)(E_n) = \mathrm{id}_{\mathbb{R}^2}(E_n) = E_n.$$

Also gilt $\varphi_\alpha \circ \varphi_\beta \in U$ und $\varphi_\alpha^{-1} \in U$, und da $\mathrm{id}_{\mathbb{R}^2} = \varphi_0 \in U$, ist $U \neq \emptyset$ und folglich ist U eine Untergruppe von $SO(2)$.

Man überzeugt sich leicht, daß U aus allen Drehungen φ_α mit $\alpha = k \cdot \frac{2\pi}{n}$, $k = 0, \ldots, n-1$, besteht. Insbesondere gilt also, $|U| = n$.

e. Die Untergruppen von $(\mathbb{Z}, +)$ sind genau die Mengen $n\mathbb{Z} := \{nz \mid z \in \mathbb{Z}\}$ aller Vielfachen von n für $n \in \mathbb{N}$, siehe Aufgabe B1.20.

f. Die Inklusionen $\mathbb{Z} \subset \mathbb{Q}$, $\mathbb{Z} \subset \mathbb{R}$ und $\mathbb{Q} \subset \mathbb{R}$ machen die Teilmenge bezüglich der Addition als Gruppenstruktur jeweils zu Untergruppen.

C) Gruppenhomomorphismen

Immer wenn man eine Struktur auf einer Menge definiert hat, spielen die *strukturerhaltenden Abbildungen* eine besondere Rolle. Diese werden (Struktur-)*Morphismen* oder (Struktur-)*Homomorphismen* genannt.

Definition B1.8

Es seien $(G, \cdot)$ und $(H, *)$ zwei Gruppen. Eine Abbildung $\alpha : G \to H$ heißt *Gruppenhomomorphismus* (oder kürzer *Homomorphismus* oder nur *Morphismus*), falls für alle $g, h \in G$ gilt:

$$\alpha(g \cdot h) = \alpha(g) * \alpha(h).$$

Wieder wollen wir uns zunächst Beispiele anschauen.

Beispiel B1.9

a. Ist $(G, \cdot)$ eine Gruppe und $U \leq G$ eine Untergruppe, dann ist die kanonische Inklusion $i_U : U \to G$ ein Gruppenhomomorphismus, da für $g, h \in U$ gilt $i_U(g \cdot h) = g \cdot h = i_U(g) \cdot i_U(h)$.

b. Sei $a \in \mathbb{R}$ und $m_a : (\mathbb{R},+) \to (\mathbb{R},+) : g \mapsto ag$ die Multiplikation mit a, dann ist m_a ein Gruppenhomomorphismus, da für $g, h \in \mathbb{R}$ gilt

$$m_a(g+h) = a(g+h) = ag + ah = m_a(g) + m_a(h).$$

c. Ist $(G, \cdot)$ eine Gruppe und $g \in G$, so definiert

$$i_g : G \to G : h \mapsto ghg^{-1} =: h^g$$

einen Gruppenhomomorphismus, die sogenannte *Konjugation* mit g, denn für $h, k \in G$ gilt

$$\begin{aligned} i_g(hk) =& g(hk)g^{-1} = g\big(hek\big)g^{-1} = g\big(h\big(g^{-1}g\big)k\big)g^{-1} \\ =& \big(ghg^{-1}\big)\big(gkg^{-1}\big) = i_g(h) \cdot i_g(k), \end{aligned}$$

Das folgende Lemma sagt, daß die Komposition zweier Homomorphismen stets wieder ein Homomorphismus ist.

Lemma B1.10
Sind $\alpha_1 : (G_1, \cdot) \to (G_2, *)$ und $\alpha_2 : (G_2, *) \to (G_3, \times)$ Gruppenhomomorphismen, so ist auch $\alpha_2 \circ \alpha_1 : (G_1, \cdot) \to (G_3, \times)$ ein Gruppenhomomorphismus.

Beweis: Seien $g, h \in G_1$, dann gilt:

$$\begin{aligned} (\alpha_2 \circ \alpha_1)(g \cdot h) = \alpha_2\big(\alpha_1(g \cdot h)\big) = \alpha_2\big(\alpha_1(g) * \alpha_1(h)\big) = \alpha_2\big(\alpha_1(g)\big) \times \alpha_2\big(\alpha_1(h)\big) \\ = (\alpha_2 \circ \alpha_1)(g) \times (\alpha_2 \circ \alpha_1)(h). \end{aligned}$$

□

Der Umstand, daß die Gruppenhomomorphismen die Gruppenstruktur *erhalten*, hat einige einfache, aber ungemein wichtige Auswirkungen.

Proposition B1.11
Es sei $\alpha : (G, \cdot) \to (H, *)$ ein Gruppenhomomorphismus. Dann gelten:

a. $\alpha(e_G) = e_H$.
b. $\alpha\big(g^{-1}\big) = \big(\alpha(g)\big)^{-1}$ *für* $g \in G$.
c. *Ist α bijektiv, so ist $\alpha^{-1} : H \to G$ ein Gruppenhomomorphismus.*
d. *Ist $U \leq G$, dann ist $\alpha(U) \leq H$. $\alpha(U)$ heißt das* Bild *von U unter α.*
e. *Ist $V \leq H$, dann ist $\alpha^{-1}(V) \leq G$. $\alpha^{-1}(V)$ heißt das* Urbild *von V unter α.*
f. Im$(\alpha) := \alpha(G)$, *das* Bild *von α, ist eine Untergruppe von H.*
g. Ker$(\alpha) := \alpha^{-1}(e_H)$, *der* Kern *von α, ist eine Untergruppe von G.*

Beweis: a. Es gilt

$$e_H * \alpha(e_G) = \alpha(e_G) = \alpha(e_G \cdot e_G) = \alpha(e_G) * \alpha(e_G).$$

Mit Hilfe der Kürzungsregel B1.4 folgt dann $e_H = \alpha(e_G)$.

b. Für $g \in G$ gilt:

$$\alpha\big(g^{-1}\big) * \alpha(g) = \alpha\big(g^{-1} \cdot g\big) = \alpha(e_G) = e_H.$$

Wegen der Eindeutigkeit der Inversen in H folgt die Behauptung.

c. Ist $\alpha : G \to H$ bijektiv, so existiert die Umkehrabbildung $\alpha^{-1} : H \to G$. Seien $u, v \in H$. Setze $g := \alpha^{-1}(u)$ und $h := \alpha^{-1}(v)$, also $u = \alpha(g)$ und $v = \alpha(h)$. Dann gilt:

$$\alpha^{-1}(u * v) = \alpha^{-1}\big(\alpha(g) * \alpha(h)\big) = \alpha^{-1}\big(\alpha(g \cdot h)\big) = g \cdot h = \alpha^{-1}(u) \cdot \alpha^{-1}(v).$$

Also ist α^{-1} ein Gruppenhomomorphismus.

d. Sind $u, v \in \alpha(U)$, dann existieren $g, h \in U$ mit $\alpha(g) = u$ und $\alpha(h) = v$. Da $g \cdot h \in U$, gilt:

$$u * v = \alpha(g) * \alpha(h) = \alpha(g \cdot h) \in \alpha(U).$$

Außerdem gilt $g^{-1} \in U$ und somit:

$$u^{-1} = \big(\alpha(g)\big)^{-1} = \alpha\big(g^{-1}\big) \in \alpha(U).$$

Da zudem $\alpha(e_G) \in \alpha(U)$, also $\alpha(U) \neq \emptyset$, folgt, daß $\alpha(U)$ eine Untergruppe von H ist.

e. Seien $g, h \in \alpha^{-1}(V)$, so gilt $\alpha(g \cdot h) = \alpha(g) * \alpha(h) \in V$, da V eine Untergruppe ist. Also gilt $g \cdot h \in \alpha^{-1}(V)$. Außerdem gilt $\alpha\big(g^{-1}\big) = \big(\alpha(g)\big)^{-1} \in V$, wieder da V eine Untergruppe ist. Somit liegt auch g^{-1} in $\alpha^{-1}(V)$. Da das Urbild von V unter α ferner nicht leer ist, weil wegen $\alpha(e_G) = e_H \in V$ gilt, daß $e_G \in \alpha^{-1}(V)$, folgt wieder, daß $\alpha^{-1}(V)$ eine Untergruppe von G ist.

f. Dies folgt aus e., da G eine Untergruppe von G ist.

g. Dies folgt aus f., da $\{e_H\}$ eine Untergruppe von H ist.

□

Nach Definition muß man für die Injektivität einer Abbildung nachprüfen, daß jedes Element im Bild nur ein Urbild hat. Bei Gruppenhomomorphismen gibt es ein einfacheres Kriterium.

Lemma B1.12
Ein Gruppenhomomorphismus $\alpha : (G, \cdot) \to (H, *)$ *ist genau dann injektiv, wenn* $\operatorname{Ker}(\alpha) = \{e_G\}$.

Beweis: Ist α injektiv, so ist $\alpha^{-1}(e_H)$ höchstens einelementig, und wegen $\alpha(e_G) = e_H$ gilt dann $\operatorname{Ker}(\alpha) = \alpha^{-1}(e_H) = \{e_G\}$.

Gilt umgekehrt $\operatorname{Ker}(\alpha) = \{e_G\}$, und sind $g, h \in G$ mit $\alpha(g) = \alpha(h)$, so folgt wegen:

$$e_H = \alpha(g) * \big(\alpha(h)\big)^{-1} = \alpha(g) * \alpha\big(h^{-1}\big) = \alpha\big(g \cdot h^{-1}\big),$$

daß $g \cdot h^{-1} = e_G$, also $g = h$. Somit ist α injektiv. □

Aufgaben

Aufgabe B1.13
Untersuche, ob die folgende zweistellige Operation die Menge $G := \mathbb{Q} \times \mathbb{Q}$ zu einer Gruppe macht:

$$G \times G \longrightarrow G : \big((a,b),(a',b')\big) \mapsto (a,b) \cdot (a',b') := (aa', bb').$$

Aufgabe B1.14
Es seien $(G, \cdot)$ und $(H, *)$ zwei Gruppen. Wir definieren auf der Menge $G \times H = \{(x,y) \mid x \in G, y \in H\}$ eine zweistellige Operation durch

$$(x,y) \circ (x',y') := (x \cdot x', y * y')$$

für $(x,y), (x',y') \in G \times H$. Zeige, daß dann $(G \times H, \circ)$ eine Gruppe ist.

Aufgabe B1.15
Untersuche, welche der folgenden zweistelligen Operationen Gruppen definieren:

a. $G = \big(\mathbb{Q}\backslash\{0\}\big) \times \big(\mathbb{Q}\backslash\{0\}\big)$ mit $(a,b)\cdot(a',b') = (ab', ba')$ für $a, a', b, b' \in \mathbb{Q}\backslash\{0\}$,
b. $G = \mathbb{R}\times\mathbb{R}\backslash\big\{(0,0)\big\}$ mit $(a,b)\cdot(a',b') = (aa'-bb', ab'+ba')$ für $a, a', b, b' \in \mathbb{R}$.

Aufgabe B1.16
Finde alle möglichen zweistelligen Operationen auf der Menge $G = \{e, a, b, c\}$, bezüglich derer G eine Gruppe mit neutralem Element e wird. Dabei sollten nur Möglichkeiten aufgelistet werden, die nicht durch Vertauschung der Buchstaben a, b und c ineinander überführt werden können.

Aufgabe B1.17 (Boolsche Gruppe)
Es sei M eine Menge.

a. Sind $X, Y, Z \subseteq M$, dann gelten

$$X \setminus \big((Y \setminus Z) \cup (Z \setminus Y)\big) = \big(X \setminus (Y \cup Z)\big) \cup (X \cap Y \cap Z)$$

und

$$\big((X \setminus Y) \cup (Y \setminus X)\big) \setminus Z = \big(X \setminus (Y \cup Z)\big) \cup \big(Y \setminus (X \cup Z)\big).$$

b. Wir definieren auf der Potenzmenge $G = \mathcal{P}(M) = \{A \mid A \subseteq M\}$ von M eine zweistellige Operation durch

$$A + B := (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B)$$

für $A, B \in G$. Zeige, $(G, +)$ ist eine abelsche Gruppe.

Aufgabe B1.18
Es sei $(G, \cdot)$ eine Gruppe, $g \in G$ und $\emptyset \neq U \subseteq G$ eine endliche Teilmenge von G.

a. Ist $\{g^n \mid n > 0\}$ endlich, so gibt es ein $n > 0$ mit $g^n = e_G$.
b. Genau dann ist U eine Untergruppe von G, wenn für alle $u, v \in U$ auch $u \cdot v \in U$.

Aufgabe B1.19
Sei M eine Menge, $m \in M$, $k \in \mathbb{N}$ und $\sigma \in \mathrm{Sym}(M)$ mit $\sigma^k(m) = m$. Zeige, dann ist auch $\sigma^{q\cdot k}(m) = m$ für alle $q \in \mathbb{Z}$.

Aufgabe B1.20
Zeige, eine Teilmenge $U \subseteq \mathbb{Z}$ ist genau dann eine Untergruppe von $(\mathbb{Z}, +)$, wenn es eine ganze Zahl $n \geq 0$ gibt mit $U = n\mathbb{Z} = \langle n \rangle$.

Aufgabe B1.21
Für zwei reelle Zahlen $a, b \in \mathbb{R}$ definieren wir die Abbildung

$$f_{a,b} : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto a \cdot x + b.$$

Welche der folgenden Mengen sind Untergruppen von $\big(\mathrm{Sym}(\mathbb{R}), \circ\big)$?

a. $U = \{f_{a,b} \mid a, b \in \mathbb{R}, a \neq 0\}$,

b. $V = \{f_{a,1} \mid a \in \mathbb{R}, a \neq 0\}$.

Aufgabe B1.22
Wir betrachten die Gruppe $U = \{f_{a,b} : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto ax + b \mid a, b \in \mathbb{R}, a \neq 0\}$ aus Aufgabe B1.21, wobei die Gruppenoperation die Verknüpfung von Abbildungen ist, sowie die Gruppe $(\mathbb{R} \setminus \{0\}, \cdot)$. Zeige, daß die Abbildung

$$\alpha : U \longrightarrow \mathbb{R} \setminus \{0\} : f_{a,b} \mapsto a$$

ein Gruppenhomomorphismus ist.

Aufgabe B1.23
Es sei $(G, \cdot)$ eine Gruppe und $h, k \in G$ gegeben. Prüfe, welche Bedingungen für h und k gelten müssen, damit α bzw. β ein Gruppenhomomorphismen ist:

a. $\alpha : G \to G : g \mapsto h \cdot g$,

b. $\beta : G \to G : g \mapsto h^{-1} \cdot g \cdot k$,

Aufgabe B1.24
Es sei $(G, \cdot)$ eine Gruppe. Zeige, $\mathrm{inv} : G \longrightarrow G : g \mapsto g^{-1}$ ist genau dann ein Gruppenhomomorphismus, wenn G abelsch ist.

Aufgabe B1.25
Es sei $\alpha : (G, \cdot) \longrightarrow (H, *)$ ein Gruppenhomomorphismus, $g \in G$ und $g' \in \mathrm{Ker}(\alpha)$. Zeige, dann gilt $g^{-1} \cdot g' \cdot g \in \mathrm{Ker}(\alpha)$.

Aufgabe B1.26
Es sei $(G, \cdot)$ eine Gruppe und $g \in G$.

a. Die Abbildung $L_g : G \longrightarrow G : h \mapsto g \cdot h$ ist bijektiv.

b. Die Abbildung $\alpha : G \longrightarrow \mathrm{Sym}(G) : g \mapsto L_g$ ist ein injektiver Gruppenhomomorphismus.

Aufgabe B1.27
Bestimme alle Gruppenhomomorphismen von $(\mathbb{Z}, +)$ nach $(\mathbb{R}, +)$.

§ B2 Die symmetrische Gruppe

Definition B2.1 (Permutationen)
Eine bijektive Abbildung $\sigma : \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}$ nennen wir eine *Permutation* der Menge $\{1, \ldots, n\}$, und wir bezeichnen mit

$$\mathbb{S}_n = \mathrm{Sym}(\{1, \ldots, n\}) = \big\{\sigma : \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\} \bigm| \sigma \text{ ist bijektiv}\big\}$$

die Menge aller Permutation der Menge $\{1, \ldots, n\}$.

Eine Permutation $\sigma \in \mathbb{S}_n$ kann durch eine *Wertetabelle* der folgenden Form beschrieben werden:

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}$$

bzw.

$$\begin{pmatrix} a_1 & a_2 & \ldots & a_n \\ \sigma(a_1) & \sigma(a_2) & \ldots & \sigma(a_n) \end{pmatrix},$$

falls $a_1, \ldots, a_n$ irgendeine Anordnung der Zahlen $1, \ldots, n$ ist.

Bemerkung B2.2 (Die symmetrische Gruppe $\mathbb{S}_n$)
In Beispiel B1.2 haben wir gezeigt, daß $\mathbb{S}_n$ mit der Komposition von Abbildungen eine Gruppe ist. Wir nennen $(\mathbb{S}_n, \circ)$ die *symmetrische Gruppe* vom Grad n. Die $\mathbb{S}_n$ enthält genau $n!$ Elemente.

Beispiel B2.3
Die Gruppe $\mathbb{S}_n$ ist für $n \geq 3$ nicht abelsch. In $\mathbb{S}_3$ gilt für die Permutationen

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \in \mathbb{S}_3$$

nämlich

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

Beachte, daß es bei dem Schema nicht darauf ankommt, in welcher Reihenfolge die Zahlen von 1 bis n in der ersten Zeile stehen. Es gilt etwa:

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

Es empfiehlt sich aber der Übersichtlichkeit halber für gewöhnlich, die Ziffern in aufsteigender Reihenfolge anzuordnen.

Bemerkung B2.4 (Invertieren einer Permutation)
Die oben eingeführte Darstellung einer Permutation hat den angenehmen Nebeneffekt, daß man das Inverse der Permutation leicht angeben kann, indem man einfach die beiden Zeilen vertauscht. Sprich, für eine Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix} \in \mathbb{S}_n$$

ist das Inverse σ^{-1} gegeben durch

$$\sigma^{-1} = \begin{pmatrix} \sigma(1) & \sigma(2) & \ldots & \sigma(n) \\ 1 & 2 & \ldots & n \end{pmatrix}.$$

Definition B2.5 (Zyklen und Transpositionen)

a. Sei $\{1, \ldots, n\} = \{a_1, \ldots, a_k\} \dot{\cup} \{b_1, \ldots, b_{n-k}\}$, $k \geq 2$, und

$$\sigma = \begin{pmatrix} a_1 & a_2 & \ldots & a_{k-1} & a_k & b_1 & \ldots & b_{n-k} \\ a_2 & a_3 & \ldots & a_k & a_1 & b_1 & \ldots & b_{n-k} \end{pmatrix} \in \mathbb{S}_n,$$

so heißt σ ein **k-Zyklus**, und wir sagen, daß sie die Zahlen $a_1, \ldots, a_k$ *zyklisch vertauscht*. Die Abbildungsvorschrift eines solchen k-Zyklus läßt sich deutlich kompakter durch das folgende einzeilige Schema repräsentieren:

$$\sigma = (a_1 \ldots a_k). \tag{61}$$

b. Ein 2-Zyklus wird auch eine **Transposition** genannt. Eine Transposition $\tau = (i\ j)$ ist mithin eine Permutation, die nur die zwei Zahlen i und j miteinander vertauscht, alle anderen aber fest läßt.

c. Das neutrale Element von $\mathbb{S}_n$, per definitionem $\mathrm{id}_{\{1,\ldots,n\}}$, wollen wir der Einfachheit halber mit id bezeichnen.

Bemerkung B2.6

a. Die Interpretation der Schreibweise in Gleichung (61) ist offensichtlich, das erste Element a_1 wird auf das zweite a_2 abgebildet, das zweite auf das dritte, und so weiter, bis schließlich das letzte, nämlich a_k, auf das erste, das heißt auf a_1, abgebildet wird – der *Kreis* schließt sich. Beachte hierbei, daß die Zyklen $(a_1 \ldots a_k)$, $(a_k a_1 \ldots a_{k-1})$, etc. übereinstimmen! Um diese Mehrdeutigkeit zu vermeiden, empfiehlt es sich, einen Zyklus stets mit der kleinsten der Zahlen $a_1, \ldots, a_k$ zu beginnen.

Bisher haben wir k-Zyklen nur für $k \geq 2$ definiert. Wir können nun auch 1-Zyklen, etwa (1) oder (3), zulassen und definieren diese in natürlicher Weise als die Identität.

b. Die Permutationen

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix} \in \mathbb{S}_4 \quad \text{und} \quad \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 3 & 2 & 5 \end{pmatrix} \in \mathbb{S}_5$$

sind jeweils 3-Zyklen, die die Zahlen $1, 4, 2$ zyklisch vertauschen. In der oben eingeführten Zyklenschreibweise gilt

$$\sigma = (1\ 4\ 2) \quad \text{und} \quad \pi = (1\ 4\ 2).$$

Damit wird der Nachteil dieser Schreibweise gegenüber dem zweizeiligen Schema deutlich – weder der Definitionsbereich noch der Wertebereich lassen sich aus der Zyklenschreibweise eindeutig ablesen. Aber diesen Preis sind wir für die gewonnene *Übersichtlichkeit* gerne bereit zu zahlen. Denn einerseits ist in Anwendungen meist zweifelsfrei bekannt, was n ist, und andererseits ist die wesentliche Information für uns letztlich, welche Zahlen durch die Permutation vertauscht werden, und nicht, welche unbewegt bleiben.

c. Für eine Transposition $\tau \in \mathbb{S}_n$ gilt $\tau^{-1} = \tau$, also $\tau^2 = \mathrm{id}$.

d. Für kleine Werte n ist $\mathbb{S}_n$ sehr übersichtlich, für große Werte n wird $\mathbb{S}_n$ jedoch riesig. $\mathbb{S}_1 = \{\mathrm{id}\}$ und $\mathbb{S}_2 = \{\mathrm{id}, (1\ 2)\}$. $\mathbb{S}_3 = \{\mathrm{id}, (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)\}$ hat schon sechs Elemente, $\mathbb{S}_4$ gar 24 und $\mathbb{S}_{60}$ ungefähr 10^{82}. Letztere Zahl entspricht in etwa der angenommenen Anzahl der Nukleone des Universums.

Satz B2.7 (Zyklenzerlegung und Signum)

a. *Jede Permutation* $\sigma \in \mathbb{S}_n$ *läßt sich als Produkt von disjunkten Zyklen schreiben.*

b. *Jede Permutation* $\sigma \in \mathbb{S}_n$ *läßt sich als Produkt von Transpositionen schreiben.*

c. *Jede Permutation* $\sigma \in \mathbb{S}_n$ *läßt sich als Produkt von Transpositionen benachbarter Zahlen schreiben.*

d. *Es gibt genau einen Gruppenhomomorphismus, das* Signum *genannt,*

$$\mathrm{sgn} : (\mathbb{S}_n, \circ) \longrightarrow \big(\{1, -1\}, \cdot\big)$$

mit $\mathrm{sgn}(\tau) = -1$ *für jede Transposition* $\tau \in \mathbb{S}_n$. *Insbesondere gilt, ist* $\sigma \in \mathbb{S}_n$ *ein Produkt von* k *Transpositionen, dann gilt mithin*

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

e. *Für* $\sigma \in \mathbb{S}_n$ *gilt* $\mathrm{sgn}(\sigma) = \mathrm{sgn}\left(\sigma^{-1}\right)$.

f. *Ist* $\mathbb{A}_n = \{\sigma \in \mathbb{S}_n \mid \mathrm{sgn}(\sigma) = 1\}$ *und ist* $\tau = (i\ j)$ *eine Transposition, so gilt*

$$\mathbb{S}_n = \mathbb{A}_n \mathbin{\dot\cup} \mathbb{A}_n\tau$$

wobei $\mathbb{A}_n\tau = \{\sigma \circ \tau \mid \sigma \in \mathbb{A}_n\}$.

Bemerkung B2.8

a. Daß die Abbildung sgn ein *Gruppenhomomorphismus* ist, heißt

$$\mathrm{sgn}(\sigma \cdot \pi) = \mathrm{sgn}(\sigma) \cdot \mathrm{sgn}(\pi)$$

für alle $\sigma, \pi \in \mathbb{S}_n$. Ist also $\sigma = \tau_1 \circ \cdots \circ \tau_k \in \mathbb{S}_n$ ein Produkt von k Transpositionen, dann gilt

$$\mathrm{sgn}(\sigma) = \mathrm{sgn}(\tau_1) \cdot \ldots \cdot \mathrm{sgn}(\tau_k) = (-1)^k.$$

b. Für den Beweis verweisen wir auf [**Mar08**, §3]. Wir wollen uns hier damit begnügen, an einem Beispiel zu zeigen, was die Aussagen bedeuten und wie man die Zerlegungen bzw. das Signum berechnen kann.

Beispiel B2.9 (Zyklenzerlegung)

Die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 3 & 1 \end{pmatrix} \in \mathbb{S}_5$$

hat die Zyklenzerlegung

$$\sigma = (1\ 2\ 5) \circ (3\ 4) = (3\ 4) \circ (1\ 2\ 5). \tag{62}$$

Eine berechtigte Frage ist, wie wir die Zyklenzerlegung in (62) gefunden haben. Wir wollen versuchen, dies so in Worte zu fassen, daß dem Leser daraus die allgemeine Vorgehensweise ersichtlich wird. Man starte mit der kleinsten Zahl, 1, und suche ihr Bild unter σ, also $\sigma(1) = 2$. Das liefert den Startteil des ersten Zyklus:

$$(1\ 2$$

Sodann betrachte man das Bild von 2 unter σ, also $\sigma(2) = 5$, und erhält:

$$(1\ 2\ 5$$

Man fährt mit dem Bild von 5 unter σ, also $\sigma(5) = 1$, fort. Da dieses das erste Element des ersten Zyklus war, schließen wir den Zyklus,

$$(1\ 2\ 5),$$

und beginnen den zweiten Zyklus mit der kleinsten Zahl in $\{1, \ldots, 5\}$, die noch nicht in dem ersten Zyklus vorkommt, also mit 3:

$$(1\ 2\ 5) \circ (3$$

Dann betrachten wir deren Bild unter σ, also $\sigma(3) = 4$, und setzen so unseren zweiten Zyklus fort:

$$(1\ 2\ 5) \circ (3\ 4$$

Da bereits alle fünf Elemente von $\{1, \ldots, 5\}$ aufgebraucht sind, muß notwendig $\sigma(4) = 3$ gelten, was es auch tut, und wir können damit auch den zweiten Zyklus schließen:

$$\sigma = (1\ 2\ 5) \circ (3\ 4).$$

Wie gesagt, da in $\{1, \ldots, 5\}$ keine Zahl mehr übrig ist, sind wir fertig und haben die Zyklenzerlegung von σ gefunden. $\square$

Beispiel B2.10 (Zerlegung in Transpositionen)
Wir wollen nun zeigen, wie man die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 8 & 7 & 1 & 4 & 5 & 6 & 9 & 2 \end{pmatrix} \in \mathbb{S}_9$$

als Produkt von Transpositionen schreiben kann und wie man ihr Signum berechnet.

Dazu zerlegen wir sie zunächst in ein Produkt disjunkter Zyklen, und mit Hilfe des Verfahrens aus Beispiel B2.9 erhalten wir

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 8 & 7 & 1 & 4 & 5 & 6 & 9 & 2 \end{pmatrix} = (1\ 3\ 7\ 6\ 5\ 4) \circ (2\ 8\ 9).$$

Dann schreiben wir die Zyklen als Produkte von Transpositionen:

$$(1\ 3\ 7\ 6\ 5\ 4) = (1\ 3) \circ (3\ 7) \circ (7\ 6) \circ (6\ 5) \circ (5\ 4)$$

und

$$(2\,8\,9) = (2\,8) \circ (8\,9).$$

Die Ergebnisse können wir dann zusammensetzen und erhalten

$$\sigma = (1\,3) \circ (3\,7) \circ (7\,6) \circ (6\,5) \circ (5\,4) \circ (2\,8) \circ (8\,9).$$

Aus dem Beispiel läßt sich leicht ein allgemeines Verfahren ableiten, um *eine* solche Zerlegung zu berechnen. Man sollte beachten, daß die Zerlegung in ein Produkt nicht eindeutig ist. Sie läßt sich auf viele Arten variieren. Wichtig ist sie allein, um das Signum zu berechnen, denn es gilt

$$\mathrm{sgn}(\sigma) = (-1)^7 = -1,$$

da σ Produkt von sieben Transpositionen ist.

Will man die Permutation gar als Produkt von Transpositionen benachbarter Zahlen schreiben, so reicht es, zu zeigen, wie man eine beliebige Transposition als Produkt solcher Transpositionen schreiben kann. Dann kann man das Verfahren auf jede Transposition in der obigen Zerlegung anwenden. Wir führen dies hier nur am Beispiel der Transposition $(3\,7)$ vor. Das allgemeine Verfahren kann man daraus leicht ablesen:

$$(3\,7) = (3\,4) \circ (4\,5) \circ (5\,6) \circ (6\,7) \circ (5\,6) \circ (4\,5) \circ (3\,4).$$

Aufgaben

Aufgabe B2.11

Betrachte die Permutationen

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 6 & 3 & 7 & 5 & 1 & 4 \end{pmatrix}, \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 2 & 7 & 1 & 4 & 3 & 6 \end{pmatrix} \in \mathbb{S}_7.$$

a. Berechne $\sigma \circ \pi$, $\pi \circ \sigma$, σ^{-1}, π^{-1}.

b. Bestimme für jede der Permutationen in a. die Zyklenzerlegung.

c. Schreibe $\sigma \circ \pi$ als ein Produkt von Transpositionen.

d. Schreibe π^{-1} als ein Produkt von Transpositionen aufeinander folgender Zahlen.

e. Berechne für jede der Permutationen in a. das Signum.

§ B3 Körper

Definition B3.1 (Körper)
Ein *Körper* ist ein Tripel $(K, +, \cdot)$ bestehend aus einer Menge K zusammen mit zwei zweistelligen Operationen

$$+ : K \times K \to K : (x, y) \mapsto x + y, \qquad (\text{“}\textit{Addition}\text{”})$$

und

$$\cdot : K \times K \to K : (x, y) \mapsto x \cdot y, \qquad (\text{“}\textit{Multiplikation}\text{”})$$

so daß folgende Axiome erfüllt sind:

a. $(K, +)$ ist eine abelsche Gruppe mit neutralem Element 0.
b. $(K \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe mit neutralem Element 1.
c. Es gilt das *Distributivgesetz* $x \cdot (y + z) = x \cdot y + x \cdot z$ für $x, y, z \in K$.

Ist eine Teilmenge $L \subseteq K$ eines Körpers mit den *gleichen* Operationen wieder selbst ein Körper, so nennen wir L einen *Teilkörper* von K.

Beispiel B3.2 (Die endlichen Körper $\mathbb{F}_p$)

a. Die rationalen Zahlen $(\mathbb{Q}, +, \cdot)$ und die reellen Zahlen $(\mathbb{R}, +, \cdot)$ mit der üblichen Addition und Multiplikation sind Körper. $\mathbb{Q}$ ist ein Teilkörper von $\mathbb{R}$.
b. Die ganzen Zahlen $(\mathbb{Z}, +, \cdot)$ sind kein Körper, da z.B. der Zahl 2 ein multiplikatives Inverses fehlt.
c. Auf der Menge $\mathbb{F}_2 := \{0, 1\}$ definieren wir zwei Operationen durch folgende Additions- und Multiplikationstafeln:

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Mit ein wenig Aufwand kann man nachrechnen, daß alle Körperaxiome erfüllt sind und daß mithin $\mathbb{F}_2$ ein Körper ist. $\mathbb{F}_2$ ist der kleinstmögliche Körper, da nach Definition ein Körper stets mindestens zwei Elemente, nämlich ein Neutrales bezüglich der Addition und ein davon verschiedenes Neutrales bezüglich der Multiplikation enthalten muß. Man beachte auch, daß aufgrund von Lemma B3.4 keine andere Möglichkeit für die obigen Verknüpfungstafeln besteht, wenn man einen Körper mit genau zwei Elementen haben möchte. — Beachte auch, daß $\mathbb{F}_2$ kein Teilkörper von $\mathbb{R}$ ist, da das Ergebnis von $1+1$ in den beiden Körpern nicht übereinstimmt.

d. Allgemeiner zeigt man, daß man für eine Primzahl p die Menge

$$\mathbb{F}_p := \{0, 1, \ldots, p-1\}$$

auf folgende Weise zu einem Körper machen kann. Für eine natürliche Zahl $a \in \mathbb{N}$ können wir Division mit Rest durch die Zahl p durchführen. Wir erhalten dann eindeutig bestimmte Zahlen $q \in \mathbb{N}$ und $0 \le r < p$ mit

$$a = q \cdot p + r.$$

Die Zahl r heißt der Rest von a bei Division mit Rest durch p, und wir bezeichnen sie $r(a : p)$.
Mit dieser Notation definieren wir für zwei Zahlen $a, b \in \mathbb{F}_p$

$$a + b := r(a + b : p)$$

und

$$a \cdot b := r(a \cdot b : p),$$

wobei das "+" bzw. das "·" auf der rechten Seite jeweils die Operation in den ganzen Zahlen bezeichnet, während das "+" und das "·" auf der linken Seite neu definierte Operationen sind. Formal wäre es besser, für diese neuen Operationen neue Symbole zu verwenden, etwa "⊕" und "⊗", aber Mathematiker sind bequeme Menschen und schreiben nur ungerne mehr als nötig. Deshalb bleiben wir bei den bewährten Symbolen und müssen nur drauf achten, wo wir gerade rechnen. Jedenfalls gilt, daß $\mathbb{F}_p$ mit diesen beiden Operationen ein Körper ist.
Man beachte auch, daß in $\mathbb{F}_p$ für jede Primzahl p stets

$$\underbrace{1 + 1 + \ldots + 1}_{p-\text{mal}} = r(p : p) = 0$$

gilt! Damit ist auch das Negative einer Zahl $a \in \mathbb{F}_p$ leicht zu berechnen als $p - a$, hingegen ist das multiplikative Inverse $\frac{1}{a}$ einer Zahl $0 \neq a \in \mathbb{F}_p$ nicht so ohne weiteres anzugeben. Man lernt in den weiterführenden Vorlesungen, wie man dieses mit Hilfe des Euklidischen Algorithmus' berechnen kann.
Z.B., gilt in $\mathbb{F}_5$

$$3 + 4 = r(3 + 4 : 5) = r(7 : 5) = 2$$

und

$$3 \cdot 4 = r(3 \cdot 4 : 5) = r(12 : 5) = 2.$$

Man schreibt übrigens meist $\mathbb{Z}/p\mathbb{Z}$ oder $\mathbb{Z}_p$ anstatt $\mathbb{F}_p$, und die Zahl a wird meist mit $\overline{a}$ oder $[a]$ bezeichnet. Das liegt daran, daß man den Körper mit der Menge der Äquivalenzklassen der Kongruenz modulo p identifizieren kann (siehe Aufgabe A6.16).

Notation B3.3
Ist K ein Körper und sind $x, y, z \in K$ mit $z \neq 0$, so schreiben wir statt $x + (-y)$ in aller Regel $x - y$, und statt $x \cdot z^{-1}$ schreiben wir oft $\frac{x}{z}$. Außerdem schreiben wir statt $x \cdot y$ meist nur xy.

Lemma B3.4 (Rechenregeln)
Es sei K *ein Körper,* $x, y, z \in K$ *und* $u, v \in K \setminus \{0\}$.

a. $-(-x) = x$,

b. $x + y = z \Leftrightarrow x = z - y$,

c. $-(x + y) = -x - y$,

d. $0 \cdot x = x \cdot 0 = 0$,

e. $(-x) \cdot y = x \cdot (-y) = -(x \cdot y)$,

f. $(-x) \cdot (-y) = x \cdot y$,

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

h. $\left(x^{-1}\right)^{-1} = x$, *für* $x \neq 0$,

i. $x \cdot y = 0 \Leftrightarrow x = 0$ *oder* $y = 0$,

j. $z \cdot x = z \cdot y,\ z \neq 0 \Rightarrow x = y$,

k. $\frac{x}{u} \cdot \frac{y}{v} = \frac{x \cdot y}{u \cdot v}$,

l. $\frac{x}{u} + \frac{y}{v} = \frac{x \cdot v + y \cdot u}{u \cdot v}$.

Beweis: Die Aussagen a., b., c. und h. folgen unmittelbar aus Bemerkung B1.3 und Lemma B1.4.

d. Für $x \in K$ gilt $0 \cdot x = (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x$, also folgt $0 \cdot x = 0$ mittels der Kürzungsregeln in $(K, +)$. Analog sieht man $x \cdot 0 = 0$.

e. Für $x, y \in K$ gilt wegen d.:
$$x \cdot y + (-x) \cdot y = (x - x) \cdot y = 0 \cdot y = 0,$$
also $-(x \cdot y) = (-x) \cdot y$. Die Gleichheit des Ausdrucks zu $x \cdot (-y)$ folgt analog.

f. Für $x, y \in K$ folgt unter Zuhilfenahme von a. und e.:
$$(-x) \cdot (-y) = -\big(x \cdot (-y)\big) = -\big(-(x \cdot y)\big) = x \cdot y.$$

g. Für $x, y, z \in K$ impliziert e.:
$$x \cdot (y - z) = x \cdot y + x \cdot (-z) = x \cdot y + \big(-(x \cdot z)\big) = x \cdot y - x \cdot z.$$

i. Ist $x = 0$ oder $y = 0$, so ist nach d. auch $x \cdot y = 0$. Ist $x \neq 0$ und $y \neq 0$, so ist $x \cdot y \in K \setminus \{0\}$, da $K \setminus \{0\}$ bezüglich der Multiplikation abgeschlossen ist.

j. Die Aussage zeigt man genau wie die Kürzungsregeln für Gruppen (siehe Lemma B1.4).

k. Unter Beachtung der Assoziativität und Kommutativität der Multiplikation sowie der Notation B3.3 gilt
$$\frac{x}{u} \cdot \frac{y}{v} = \left(x \cdot u^{-1}\right) \cdot \left(y \cdot v^{-1}\right) = (x \cdot y) \cdot (u \cdot v)^{-1} = \frac{x \cdot y}{u \cdot v}.$$

l. Dies geht analog zu k. mit etwas mehr Schreibarbeit.

□

Notation B3.5 (Produkte und Summen)
Es sei K ein Körper und $x_0, \ldots, x_n \in K$ seien $n+1$ Elemente in K, $n \in \mathbb{N}$. Wir schreiben

$$\prod_{i=0}^{n} x_i := x_0 \cdot \ldots \cdot x_n$$

für das *Produkt* der Zahlen $x_0, \ldots, x_n$ und

$$\sum_{i=0}^{n} x_i := x_0 + \ldots + x_n$$

für die *Summe* der Zahlen $x_0, \ldots, x_n$.

Außerdem definieren wir für $x \in K$ und $n \in \mathbb{N}$ die *Potenzen* von x durch

$$x^n := \underbrace{x \cdot \ldots \cdot x}_{n-\text{mal}} = \prod_{i=1}^{n} x$$

falls $n \geq 1$ sowie $x^0 := 1$. Ist zudem $x \neq 0$, so definieren wir

$$x^{-n} := \left(x^{-1}\right)^n = \underbrace{x^{-1} \cdot \ldots \cdot x^{-1}}_{n-\text{mal}} = \frac{1}{x^n}.$$

Analog dazu setzen wir

$$n \cdot x := \underbrace{x + \ldots + x}_{n-\text{mal}} = \sum_{i=1}^{n} x$$

und

$$(-n) \cdot x := n \cdot (-x) = \underbrace{(-x) + \ldots + (-x)}_{n-\text{mal}}$$

für $n \geq 1$, sowie $0 \cdot x = 0$.

Bemerkung B3.6 (Rekursionsprinzip)
Dem Prinzip der vollständigen Induktion ist das *Rekursionsprinzip* eng verwandt. Wollen wir einen Ausdruck für alle natürlichen Zahlen definieren, so definieren wir ihn für die Zahl 0 und führen die Definition für die Zahl n auf die Definition für die Zahl $n-1$ zurück.

Die Notation mit Punkten "$\ldots$" in Notation B3.5 ist stets eine versteckte Induktion oder Rekursion. Formal korrekt wäre es das Produkt rekursiv zu definieren durch $\prod_{i=0}^{0} x_i := x_0$ und $\prod_{i=0}^{n} x_i := \left(\prod_{i=0}^{n-1} x_i\right) \cdot x_n$. Analog sollte man die Summe rekursiv definieren durch $\sum_{i=0}^{0} x_i := x_0$ und $\sum_{i=0}^{n} x_i := \left(\sum_{i=0}^{n-1} x_i\right) + x_n$. Und für die Definition von x^n und $n \cdot x$ gilt Entsprechendes.

Beispiel B3.7 (Gauß)
Die Summe der natürlichen Zahlen bis zu einer gegebenen Zahl n ist

$$\sum_{k=0}^{n} k = \frac{n \cdot (n+1)}{2}.$$

Man beweist die Aussage durch Induktion nach n, wobei sie für $n = 0$ offenbar richtig ist. Nehmen wir nun an, daß sie für n gilt, so folgt

$$\sum_{k=0}^{n+1} k = \sum_{k=0}^{n} k + (n+1) \overset{\text{Ind}}{=} \frac{n \cdot (n+1)}{2} + (n+1) = \frac{(n+1) \cdot (n+2)}{2}.$$

Also gilt die Aussage für alle $n \in \mathbb{N}$ nach dem Prinzip der vollständigen Induktion.

Satz B3.8 (Endliche geometrische Reihe)
Ist K *ein Körper,* $1 \neq q \in K$ *und* $n \in \mathbb{N}$, *so gilt*

$$\sum_{k=0}^{n} q^k = \frac{1 - q^{n+1}}{1 - q}.$$

Beweis: Der Beweis ist eine einfache Anwendung des Prinzips der vollständigen Induktion. □

Definition B3.9 (Fakultät)
Für eine natürliche Zahl $n \in \mathbb{N}$ definieren wir die *Fakultät* durch

$$n! := \prod_{i=1}^{n} i = 1 \cdot \ldots \cdot n,$$

falls $n \geq 1$, und durch $0! := 1$.

Für zwei natürliche Zahlen $k, n \in \mathbb{N}$ erklären wir den *Binomialkoeffizienten* von n über k durch

$$\binom{n}{k} := \frac{n!}{(n-k)! \cdot k!} = \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{k \cdot (k-1) \cdot \ldots \cdot 1},$$

falls $0 \leq k \leq n$, und durch $\binom{n}{k} := 0$ sonst.

Proposition B3.10 (Binomialkoeffizienten)
Es seien $n, k \in \mathbb{N}$ *natürliche Zahlen. Dann gilt*

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

Beweis: Wir unterscheiden mehrere Fälle.

1. Fall: $k = 0$:

$$\binom{n+1}{k} = \binom{n+1}{0} = 1 = 0 + 1 = \binom{n}{-1} + \binom{n}{0} = \binom{n}{k-1} + \binom{n}{k}.$$

2. Fall: $k = n+1$:

$$\binom{n+1}{k} = \binom{n+1}{n+1} = 1 = 1 + 0 = \binom{n}{n} + \binom{n}{n+1} = \binom{n}{k-1} + \binom{n}{k}.$$

3. Fall: $k < 0$ **oder** $k > n+1$**:**

$$\binom{n+1}{k} = 0 = 0 + 0 = \binom{n}{k-1} + \binom{n}{k}.$$

4. Fall: $1 \leq k \leq n$**:**

$$\begin{aligned}\binom{n}{k-1} + \binom{n}{k} &= \frac{n!}{(n+1-k)! \cdot (k-1)!} + \frac{n!}{(n-k)! \cdot k!}\\ &= \frac{n! \cdot k}{(n+1-k)! \cdot k!} + \frac{n! \cdot (n+1-k)}{(n+1-k)! \cdot k!}\\ &= \frac{n! \cdot (k+n+1-k)}{(n+1-k)! \cdot k!} = \frac{(n+1)!}{(n+1-k)! \cdot k!} = \binom{n+1}{k}.\end{aligned}$$

□

Satz B3.11 (Binomischer Lehrsatz)
Es sei K *ein Körper,* $x, y \in K$ *und* $n \in \mathbb{N}$*, so gilt*

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

Beweis: Wir führen den Beweis durch Induktion nach n.

Induktionsanfang: $n = 0$**:** Nach Definition gilt

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

Induktionsschluß: $n \mapsto n+1$**:** Es gilt

$$\begin{aligned}(x+y)^{n+1} &= (x+y)^n \cdot (x+y) = (x+y)^n \cdot x + (x+y)^n \cdot y\\ &\overset{\text{Ind.}}{=} \sum_{k=0}^{n} \binom{n}{k} \cdot x^{k+1} \cdot y^{n-k} + \sum_{k=0}^{n} \binom{n}{k} \cdot x^k \cdot y^{n+1-k}\\ &= x^{n+1} + \sum_{k=0}^{n-1} \binom{n}{k} \cdot x^{k+1} \cdot y^{n-k} + \sum_{k=1}^{n} \binom{n}{k} \cdot x^k \cdot y^{n+1-k} + y^{n+1}\\ &= x^{n+1} + \sum_{k=1}^{n} \binom{n}{k-1} \cdot x^k \cdot y^{n+1-k} + \sum_{k=1}^{n} \binom{n}{k} \cdot x^k \cdot y^{n+1-k} + y^{n+1}\\ &\overset{\text{B3.10}}{=} x^{n+1} + \sum_{k=1}^{n} \binom{n+1}{k} \cdot x^k \cdot y^{n+1-k} + y^{n+1}\\ &= \sum_{k=0}^{n+1} \binom{n+1}{k} \cdot x^k \cdot y^{n+1-k}\end{aligned}$$

Die Aussage folgt damit aus dem Prinzip der vollständigen Induktion. □

Bemerkung B3.12 (Pascalsches Dreieck)
Man ordnet die Binomialkoeffizienten gerne in der folgenden Form an, die als Pascalsches Dreieck bekannt ist:

$$\begin{array}{ccccccccc}
 & & & & \binom{0}{0} & & & & \\
 & & & \binom{1}{0} & & \binom{1}{1} & & & \\
 & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & \\
 & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & \\
\binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4}
\end{array}$$

Berechnet man die Werte der Binomialkoeffizienten, erhält man die folgende Gestalt:

$$\begin{array}{lccccccccc}
\text{0. Zeile:} & & & & & 1 & & & & \\
\text{1. Zeile:} & & & & 1 & & 1 & & & \\
\text{2. Zeile:} & & & 1 & & 2 & & 1 & & \\
\text{3. Zeile:} & & 1 & & 3 & & 3 & & 1 & \\
\text{4. Zeile:} & 1 & & 4 & & 6 & & 4 & & 1
\end{array}$$

Aufgrund von Proposition B3.10 kann man die Einträge der $n+1$-ten Zeile aus den Einträgen der n-ten Zeile berechnen. Graphisch im Pascalschen Dreieck nimmt die Proposition folgende Gestalt an:

$$\begin{tikzcd}
\binom{n}{k-1} \arrow[dr] & + & \binom{n}{k} \arrow[dl] \\
 & \binom{n+1}{k} &
\end{tikzcd}$$

D.h. die Summe zweier benachbarter Einträge der n-ten Zeile liefert den mittig unter ihnen stehenden Eintrag der $n+1$-ten Zeile.

Aufgrund des binomischen Lehrsatzes sind die Einträge der n-ten Zeile des Pascalschen Dreiecks genau die Koeffizienten, die wir erhalten, wenn wir $(x+y)^n$ ausschreiben. Z.B.

$$(x+y)^3 = 1\cdot x^3 + 3\cdot x^2y + 3\cdot xy^2 + 1\cdot y^3.$$

Aufgaben

Aufgabe B3.13
Es sei K ein Körper und $x \in K$. Zeige, $x^2 = 1$ genau dann, wenn $x \in \{1, -1\}$.

Aufgabe B3.14

a. Auf der Menge $G := \mathbb{R} \times \mathbb{R}$ definieren wir eine zweistellige Operation

$$+ : G \times G \longrightarrow G : \big((x,y),(u,v)\big) \mapsto (x+u, y+v).$$

Zeige, $(G,+)$ ist eine abelsche Gruppe mit neutralem Element $(0,0)$.

b. Auf der Menge $H := (\mathbb{R}\times\mathbb{R})\setminus\{(0,0)\}$ definieren wir eine zweistellige Operation

$$\cdot : H \times H \longrightarrow H : \big((x,y),(u,v)\big) \mapsto (xu - yv, xv + yu).$$

Zeige, $(H,\cdot)$ ist eine abelsche Gruppe mit neutralem Element $(1,0)$.

c. Zeige, daß $(\mathbb{R}\times\mathbb{R},+,\cdot)$ ein Körper ist, wenn die Operationen "+" und "·" wie in a. und b. definiert sind.

Aufgabe B3.15
Zeige durch vollständige Induktion, daß

$$\sum_{k=0}^{n}(k+1)\cdot\binom{n}{k} = 2^{n-1}\cdot(n+2)$$

für $n \in \mathbb{N}$ gilt.

Aufgabe B3.16 (Die projektive Gerade als Gruppe)
Wir haben in Aufgabe A6.15 die Projektive Gerade $\mathbb{P}^1_{\mathbb{R}}$ als Menge von Äquivalenzklassen auf $\mathbb{R}^2\setminus\{(0,0)\}$ eingeführt.

Zeige, daß die zweistellige Operation

$$(v_1 : v_2)\cdot(w_1 : w_2) := (v_1\cdot w_1 - v_2\cdot w_2 : v_1\cdot w_2 + v_2\cdot w_1).$$

wohldefiniert ist, d.h. nicht von der Wahl der Repräsentanten für die Äquivalenzklasse abhängt, und daß $\mathbb{P}^1_{\mathbb{R}}$ mit dieser Operation eine Gruppe ist.

§ B4 Ordnungsrelationen

A) Ordnungsrelationen

Definition B4.1 (Ordnungsrelation)
Es sei M eine Menge. Eine *Ordnungsrelation* auf M, auch *Halbordnung* oder *partielle Ordnung* genannt, ist eine Relation $R \subseteq M \times M$, so daß für alle $x, y, z \in M$ gilt:

O1: $(x, x) \in R$, ("Reflexivität")
O2: $(x, y), (y, x) \in R \implies x = y$, ("Antisymmetrie")
O3: $(x, y), (y, z) \in R \implies (x, z) \in R$. ("Transitivität")

Beispiel B4.2
Es sei $M = \mathbb{N}$.

a. Die übliche Größerrelation
$$R = \{(x, y) \in \mathbb{N} \times \mathbb{N} \mid x \leq y\}$$
ist eine Ordnungsrelation auf $\mathbb{N}$.

b. Die Relation
$$R = \{(x, y) \in \mathbb{N} \times \mathbb{N} \mid x \text{ teilt } y\}$$
ist eine weitere Ordnungsrelation auf $\mathbb{N}$ (siehe Aufgabe B4.21).

Notation B4.3 (Schreibweise $\leq$ für Ordnungsrelationen)
Es sei M eine Menge und R eine Ordnungsrelation auf M. Wir definieren für $x, y \in M$
$$x \leq y \;:\Leftrightarrow\; (x, y) \in R,$$
und sprechen in aller Regel von der Ordnungsrelation "$\leq$" statt R, sofern keine Mißverständnisse zu befürchten sind. Ferner sprechen wir von der *partiell* oder *teilgeordneten Menge* $(M, \leq)$.

Mit dieser Schreibweise lassen sich die drei Axiome in Definition B4.1 wie folgt formulieren. Für $x, y, z \in M$ soll gelten:

O1: $x \leq x$, ("Reflexivität")
O2: $x \leq y \wedge y \leq x \implies x = y$, ("Antisymmetrie")
O3: $x \leq y \wedge y \leq z \implies x \leq z$. ("Transitivität")

Gilt für $x, y \in M$, daß $x \leq y$ und $x \neq y$, so schreiben wir auch $x < y$.

Beispiel B4.4
Ist M eine Menge, so ist die Potenzmenge $\mathcal{P}(M)$ von M durch
$$A \leq B \;:\Longleftrightarrow\; A \subseteq B, \text{ für } A, B \in \mathcal{P}(M),$$
partiell geordnet, aber im allgemeinen sind zwei Elemente von $\mathcal{P}(M)$ nicht unbedingt vergleichbar bezüglich dieser Ordnungsrelation. Z. B. sind im Fall $M = \mathbb{N}$ die Elemente $\{2\}$ und $\{3\}$ in $\mathcal{P}(\mathbb{N})$ nicht vergleichbar.

Allgemeiner gilt, ist N eine Menge, deren Elemente wieder Mengen sind, so wird N mit der analogen Definition von "$\leq$" eine partiell geordnete Menge.

Definition B4.5 (Total- und Wohlordnungen)
Es sei M ein Menge.

a. Eine Ordnungsrelation "$\leq$" auf M heißt *Totalordnung* oder *lineare Ordnung*, falls je zwei Elemente aus M vergleichbar sind, d. h. für je zwei Elemente $x, y \in M$ gilt $x \leq y$ oder $y \leq x$.
b. Ist "$\leq$" eine Ordnungsrelation auf M, $A \subseteq M$ und $x \in A$, so heißt x *minimal* (bzw. *maximal*) in A, falls für alle $y \in A$ mit $y \leq x$ (bzw. $x \leq y$) gilt $x = y$.
c. Eine Totalordnung heißt *Wohlordnung*, falls jede nicht-leere Teilmenge von M ein minimales Element besitzt.

Bemerkung B4.6 (Minimum und Maximum)
Das Minimum bzw. Maximum einer Menge M bezüglich einer Totalordnung ist offenbar eindeutig bestimmt, sofern es existiert. Wir bezeichnen es mit $\min(M)$ bzw. mit $\max(M)$.

Beispiel B4.7

a. Die reellen Zahlen $(\mathbb{R}, \leq)$ mit der üblichen Kleiner-Gleich-Relation $\leq$ sind total geordnet, aber nicht wohlgeordnet.
b. Gleiches trifft auf $(\mathbb{Z}, \leq)$ mit der üblichen Kleiner-Gleich-Relation
$$\ldots - 2 < -1 < 0 < 1 < 2 < \ldots$$
zu. Allerdings definiert die "unübliche" Anordnung
$$0 < -1 < 1 < -2 < 2 < -3 < 3 < \ldots$$
in der Tat eine Wohlordnung auf $\mathbb{Z}$.

Bemerkung B4.8 (Archimedisches Prinzip)
Die natürlichen Zahlen sind bezüglich der üblichen Ordnungsrelation "$\leq$" wohlgeordnet, d.h.:

Jede nicht-leere Menge natürlicher Zahlen enthält eine kleinste Zahl.

Diese wohlbekannte Eigenschaft der natürlichen Zahlen nennen wir auch das *archimedische Prinzip.*

Definition B4.9 (Charakteristik eine Körpers)
Es sei K ein Körper. Gibt es eine positive ganze Zahl $n \in \mathbb{Z}_{>0}$ mit $n \cdot 1_K = 0_K$, so definieren wir
$$\operatorname{char}(K) := \min\{m > 0 \mid m \cdot 1_K = 0_K\} \in \mathbb{N},$$
sonst setzen wir $\operatorname{char}(K) := 0$. Die Zahl $\operatorname{char}(K)$ heißt die *Charakteristik* von K.

Proposition B4.10 (Die Charakteristik eines Körpers ist eine Primzahl oder Null.)
Ist K *ein Körper mit* $\mathrm{char}(K) \neq 0$*, so ist* $\mathrm{char}(K)$ *eine Primzahl.*

Beweis: Angenommen, $n := \mathrm{char}(K)$ sei keine Primzahl. Dann gibt es zwei Zahlen $1 < a, b < n$ mit $n = a \cdot b$. Setzen wir $x = a \cdot 1_K$ und $y = b \cdot 1_K$, so gilt

$$x \cdot y = (a \cdot 1_K) \cdot (b \cdot 1_K) = (a \cdot b) \cdot 1_K = n \cdot 1_K = 0.$$

Aus Lemma B3.4 folgt dann aber $a \cdot 1_K = x = 0$ oder $b \cdot 1_K = y = 0$, im Widerspruch zur Minimalität von $n = \mathrm{char}(K)$. Also muß n eine Primzahl sein. □

Beispiel B4.11
Ist p eine Primzahl, so hat der Körper $\mathbb{F}_p$ aus Beispiel B3.2 die Charakterisitik $\mathrm{char}(\mathbb{F}_p) = p$. Die Körper $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ haben Charakteristik null.

Definition B4.12 (Supremum und Infimum)
Es sei "$\leq$" eine Totalordnung auf einer Menge M und $\emptyset \neq A \subseteq M$ eine nicht-leere Teilmenge von M.

a. Wir nennen $s \in M$ eine *obere Schranke* von A, falls $s \geq x$ für alle $x \in A$.
b. Wir nennen A *nach oben beschränkt*, falls A eine obere Schranke besitzt.
c. Wir nennen $s \in M$ das *Supremum* von A, falls s das Minimum der Menge der oberen Schranken von A ist. Dieses Minimum ist eindeutig bestimmt, wenn es existiert, und wir bezeichnen es dann mit $\sup(A)$.
d. Wir nennen $s \in M$ eine *untere Schranke* von A, falls $s \leq x$ für alle $x \in A$.
e. Wir nennen A *nach unten beschränkt*, falls A eine untere Schranke besitzt.
f. Wir nennen $s \in M$ das *Infimum* von A, falls s das Maximum der Menge aller unteren Schranken von A ist. Dieses Maximum ist eindeutig bestimmt, wenn es existiert, und wir bezeichnen es dann mit $\inf(A)$.
g. Wir nennen A *beschränkt*, wenn A nach oben und nach unten beschränkt ist.

Beispiel B4.13

a. Besitzt eine Teilmenge A einer totalgeordneten Menge M ein Maximum, so ist dieses offenbar auch das Supremum von A. Analog ist das Minimum einer Menge A auch ihr Infimum.
b. Betrachten wir die reellen Zahlen mit ihrer üblichen Ordnung und die Menge $A = \{x \in \mathbb{R} \mid 0 < x \leq 1\}$, so ist $1 = \sup(A) = \max(A)$ das Supremum von A, das zugleich ein Maximum ist, und $0 = \inf(A)$ ist ein Infimum von A, das kein Minimum ist.
c. Betrachten wir die rationalen Zahlen mit ihrer üblichen Ordnungsrelation, so ist
$$\{x \in \mathbb{Q} \mid x > 0 \text{ und } x^2 \leq 2\}$$
nach oben beschränkt, besitzt aber kein Supremum in $\mathbb{Q}$ (siehe Satz B5.10).

Bemerkung B4.14 (Supremumsaxiom)
Die reellen Zahlen sind bezüglich ihrer üblichen Ordnungsrelation nicht wohlgeordnet, d.h. nicht jede nicht-leere Teilmenge besitzt ein kleinstes Element. Selbst, wenn wir voraussetzen, daß die Teilmenge nach unten beschränkt ist, muß sie kein kleinstes Element besitzen, d.h. kein Minimum enthalten, wie wir in Beispiel B4.13 gesehen haben. Es gilt aber, daß zu jeder nicht-leeren, nach unten beschränkten Teilmenge von $\mathbb{R}$ ein Infimum in $\mathbb{R}$ existiert. Äquivalent dazu ist die duale Aussage für das Supremum:

> Jede nicht-leere, nach oben beschränkte Teilmenge von $\mathbb{R}$ besitzt ein Supremum in $\mathbb{R}$.

Diese Eigenschaft ist als *Supremumsaxiom* der reellen Zahlen bekannt. Auch wenn sich die Korrektheit der Aussage nicht unmittelbar aus unserer Alltagserfahrung mit den reellen Zahlen als Dezimalzahlen erschließt, wollen wir sie ohne weiteren Beweis als gegeben voraussetzen.

B) Angeordnete Körper

Definition B4.15 (Angeordnete Körper)
Es sei K ein Körper und "$\leq$" eine Totalordnung auf K. Wir nennen das Quadrupel $(K, +, \cdot, \leq)$ einen *angeordneten Körper*, wenn die Totalordnung mit der Addition und der Multiplikation verträglich ist, d.h. wenn für alle $x, y, z \in K$

$$x < y \quad \Longrightarrow \quad x + z < y + z$$

und

$$x < y, \;\; 0 < z \quad \Longrightarrow \quad x \cdot z < y \cdot z$$

gilt. Ist $x \in K$ und $x > 0$, so nennen wir x *positiv*, ist $x < 0$, so nennen wir x *negativ*.

Beispiel B4.16

a. Die rationalen Zahlen $\mathbb{Q}$ und die reellen Zahlen $\mathbb{R}$ mit der üblichen Ordnungsrelation sind Beispiele für angeordnete Körper. $\mathbb{Q}$ erfüllt das Supremumsaxiom nicht (siehe Beispiel B4.13), $\mathbb{R}$ erfüllt es.
b. Es gibt keine Totalordnung auf $\mathbb{F}_2$, durch die $\mathbb{F}_2$ ein angeordneter Körper würde. Denn würde es eine solche Totalordnung "$\leq$" geben, so wäre entweder $0 < 1$, was zum Widerspruch $1 = 0 + 1 < 1 + 1 = 0$ führt, oder es wäre $1 < 0$, was zum Widerspruch $0 = 1 + 1 < 0 + 1 = 1$ führt.

Lemma B4.17 (Rechenregeln in angeordneten Körpern)
Es sei $(K, +, \cdot, \leq)$ *ein angeordneter Körper und* $x, y, u, v \in K$.

a. $x > 0 \quad \Longleftrightarrow \quad -x < 0.$
b. *Ist* $x \neq 0$, *so ist* $x^2 > 0$.
c. $1 > 0$.

d. *Ist* $0 < x < y$, *so ist* $0 < \frac{1}{y} < \frac{1}{x}$.

e. *Ist* $x < y$ *und* $u < v$, *so ist* $x + u < y + v$.

f. *Ist* $0 < x$ *und* $n \in \mathbb{N}$, *so ist* $0 < x^n$.

g. *Ist* $0 \leq x, y$ *und* $n \in \mathbb{N}$ *mit* $n \geq 1$, *so gilt*

$$x < y \iff x^n < y^n.$$

Beweis:

a. Aus $0 < x$ folgt durch Addition von $-x$

$$-x = 0 + (-x) < x + (-x) = 0.$$

Umgekehrt folgt aus $-x < 0$ durch Addition von x

$$0 = -x + x < 0 + x = x.$$

b. Ist $x > 0$, so folgt unmittelbar

$$0 = 0 \cdot x < x \cdot x = x^2.$$

Ist $x < 0$, so ist $0 < -x$ und es gilt

$$0 = 0 \cdot (-x) < (-x) \cdot (-x) = x \cdot x = x^2.$$

c. $1 = 1^2 > 0$.

d. Nach Voraussetzung ist $y > 0$. Nehmen wir an, $\frac{1}{y} < 0$, so folgt

$$1 = \frac{1}{y} \cdot y < 0 \cdot y = 0$$

im Widerspruch zu Teil c., also ist $0 < \frac{1}{y}$. Entsprechend gilt $0 < \frac{1}{x}$, so daß auch

$$0 = 0 \cdot \frac{1}{y} < \frac{1}{x} \cdot \frac{1}{y} = \frac{1}{xy}$$

und somit wegen $x < y$ auch

$$\frac{1}{y} = x \cdot \frac{1}{xy} < y \cdot \frac{1}{xy} = \frac{1}{x}.$$

e. Wir wenden die Verträglichkeit der Totalordnung mit der Addition mehrfach an:

$$x + u < y + u < y + v.$$

f./g. Den Beweis überlassen wir dem Leser als Übungsaufgabe.

□

Proposition B4.18 (Charakerisierung des Supremums und Infimums)
Ist $(K, +, \cdot, \leq)$ *ein angeordneter Körper,* $A \subseteq K$ *und* $s \in K$, *dann gelten*

$$s = \sup(A) \iff \begin{array}{l} 1)\ \forall\, x \in A\ :\ x \leq s \text{ und} \\ 2)\ \forall\, 0 < \varepsilon \in K\ :\ \exists\, x \in A\ :\ s - \varepsilon < x \end{array}$$

sowie

$$s = \inf(A) \iff \begin{array}{l} 1) \ \ \forall\, x \in A \ : \ x \geq s \ \textit{und} \\ 2) \ \ \forall\, 0 < \varepsilon \in K \ : \ \exists\, x \in A \ : \ s + \varepsilon > x. \end{array}$$

Beweis: Ist $s = \sup(A)$, so ist s ein obere Schranke von A und somit gilt Bedingung 1). Sei also $0 < \varepsilon \in K$, so ist $s - \varepsilon < s$ und mithin ist $s - \varepsilon$ keine obere Schranke von A. Also gibt es ein $x \in A$ mit $x > s - \varepsilon$ und Bedingung 2) ist erfüllt.

Nehmen wir nun umgekehrt an, daß die Bedingungen 1) und 2) gelten. Wegen 1) ist s dann eine obere Schranke von A, und wir müssen nur noch zeigen, daß es keine kleinere obere Schranke geben kann. Dazu betrachten wir eine beliebige kleinere Zahl $t \in K$ mit $t < s$. Für $\varepsilon := s - t \in K$ gilt $\varepsilon > 0$ und wegen 2) gibt es dann ein $x \in A$ mit $x > s - \varepsilon = t$. Also ist t keine obere Schranke von A.

Die Aussage für das Infimum zeigt man analog. □

Das folgende Lemma ist interessant bei der Definition des Riemann-Integrals einer Funktion

Lemma B4.19
Seien $A, B \subseteq \mathbb{R}$ *zwei nicht-leere Teilmengen von* $\mathbb{R}$ *mit* $a \leq b$ *für alle* $a \in A$*,* $b \in B$*. Dann gilt*

$$\sup(A) \leq \inf(B).$$

Beweis: Aus der Voraussetzung folgt unmittelbar, daß A nach oben und B nach unten beschränkt ist, so daß $\sup(A) \in \mathbb{R}$ und $\inf(B) \in \mathbb{R}$ existieren.

Angenommen, $\sup(A) > \inf(B)$, so ist $\varepsilon := \frac{\sup(A)-\inf(B)}{2} > 0$. Somit ist $\sup(A) - \varepsilon$ keine obere Schranke von A und $\inf(B) + \varepsilon$ keine untere Schranke von B. Es gibt also ein $a \in A$ und ein $b \in B$ mit

$$a > \sup(A) - \varepsilon = \frac{\sup(A) + \inf(B)}{2} = \inf(B) + \varepsilon > b,$$

was im Widerspruch zur Voraussetzung steht. □

Aufgaben

Aufgabe B4.20
Ist M eine endliche Menge, so gilt

$$|M| = \min\{n \in \mathbb{N} \mid \exists\, f : M \longrightarrow \{1, \ldots, n\} \text{ injektiv}\},$$

und jede injektive Abbildung $f : M \longrightarrow \{1, \ldots, |M|\}$ ist bijektiv.

Aufgabe B4.21
Zeige, daß durch

$$R := \{(x, y) \in \mathbb{N} \times \mathbb{N} \mid x \text{ teilt } y\}$$

eine Ordnungsrelation auf $\mathbb{N}$ definiert wird. Ist R eine Totalordnung?

Aufgabe B4.22

Definiere auf $M = \mathbb{N} \times \mathbb{N}$ eine Relation durch

$$(m,n) \leq (k,l) \iff \begin{array}{l} 1.\ \max\{m,n\} < \max\{k,l\} \text{ oder} \\ 2.\ (\max\{m,n\} = \max\{k,l\} \text{ und } m < k) \text{ oder} \\ 3.\ (\max\{m,n\} = \max\{k,l\} \text{ und } m = k \text{ und } n > l) \text{ oder} \\ 4.\ (m,n) = (k,l). \end{array}$$

Zeige, daß "$\leq$" eine Totalordnung auf M definiert. Stelle graphisch in der Zahlenebene $\mathbb{R}^2$ dar, wie die Elemente (m,n) in M mit $\max\{m,n\} \leq 4$ angeordnet sind.

Aufgabe B4.23

Sei K ein angeordneter Körper und $A, B \subseteq K$ Teilmengen, so daß $\sup(A)$ und $\sup(B)$ existieren. Wir setzen $A+B := \{a+b \mid a \in A, b \in B\}$. Beweise, daß auch $\sup(A+B)$ existiert und $\sup(A+B) = \sup(A) + \sup(B)$ gilt.

Aufgabe B4.24

Bestimme Supremum, Infimum, Maximum und Minimum (sofern sie existieren) der Mengen:

$$A = \left\{ \frac{m+n}{m \cdot n} \;\middle|\; m, n \in \mathbb{N}_{>0} \right\} \subseteq \mathbb{R}$$

und

$$B = \left\{ n + \frac{(-1)^n}{n} \;\middle|\; n \in \mathbb{N}_{>0} \right\} \subseteq \mathbb{R}.$$

§ B5 Eigenschaften der reellen Zahlen $\mathbb{R}$

Theorem B5.1 (Charakterisierung der reellen Zahlen)
Der Körper $\mathbb{R}$ der reellen Zahlen mit der üblichen Ordnungsrelation ist der einzige angeordnete Körper, in dem jede nicht-leere, nach oben beschränkte Menge ein Supremum besitzt.

Bemerkung B5.2
Die Aussage in Theorem B5.1 besagt zweierlei. Zum einen wird festgestellt, daß $\mathbb{R}$ ein angeordneter Körper ist und dem Supremumsaxiom genügt. Zum anderen wird festgestellt, daß dies für keinen *anderen* angeordneten Körper gilt. Das soll heißen, wenn es einen anderen angeordneten Körper $(K, +, \cdot, \leq)$ mit diesen Eigenschaften gibt, dann gibt es eine *bijektive* Abbildung

$$f : \mathbb{R} \longrightarrow K,$$

so daß $f(x + y) = f(x) + f(y)$, $f(x \cdot y) = f(x) \cdot f(y)$ und

$$x \leq y \quad \Longleftrightarrow \quad f(x) \leq f(y)$$

für alle $x, y \in \mathbb{R}$ gilt. In dem Fall kann man die beiden Körper nicht mehr unterscheiden. Man sagt deshalb auch, daß die reellen Zahlen durch die Eigenschaften in Theorem B5.1 charakterisiert sind, und man könnte die reellen Zahlen axiomatisch durch Angabe der Eigenschaften einführen.

Wir wollen Theorem B5.1 in dieser Vorlesung *nicht* beweisen. Stattdessen werden wir von den reellen Zahlen von nun an nur noch die im Satz angegebenen Eigenschaften wirklich verwenden. Wenn wir uns also $\mathbb{R}$ als einen beliebigen angeordneten Körper mit Supremumsaxiom denken, dann wird alles, was wir von nun an beweisen, dort genauso gelten. Wir müßten die reellen Zahlen also noch gar nicht kennen, um die weitere Theorie betreiben zu können. Die wenigen oben gegebenen Axiome reichen uns aus. Insofern befinden wir uns von jetzt an auf wesentlich sichererem Grund und müssen nicht mehr immer wieder Bezug auf unser Vorwissen zu den Zahlsystemen nehmen.

Satz B5.3 ($\mathbb{R}$ ist archimedisch angeordnet.)
Für $x, y \in \mathbb{R}$ mit $0 < x < y$ gibt es eine natürliche Zahl $n \in \mathbb{N}$, so daß $y < n \cdot x$.

Beweis: Wir betrachten die nicht-leere Teilmenge

$$A := \{n \cdot x \mid n \in \mathbb{N}\} \subsetneqq \mathbb{R}$$

der reellen Zahlen und müssen zeigen, daß y keine obere Schranke dieser Menge ist.

Nehmen wir an, dies wäre doch der Fall, dann ist A nach oben beschränkt und somit existiert das Supremum

$$s := \sup(A).$$

Da $x > 0$ ist, ist $s - x < s$ und somit ist $s - x$ keine obere Schranke von A, d.h. es gibt eine natürliche Zahl $n \in \mathbb{N}$ mit

$$s - x < n \cdot x.$$

Dann ist aber auch

$$s = (s - x) + x < n \cdot x + x = (n + 1) \cdot x,$$

im Widerspruch dazu, daß s eine obere Schranke von A ist.

Damit haben wir gezeigt, daß A keine obere Schranke besitzt und insbesondere, daß y keine solche ist, d.h. es gibt eine natürliche Zahl $n \in \mathbb{N}$ mit $y < n \cdot x$. □

Korollar B5.4 (Konsequenzen der archimedischen Anordnung)

a. *Für alle $x \in \mathbb{R}$ gibt es eine ganze Zahl n, so daß $n \leq x < n + 1$.*

b. *Für alle $\varepsilon \in \mathbb{R}$ mit $\varepsilon > 0$ gibt es eine natürliche Zahl n, so daß $0 < \frac{1}{n} < \varepsilon$.*

Beweis:

a. Ist $0 \leq x < 1$, so ist $n = 0$. Ist $1 \leq x$, so gibt es nach Satz B5.3 eine Zahl $m \in \mathbb{N}$ mit $x < m \cdot 1 = m$. Nach dem Archimedischen Prinzip B4.8 besitzt dann die nicht-leere Menge

$$M := \{k \in \mathbb{N} \mid x < k\}$$

ein Minimum $m_0 = \min(M)$, und für $n := m_0 - 1 < m_0$ gilt mithin

$$n \leq x < m_0 = n + 1.$$

Ist $x < 0$, so ist $-x > 0$ und wir haben schon gezeigt, daß es eine natürliche Zahl $m \in \mathbb{N}$ mit $m \leq -x < m + 1$ gibt. Dann ist aber

$$-m - 1 < x \leq -m.$$

Falls $x = -m$, so setzen wir $n := -m$, und sonst setzen wir $n := -m - 1$.

b. Wegen $\varepsilon > 0$ ist nach Lemma B4.17 auch $\frac{1}{\varepsilon} > 0$, und nach a. gibt es dann eine natürliche Zahl $n \in \mathbb{N}$ so, daß

$$0 < \frac{1}{\varepsilon} < n.$$

Mit Lemma B4.17 folgt dann

$$0 < \frac{1}{n} < \frac{1}{\frac{1}{\varepsilon}} = \varepsilon.$$

□

Definition B5.5 (Intervalle)
Es seien $a, b \in \mathbb{R}$. Wir nennen eine Menge der Form

$$[a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}$$

ein *abgeschlossenes Intervall*, eine Menge der Form

$$(a,b) := \{x \in \mathbb{R} \mid a < x < b\}$$

ein *offenes Intervall* und Mengen der Form

$$[a,b) := \{x \in \mathbb{R} \mid a \leq x < b\}$$

bzw.

$$(a,b] := \{x \in \mathbb{R} \mid a < x \leq b\}$$

halboffene Intervalle. Mengen der Form

$$[a,\infty) := \{x \in \mathbb{R} \mid a \leq x\},$$

$$(a,\infty) := \{x \in \mathbb{R} \mid a < x\},$$

$$(-\infty,a] := \{x \in \mathbb{R} \mid x \leq a\},$$

$$(-\infty,a) := \{x \in \mathbb{R} \mid x < a\}$$

$$(-\infty,\infty) := \mathbb{R}$$

heißen *uneigentliche Intervalle*.

Satz B5.6 ($\mathbb{Q}$ liegt dicht in $\mathbb{R}$.)
Sind $a, b \in \mathbb{R}$ *mit* $a < b$*, so gibt es eine rationale Zahl im Intervall* (a, b).

Beweis: Wegen $b - a > 0$ gibt es nach Korollar B5.4 eine natürliche Zahl $n \in \mathbb{N}$ mit

$$0 < \frac{1}{n} < b - a. \tag{63}$$

Zudem gibt es nach Korollar B5.4 eine ganze Zahl $m \in \mathbb{Z}$ mit

$$m \leq n \cdot a < m + 1. \tag{64}$$

Damit gilt dann

$$a \overset{(64)}{<} \frac{m+1}{n} = \frac{m}{n} + \frac{1}{n} \overset{(64)}{\leq} a + \frac{1}{n} \overset{(63)}{<} b$$

und $\frac{m+1}{n} \in \mathbb{Q}$ ist eine rationale Zahl. □

Satz B5.7 (Bernoullische Ungleichung)
Es sei $x \in \mathbb{R}$ *mit* $x \geq -1$ *und* $n \in \mathbb{N}$*, dann gilt*

$$(1 + x)^n \geq 1 + n \cdot x.$$

Beweis: Wir führen den Beweis durch Induktion nach n.

Induktionsanfang: $n = 0$: $(1 + x)^0 = 1 = 1 + 0 \cdot x$.

Induktionsschluß: $n \mapsto n + 1$: Nach Lemma B4.17 b. ist $x^2 \geq 0$ und nach Voraussetzung gilt zudem $1 + x \geq 0$. Damit erhalten wir dann:

$$(1 + x)^{n+1} = (1 + x)^n \cdot (1 + x) \overset{\text{Ind}}{\geq}$$

$$(1 + n \cdot x) \cdot (1 + x) = 1 + (n + 1) \cdot x + n \cdot x^2 \overset{\text{B4.17b.}}{\geq} 1 + (n + 1) \cdot x.$$

Die Aussage ist damit also mittels Induktion gezeigt. □

Satz B5.8 (Existenz von n-ten Wurzeln in $\mathbb{R}$)
Zu jeder reellen Zahl $x \in \mathbb{R}$ mit $x \geq 0$ und jeder natürlichen Zahl $n \in \mathbb{N}$ mit $n \geq 2$ gibt es genau eine reelle Zahl $a \in \mathbb{R}$ mit $a \geq 0$ und $a^n = x$.
Wir nennen diese Zahl die n-te Wurzel *aus x und bezeichnen sie mit $\sqrt[n]{x}$ oder $x^{\frac{1}{n}}$.*

Beweis: Wir wollen uns zunächst der Eindeutigkeit der Lösung zuwenden, sofern sie existiert. Nehmen wir also an, es würde zwei verschiedene nicht-negative reelle Zahlen $a, b \in \mathbb{R}$ mit $a^n = b^n = x$ geben. Dann ist eine der beiden echt kleiner als die andere und wir können ohne Einschränkung annehmen, daß dies a ist, d.h. $0 \leq a < b$. Aus Lemma B4.17 g. folgt dann $x = a^n < b^n = x$, was ein offensichtlicher Widerspruch ist. Mithin haben wir gezeigt, daß es höchstens eine nicht-negative Zahl $a \in \mathbb{R}$ mit $a^n = x$ geben kann.

Es bleibt noch zu zeigen, daß es auch wirklich eine solche nicht-negative Zahl a gibt. Ist $x = 0$, so ist $a = 0$ eine Lösung für $a^n = 0$. Wir können im weiteren Verlauf des Beweises also voraussetzen, daß $x > 0$.

Wir betrachten dann die Teilmenge

$$A := \{y \in \mathbb{R} \mid y \geq 0, y^n \leq x\}$$

der reellen Zahlen, und wir behaupten, daß $1+x$ eine obere Schranke für A ist. Dazu betrachten wir eine reelle Zahl $y \in \mathbb{R}$ mit $y \geq 1 + x > 0$. Aus der Bernoullischen Ungleichung folgt dann

$$y^n \overset{\text{B4.17g.}}{\geq} (1+x)^n \overset{\text{B5.7}}{\geq} 1 + n \cdot x > x,$$

und somit ist $y \notin A$. Also ist A nach oben beschränkt durch $x + 1$. Wegen $0 \in A$ ist A zudem nicht-leer und deshalb existiert das Supremum

$$a := \sup(A) \geq 0.$$

Wir wollen nun zeigen, daß $a^n = x$ gilt.

Zeige: $a^n \geq x$: Nehmen wir an, es gelte $a^n < x$.
Idee: *Finde eine reelle Zahl $\varepsilon > 0$, so daß $a + \varepsilon \in A$. – ↯*
Wegen $a \geq 0$ ist

$$c := \sum_{k=1}^{n} \binom{n}{k} \cdot a^{n-k} \geq \binom{n}{n} = 1 > 0$$

und somit auch $\frac{1}{c} > 0$ nach Lemma B4.17. Aus unserer Annahme folgt dann

$$\frac{x - a^n}{c} > 0.$$

Somit ist auch

$$\varepsilon := \min\left\{\frac{x - a^n}{c},\ 1\right\} > 0$$

und es folgt
$$a^n + c \cdot \varepsilon \leq x. \tag{65}$$
Wegen $0 < \varepsilon \leq 1$ ist $\varepsilon^k \leq \varepsilon$ für alle $k \geq 1$, und aus dem Binomischen Lehrsatz B3.11 folgt dann
$$\begin{aligned}(a+\varepsilon)^n =& a^n + \sum_{k=1}^{n} \binom{n}{k} \cdot a^{n-k} \cdot \varepsilon^k \\ \leq& a^n + \sum_{k=1}^{n} \binom{n}{k} \cdot a^{n-k} \cdot \varepsilon \;=\; a^n + c \cdot \varepsilon \overset{(65)}{\leq} x.\end{aligned}$$
Somit ist $a+\varepsilon \in A$ und $a+\varepsilon > a$ im Widerspruch dazu, daß a das Supremum von A ist. Mithin muß $a^n \geq x$ sein.

Zeige: $a^n \leq x$: Nehmen wir an, es gelte $a^n > x$.

Idee: *Finde ein $\varepsilon > 0$ und ein $y \in A$, so daß $y^n > (a-\varepsilon)^n \geq x$.* – ↯

Wegen $a^n > x$ ist $a > 0$ und dann ist auch die Zahl
$$\frac{a \cdot (a^n - x)}{n \cdot a^n} > 0$$
positiv. Wir setzen nun
$$\varepsilon := \min\left\{ \frac{a \cdot (a^n - x)}{n \cdot a^n}\;,\;\; a \right\} > 0.$$
Aus der Definition von ε folgt zum einen
$$-\frac{\varepsilon}{a} \geq -1 \tag{66}$$
und zum anderen unter Anwendung der Bernoullischen Ungleichung
$$x \leq a^n \cdot \left(1 + n \cdot \frac{-\varepsilon}{a}\right) \overset{\text{B5.7}}{\leq} a^n \cdot \left(1 - \frac{\varepsilon}{a}\right)^n = (a-\varepsilon)^n\,; \tag{67}$$
dabei beachten wir die Bernoullische Ungleichung wegen (66) anwenden können.

Da a das Supremum von A ist und $a - \varepsilon < a$ ist, muß es eine Zahl $y \in A$ geben mit
$$y > a - \varepsilon > 0.$$
Dann gilt nach Lemma B4.17 auch
$$y^n > (a-\varepsilon)^n \overset{(67)}{\geq} x,$$
im Widerspruch dazu, daß $y \in A$. Also muß auch $a^n \leq x$ gelten.

Da sowohl $a^n \geq x$, als auch $a^n \leq x$ gilt, folgt aus der Antisymmetrie der Ordnungsrelation, daß $a^n = x$, und wir haben die n-te Wurzel von x gefunden. □

Bemerkung B5.9

In $\mathbb{R}$ besitzt also insbesondere jede nicht-negative Zahl eine Quadratwurzel. Dies gilt in den rationalen Zahlen nicht (siehe Satz B5.10), und man kann die reellen Zahlen als eine Erweiterung des Zahlbereichs der rationalen Zahlen ansehen, die unter anderem deshalb notwendig war. Negative Zahlen besitzen aber auch in $\mathbb{R}$ noch keine Quadratwurzeln, und wir werden im folgenden Kapitel deshalb unseren Zahlbereich noch einmal erweitern zu den sogenannten komplexen Zahlen, die dieses Manko dann beheben.

Satz B5.10 ($\sqrt{2}$ ist irrational.)
Es gibt keine rationale Zahl $a \in \mathbb{Q}$ *mit* $a^2 = 2$.

Beweis: Nehmen wir an, es wäre $a = \frac{p}{q} \in \mathbb{Q}$ eine solche Zahl. Wir können ohne weiteres annehmen, daß der Bruch in gekürzter Form vorliegt. Aus

$$\frac{p^2}{q^2} = a^2 = 2$$

folgt dann

$$p^2 = q^2 \cdot 2.$$

Also ist p^2 eine gerade Zahl, und dann muß notwendigerweise auch p eine gerade Zahl sein. D.h. es gibt ein $b \in \mathbb{Z}$ mit $p = 2 \cdot b$. Also ist

$$4 \cdot b^2 = p^2 = 2 \cdot q^2,$$

und somit

$$2 \cdot b^2 = q^2.$$

Mit dem gleichen Argument sind dann auch q^2 und q gerade Zahlen, und somit ist q von der Form $q = 2 \cdot c$. Aber das widerspricht der Voraussetzung, daß der Bruch $\frac{p}{q}$ in gekürzter Form vorgelegen hat. □

Aufgaben

Aufgabe B5.11
Zeige durch vollständige Induktion, daß

$$\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \ldots + \frac{1}{\sqrt{n-1}} + \frac{1}{\sqrt{n}} > \sqrt{n}$$

für alle $n \geq 2$ gilt.

Aufgabe B5.12
Zeige, für je drei reelle Zahlen $a, b, c \in \mathbb{R}$ gilt

$$\frac{|a-c|}{1+|a-c|} \leq \frac{|a-b|}{1+|a-b|} + \frac{|b-c|}{1+|b-c|}$$

§ B6 Der Körper der komplexen Zahlen

Wir kommen jetzt zum Körper $\mathbb{C}$ der komplexen Zahlen, dem neben $\mathbb{R}$ wichtigsten Körper. Warum reichen eigentlich die reellen Zahlen nicht aus, wozu braucht man die komplexen Zahlen? Ja, man kann sogar fragen, warum wir überhaupt die reellen Zahlen benötigen, wenn wir doch ohnehin nur mit endlichen Dezimalbrüchen, also rationalen Zahlen, rechnen können? Die Antwort auf die zweite Frage ist schnell gegeben. Wir wissen alle, daß etwa ganz natürlich auftretende Größen wie die Länge der Diagonalen eines Quadrates mit Seitenlänge eins, sprich die Zahl $\sqrt{2}$, oder das Verhältnis von Umfang zum Durchmesser eines Kreises, sprich die Kreiszahl π, keine rationalen Zahlen sind. Sie sind aber reelle Zahlen und die reellen Zahlen sind in gewissem Sinne, eine Vervollständigung der rationalen Zahlen. Wir brauchen also die reellen Zahlen, da die rationalen Zahlen Lücken aufweisen. Die komplexen Zahlen werden nun deshalb eingeführt, um einen Mangel, den die reellen Zahlen immer noch haben, zu beheben. Hierbei geht es um das Lösen von Gleichungen, aber nicht mehr linearen, sondern quadratischen. Es ist bekannt, daß das Quadrat einer reellen Zahl stets nicht-negativ ist. Also kann es keine reelle Zahl x geben, die die Gleichung $x^2 = -1$ löst.

Als Lösung genau dieser Gleichung wird nun eine neue Größe eingeführt, die *imaginäre Einheit* i. Definitionsgemäß ist sie diejenige Zahl, für die $i^2 = -1$ gilt. Wenn man nun eine solche Größe i einführt, dann ist damit alleine gar nichts gewonnen. Man will ja mit i auch rechnen können, und zwar will man möglichst alle Rechenregeln von $\mathbb{R}$ übertragen. Man will nicht nur $i^2 = i \cdot i$, sondern auch $i+i$ oder Ausdrücke wie $37+42i$ bilden können. Dabei sollen die so zu konstruierenden *komplexen Zahlen* die reellen Zahlen als Teilmenge enthalten.

Daß es wirklich ein solches Zahlsystem komplexer Zahlen, in unserer Sprache den Körper der komplexen Zahlen, gibt, ist überhaupt nicht klar und wurde historisch erst spät realisiert und auch akzeptiert.[2] Gauß hat die Zahlen geometrisch, als Punkte in der Ebene, eingeführt, weshalb die komplexen Zahlen heute noch *Gaußsche Zahlenebene* heißen. Wir führen die komplexen Zahlen ebenfalls als reelle Zahlenpaare ein, definieren die Addition und die Multiplikation aber algebraisch und werden die Definitionen erst im Anschluß daran geometrisch interpretieren.

Bemerkung B6.1 (Konstruktion der komplexen Zahlen)
Es ist unser erklärtes Ziel, auf der reellen Zahlenebene $\mathbb{R}^2$ mit der Vektoraddtion

$$(x, y) + (u, v) := (x + u, y + v)$$

eine *Multiplikation* zu definieren, so daß einerseits die üblichen Rechenregeln (Assoziativgesetze, Kommutativgesetze und Distributivgesetze) gelten und daß außerdem

[2]Erstmals tauchte $\sqrt{-1}$ wohl um 1540 bei Cardano auf. Wirklich als Zahlsystem wurden die komplexen Zahlen aber erst durch Gauß, 1777-1855, etabliert. Hierzu und zu vielen weiteren interessanten Tatsachen um die komplexen Zahlen vgl. [**Ebb92**] § 3.

der Vektor

$$i := (0, 1)$$

eine Lösung der Gleichung

$$z^2 = -1$$

ist. Um letzteres richtig zu interpretieren, denken wir uns die reelle Zahlengerade $\mathbb{R}$ als Teilmenge von $\mathbb{R}^2$, indem wir sie mit der x-Achse identifizieren, d.h.

$$\mathbb{R} \mathrel{\widehat{=}} \{(a, 0) \mid a \in \mathbb{R}\} = x\text{-Achse}.$$

Die Multiplikation soll also der Bedingung

$$i^2 = (0, 1) \cdot (0, 1) = (-1, 0) \mathrel{\widehat{=}} -1$$

genügen. Außerdem würden wir uns sicher wünschen, daß die Multiplikation eines Vektors mit der reellen Zahl

$$a \mathrel{\widehat{=}} (a, 0)$$

wie die Streckung des Vektors um den Faktor a funktioniert, d.h.

$$(a, 0) \cdot (x, y) \mathrel{\widehat{=}} a \cdot (x, y) = (ax, ay).$$

Wenn eine Multiplikation diese Wunschliste erfüllt, so gilt offenbar:

$$\begin{aligned}
(x, y) \cdot (u, v) &= \big((x, 0) + (0, y)\big) \cdot \big((u, 0) + (0, v)\big)\\
&= \big((x, 0) + (y, 0) \cdot (0, 1)\big) \cdot \big((u, 0) + (v, 0) \cdot (0, 1)\big)\\
&= (x, 0) \cdot (u, 0) + (y, 0) \cdot (0, 1) \cdot (u, 0) + (x, 0) \cdot (v, 0) \cdot (0, 1)\\
&\qquad\qquad + (y, 0) \cdot (0, 1) \cdot (v, 0) \cdot (0, 1)\\
&= (xu, 0) + (yu, 0) \cdot (0, 1) + (xv, 0) \cdot (0, 1) + (yv, 0) \cdot (0, 1) \cdot (0, 1)\\
&= (xu, 0) + (yu, 0) \cdot (0, 1) + (xv, 0) \cdot (0, 1) + (yv, 0) \cdot (-1, 0)\\
&= (xu, 0) + (0, yu) + (0, xv) + (-yv, 0)\\
&= (xu - yv, xv + yu).
\end{aligned}$$

Wir haben für die Definition der Multiplikation also nur *eine einzige* Möglichkeit, und die funktioniert zum Glück auch.

Satz B6.2 (Der Körper der komplexen Zahlen)
Die Menge $\mathbb{C} := \{(x, y) \mid x, y \in \mathbb{R}\}$ *zusammen mit der durch*

$$(x, y) + (u, v) := (x + u, y + v), \quad \textit{für } (x, y), (u, v) \in \mathbb{C},$$

und

$$(x, y) \cdot (u, v) := (xu - yv, xv + yu), \quad \textit{für } (x, y), (u, v) \in \mathbb{C},$$

definierten Addition und Multiplikation ist ein Körper, den wir den Körper der komplexen Zahlen *nennen. .*

Beweis: Dies folgt aus Aufgabe B3.14. □

Bemerkung B6.3

a. Daß $\mathbb{C}$ mit den beiden Operationen ein *Körper* ist, bedeutet, daß die oben erwähnten üblichen Rechenregeln bezüglich der Addition, Subtraktion, Multiplikation und Division gelten, so wie wir sie von den reellen Zahlen her kennen. Man beachte dabei, daß die reelle Zahl $0 \widehat{=} (0,0)$ bei der Addtion nichts tut und die reelle Zahl $1 \widehat{=} (1,0)$ bei der Multiplikation ohne Wirkung ist:

$$(x,y) + (0,0) = (x+0, y+0) = (x,y)$$

und

$$(x,y) \cdot (1,0) = (x \cdot 1 - y \cdot 0, x \cdot 0 + y \cdot 1) = (x,y).$$

Das multiplikative Inverse der Zahl $(0,0) \neq (x,y) \in \mathbb{C}$ ist

$$(x,y)^{-1} = \left(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2}\right).$$

b. Die Abbildung

$$\iota : \mathbb{R} \longrightarrow \mathbb{C} : x \mapsto (x,0)$$

ist mit der Addition und der Multiplikation verträglich und identifiziert den Körper der reellen Zahlen $\mathbb{R}$ mit dem Teilkörper $\mathbb{R} \times \{0\}$ von $\mathbb{C}$. Wir fassen $\mathbb{R}$ in diesem Sinne als Teilmenge von $\mathbb{C}$ auf.

c. Praktischer als das Rechnen mit Paaren von Zahlen ist die folgende Notation für komplexe Zahlen. Wir setzen $x := (x,0)$ für $x \in \mathbb{R}$ und $i := (0,1)$. Dann gilt für $z = (x,y) \in \mathbb{C}$

$$z = (x,y) = (x,0) + (0,y) = (x,0) + (0,1) \cdot (y,0) \widehat{=} x + iy.$$

Diese Schreibweise wollen wir künftig für komplexe Zahlen verwenden. Damit gilt dann:

$$i^2 = (0,1) \cdot (0,1) = -1.$$

Ferner ergibt sich die etwas willkürlich anmutende Definition der Multiplikation ganz "natürlich" aus

$$(x+iy)(u+iv) = \big(xu + i^2yv\big) + i(xv + yu) = (xu - yv) + i(xv + yu).$$

Lemma B6.4 ($\mathbb{C}$ ist nicht angeordnet.)

Es gibt keine Totalordnung "$\leq$" auf $\mathbb{C}$, die $\mathbb{C}$ zu einem angeordneten Körper macht.

Beweis: Angenommen, es gäbe eine Totalordnung "$\leq$", die $\mathbb{C}$ zu einem angeordneten Körper macht. Nach Lemma B4.17 muß dann $0 < i^2 = -1$ gelten, was im Widerspruch zu $0 < 1$ steht. $\square$

Definition B6.5 (Der Betrag und die komplexe Konjugation)

a. Wir definieren die *Betragsfunktion* auf $\mathbb{C}$ durch

$$|\cdot| : \mathbb{C} \longrightarrow \mathbb{R}_{\geq 0} : x + iy \mapsto \sqrt{x^2+y^2}$$

und nennen $|x|$ auch den *Absolutbetrag* von x. Wegen Satz B5.8 ist der Betrag einer komplexen Zahl definiert und ist stets eine nicht-negative reelle Zahl.

Beachte zudem, für $x \in \mathbb{R}$ gilt

$$|x| := \begin{cases} x, & \text{falls } x \geq 0 \\ -x, & \text{falls } x < 0. \end{cases}$$

b. Wir definieren die *komplexe Konjugation* als

$$\bar{\cdot} : \mathbb{C} \longrightarrow \mathbb{C} : z = x + iy \mapsto \overline{z} := x - iy.$$

Für $z \in \mathbb{C}$ heißt $\overline{z}$ die zu z *konjugiert komplexe Zahl*.

c. Wir definieren die Abbildungen *Realteil*

$$\mathrm{Re} : \mathbb{C} \longrightarrow \mathbb{R} : x + iy \mapsto x$$

und *Imaginärteil*

$$\mathrm{Im} : \mathbb{C} \longrightarrow \mathbb{R} : x + iy \mapsto y$$

und nennen $\mathrm{Re}(x + iy) = x$ den *Realteil* von z und $\mathrm{Im}(x + iy) = y$ den *Imaginärteil* von z.

Beispiel B6.6
Wir betrachten die komplexe Zahl

$$z = i - 1 = -1 + i.$$

Dann gilt $\mathrm{Re}(z) = -1$, $\mathrm{Im}(z) = 1$ und

$$\overline{z} = -1 - i = -(1 + i).$$

Für den Betrag von z rechnen wir

$$|z| = \sqrt{\mathrm{Re}(z)^2 + \mathrm{Im}(z)^2} = \sqrt{1 + 1} = \sqrt{2}$$

und damit erhalten wir die Gleichung

$$z \cdot \overline{z} = (-1 + i) \cdot (-1 - i) = 2 = |z|^2.$$

Lemma B6.7 (Einfache Rechenregeln in $\mathbb{C}$)
Es seien $z, w \in \mathbb{C}$.

a. *Der Betrag ist multiplikativ, d.h.*

$$|z| \cdot |w| = |zw|.$$

b. *Der Betrag erfüllt die* Dreiecksungleichung*, d.h.*

$$|z + w| \leq |z| + |w|,$$

und es gilt stets

$$\big||z| - |w|\big| \leq |z - w|.$$

c. $z = 0 \iff |z| = 0$.

d. $z \cdot \overline{z} = |z|^2$.

e. *Wenn $z \neq 0$, dann ist $z^{-1} = \frac{1}{z} = \frac{\overline{z}}{|z|^2}$.*

f. *Die komplexe Konjugation ist additiv, d.h.*

$$\overline{z} + \overline{w} = \overline{z + w}.$$

g. *Die komplexe Konjugation ist multiplikativ, d.h.*

$$\overline{z} \cdot \overline{w} = \overline{z \cdot w}.$$

h. $\overline{\overline{z}} = z$.

i. $\overline{z} = z \iff z \in \mathbb{R}$.

j. $\mathrm{Re}(z) = \frac{z+\overline{z}}{2} \leq |z|$.

k. $\mathrm{Im}(z) = \frac{z-\overline{z}}{2i} \leq |z|$.

l. $|z| = \left|\overline{z}\right| = |-z|$.

Beweis: Die Aussagen in den Teilen c.-l. überlassen wir dem Leser als Übungsaufgabe.

a. Seien $z = x + iy$ und $w = u + iv$ mit $x, y, u, v \in \mathbb{R}$. Dann gilt

$$\begin{aligned}
|zw|^2 =& |(xu - yv) + i \cdot (xv + yu)|^2 = (xu - yv)^2 + (xv + yu)^2 \\
=& x^2u^2 - 2xuyv + y^2v^2 + x^2v^2 + 2xvyu + y^2u^2 \\
=& x^2u^2 + y^2v^2 + x^2v^2 + y^2u^2 = (x^2 + y^2) \cdot (u^2 + v^2) \\
=& |z|^2 \cdot |w|^2 = \big(|z| \cdot |w|\big)^2.
\end{aligned}$$

Aus der Eindeutigkeit der nicht-negativen Quadratwurzel (Satz B5.8) folgt dann

$$|zw| = |z| \cdot |w|.$$

b. Wir wollen nun die Dreiecksungleichung unter Verwendung der übrigen Aussagen zeigen. Es gilt

$$\begin{aligned}
|z + w|^2 &\overset{\text{d.}}{=} (z + w) \cdot \overline{(z + w)} \\
&\overset{\text{f.}}{=} z \cdot \overline{z} + (z \cdot \overline{w} + \overline{z} \cdot w) + w \cdot \overline{w} \\
&\overset{\text{d.,h.}}{=} |z|^2 + \big(z \cdot \overline{w} + \overline{z \cdot \overline{w}}\big) + |w|^2 \\
&\overset{\text{j.}}{=} |z|^2 + 2 \cdot \mathrm{Re}(z \cdot \overline{w}) + |w|^2 \\
&\overset{\text{j.}}{\leq} |z|^2 + 2 \cdot |z \cdot \overline{w}| + |w|^2 \\
&\overset{\text{a.}}{=} |z|^2 + 2 \cdot |z| \cdot |\overline{w}| + |w|^2 \\
&\overset{\text{l.}}{=} |z|^2 + 2 \cdot |z| \cdot |w| + |w|^2 \\
&= \big(|z| + |w|\big)^2.
\end{aligned}$$

Da dies eine Ungleichung von nicht-negativen Zahlen in dem angeordneten Körper $\mathbb{R}$ ist, folgt aus Lemma B4.17, daß

$$|z+w| \le |z| + |w|.$$

Es bleibt, die zweite Aussage in Teil b. zu zeigen. Aus der Dreiecksungleichung erhalten wir

$$|z| = |(z-w)+w| \le |z-w| + |w|,$$

und somit

$$|z| - |w| \le |z-w|.$$

Analog folgt

$$-(|z|-|w|) = |w| - |z| \le |w-z| = |-(w-z)| = |z-w|.$$

Wegen

$$\big||z|-|w|\big| = \begin{cases} |z|-|w|, & \text{falls } |z|-|w| \ge 0, \\ -(|z|-|w|) & \text{falls } |z|-|w| < 0, \end{cases}$$

folgt dann $\big||z|-|w|\big| \le |z-w|$.

$\square$

Beispiel B6.8

a. Gegeben seien $z = 3+2i$ und $w = 5-i$. Dann gelten

$$z \cdot w = (3\cdot 5 - 2\cdot(-1)) + (3\cdot(-1) + 2\cdot 5)\cdot i = 17 + 7i$$

sowie

$$|w| = \sqrt{5^2 + (-1)^2} = \sqrt{26}$$

und

$$\begin{aligned} \frac{z}{w} &= z \cdot \frac{\overline{w}}{|w|^2} = (3+2i)\cdot\left(\frac{5}{26} + \frac{1}{26}\cdot i\right) \\ &= \left(3\cdot\frac{5}{26} - 2\cdot\frac{1}{26}\right) + \left(3\cdot\frac{1}{26} + 2\cdot\frac{5}{26}\right)\cdot i \\ &= \frac{13}{26} + \frac{13}{26}\cdot i = \frac{1}{2} + \frac{1}{2}\cdot i. \end{aligned}$$

b. Für die komplexen Zahlen $z = 3+4i$ und $w = 5-12i$ gilt

$$z + w = (3+5) + (4-12)\cdot i = 8 - 8i$$

und somit

$$\begin{aligned} |z+w| &= \sqrt{8^2+8^2} = \sqrt{2}\cdot 8 < 16 < 18 = 5 + 13 \\ &= \sqrt{25} + \sqrt{169} = \sqrt{3^2+4^2} + \sqrt{5^2+12^2} = |z| + |w|. \end{aligned}$$

Außerdem gilt

$$\frac{z+\overline{z}}{2} = \frac{(3+4i)+(3-4i)}{2} = \frac{6}{2} = 3 = \mathrm{Re}(z).$$

Bemerkung B6.9 (Geometrische Deutung und Polarkoordinaten)
Wir wollen hier einige der bisher eingeführten Operationen auf den komplexen Zahlen und der angeführten Eigenschaften derselben geometrisch interpretieren.

- Die Addition ist einfach die komponentenweise Addition, also die Addition der Vektoren (siehe Abbildung 1).

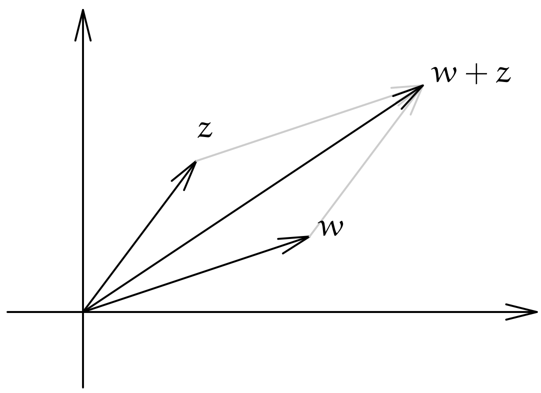


ABBILDUNG 1. Addition in $\mathbb{C}$ als Vektoraddition

- Die komplexe Konjugation ist die Spiegelung an der x-Achse.
- Der Realteil ist die orthogonale Projektion auf die x-Achse und der Imaginärteil die orthogonale Projektion auf die y-Achse.
- Der Betrag $|z| = \sqrt{x^2+y^2}$ einer komplexen Zahl $z = x+iy$ ist die euklidische Länge des Vektors z, d.h. der Abstand von z zum Ursprung. Dies ergibt sich unmittelbar aus dem Satz von Pythagoras (siehe Abbildung 2).

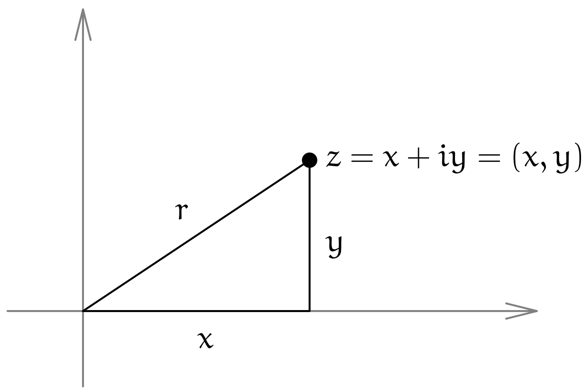


ABBILDUNG 2. Pythagoras: $x^2+y^2=r^2$

- Die Dreiecksungleichung besagt deshalb im wesentlichen, daß in einem Dreieck die Summe der Seitenlängen von zwei Seiten stets eine obere Schranke für die Seitenlänge der dritten Seite ist.
- Damit ist die Menge
$$K := \big\{z \in \mathbb{C} \;\big|\; |z| = 1\big\}$$
die Menge der Punkte in der Ebene, deren Abstand zum Ursprung genau 1 ist, d.h. K ist der Einheitskreis um den Ursprung. Man beachte, daß bei einem

Punkt

$$z = x + iy,$$

der auf dem Einheitskreis liegt, die kartesichen Koordinaten x und y schon vollständig durch den Winkel $\alpha \in [0, 1\pi)$ bestimmt sind, den der Vektor z mit der x-Achse einschließt. Es gilt nämlich (siehe Abbildung 3)

$$x = \cos(\alpha)$$

und

$$y = \sin(\alpha)$$

und somit

$$z = \cos(\alpha) + i \cdot \sin(\alpha).$$

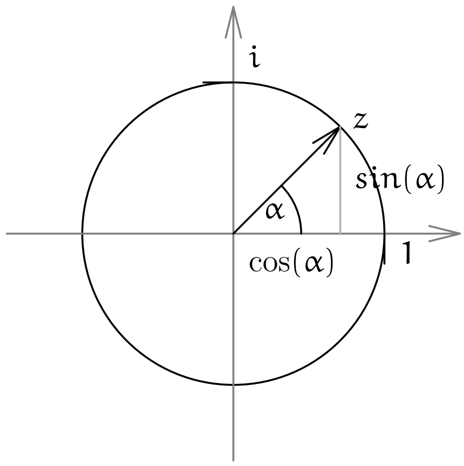


Abbildung 3. Koordinaten eines Punktes $z = \cos(\alpha) + i \cdot \sin(\alpha)$ auf dem Einheitskreis

- Es bleibt, die Multiplikation zweier komplexer Zahlen $0 \neq z, w \in \mathbb{C}$ geometrisch zu deuten. Dazu schreiben wir die Zahl z als

$$z = |z| \cdot \frac{z}{|z|} = r \cdot z'$$

mit $r = |z|$ und $z' = \frac{z}{|z|}$. Man beachte, daß die Zahl z' den Betrag 1 hat, so daß es genau einen Winkel $\alpha \in [0, 2\pi)$ gibt mit

$$z' = (\cos(\alpha), \sin(\alpha)) = \cos(\alpha) + i \cdot \sin(\alpha).$$

Die komplexe Zahl $z \neq 0$ ist also eindeutig durch ihren Betrag und den Winkel α bestimmt. Wir nennen

$$\arg(z) := \alpha = \arctan\left(\frac{\mathrm{Im}(z)}{\mathrm{Re}(z)}\right)$$

das *Argument* von z und das Paar

$$(r, \alpha) = \big(|z|, \arg(z)\big)$$

die *Polarkoordinaten* von z.

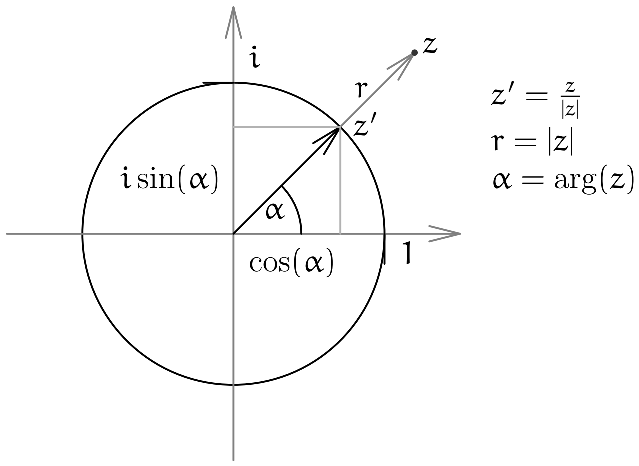


ABBILDUNG 4. Polarkoordinaten von $z = r \cdot (\cos(\alpha) + i \cdot \sin(\alpha))$

Wir erinnern hier an die beiden Additionstheoreme für den Sinus und den Cosinus:

$$\cos(\alpha + \beta) = \cos(\alpha) \cdot \cos(\beta) - \sin(\alpha) \cdot \sin(\beta) \tag{68}$$

und

$$\sin(\alpha + \beta) = \cos(\alpha) \cdot \sin(\beta) + \sin(\alpha) \cdot \cos(\beta). \tag{69}$$

Betrachten wir zunächst die Multiplikation von zwei komplexen Zahlen $z = |z| \cdot \big(\cos(\alpha) + i \cdot \sin(\alpha)\big)$ und $w = |w| ot\big(\cos(\beta) + i \cdot \sin(\beta)\big)$:

$$\begin{aligned} z \cdot w &= |z| \cdot |w| \cdot \big(\cos(\alpha) + i \cdot \sin(\alpha)\big) \cdot \big(\cos(\beta) + i \cdot \sin(\beta)\big) \\ &= |z| \cdot |w| \cdot \big(\cos(\alpha) \cdot \cos(\beta) - \sin(\alpha) \cdot \sin(\beta)\big) + i \cdot \big(\cos(\alpha) \cdot \sin(\beta) + \sin(\alpha) \cdot \cos(\beta)\big) \\ &\overset{(68),(69)}{=} |z| \cdot |w| \cdot \big(\cos(\alpha + \beta) + i \cdot \sin(\alpha + \beta)\big). \end{aligned}$$

Die beiden Zahlen werden also multipliziert, indem man die Argumente addiert und die Beträge multipliziert (siehe Abbildung 5).

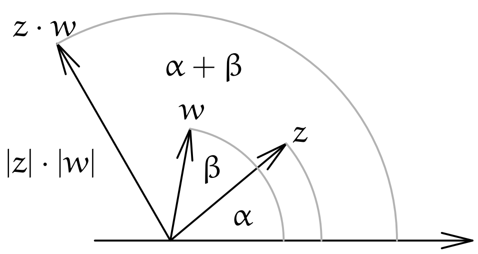


ABBILDUNG 5. Multiplikation zweier komplexer Zahlen

In Polarkoordinaten könnte man dies schreiben als

$$(r, \alpha) \cdot (s, \beta) = (r \cdot s, \alpha + \beta).$$

Beispiel B6.10
Zur Ermittlung von $\alpha = \arg(z)$ für $z = i - 1$ betrachten wir die Zahl

$$\frac{z}{|z|} = -\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}i$$

vom Betrag 1, für die gilt

$$\frac{z}{|z|} = \cos(\alpha) + i\sin(\alpha),$$

d.h. $\cos(\alpha) = -\frac{\sqrt{2}}{2}$ und $\sin(\alpha) = \frac{\sqrt{2}}{2}$, also $\alpha = \frac{3}{4}\pi$.

Bemerkung B6.11 (n-te Wurzeln)
Aus der Polarkoordinatendarstellung einer komplexen Zahl

$$w = r \cdot \big(\cos(\alpha) + i \cdot \sin(\alpha)\big)$$

läßt sich leicht ableiten, daß die Zahl

$$a = \sqrt[n]{r} \cdot \left(\cos\left(\tfrac{\alpha}{n}\right) + i \cdot \sin\left(\tfrac{\alpha}{n}\right)\right)$$

eine n-te Wurzel aus w ist, d.h.

$$a^n = w.$$

Dabei ist $\sqrt[n]{r}$ die eindeutig bestimmte nicht-negative n-te Wurzel der nicht-negativen Zahl r.

Die obige Zahl a ist aber nicht die einzige Lösung der Gleichung

$$z^n = w$$

in $\mathbb{C}$. Denn addiert man zum Argument einen der folgenden Winkel

$$\frac{2\pi k}{n}, \quad \text{mit } k = 1, \ldots, n-1,$$

so erhalten wir

$$\begin{aligned}\left(\sqrt[n]{r} \cdot \left(\cos\left(\tfrac{\alpha+2\pi k}{n}\right) + i \cdot \sin\left(\tfrac{\alpha+2\pi k}{n}\right)\right)\right)^n &= \sqrt[n]{r}^{\,n} \cdot (\cos(\alpha + 2\pi k) + i \cdot \sin(\alpha + 2\pi k)) \\ &= \sqrt[n]{r}^{\,n} \cdot (\cos(\alpha) + i \cdot \sin(\alpha)) = w.\end{aligned}$$

Wir haben also in der Tat n verschiedene n-te Wurzeln von w gefunden:

$$a_k = \sqrt[n]{r} \cdot \left(\cos\left(\tfrac{\alpha+2\pi \cdot k}{n}\right) + i \cdot \sin\left(\tfrac{\alpha+2\pi \cdot k}{n}\right)\right), \quad k = 0, \ldots, n-1.$$

Damit sehen wir, daß die Polynomgleichung

$$z^n = 1$$

in $\mathbb{C}$ genau n Lösungen hat, wobei n der Grad der Gleichung ist. Das ist ein Spezialfall des Fundamentalsatzes der Algebra.

Aufgaben

Aufgabe B6.12

Bestimme für die folgenden komplexen Zahlen $\operatorname{Re} z$, $\operatorname{Im} z$, $\arg z$, $|z|$, $\overline{z}$ und z^{-1}:

(a) $z = i - 1$.

(b) $z = \frac{4i}{1+i}$.

(c) $z = \frac{(2+2i)^7}{(1-i)^3}$.

§ B7 Der Polynomring $K[t]$

Wir führen hier den Begriff des Polynoms noch mal ausführlich ein. Auf Beweise werden wir weitgehend verzichten, da diese Bestandteil weiterführender Vorlesungen sind.

Definition B7.1 (Der Polynomring)
Wir nennen einen Ausdruck der Form

$$\sum_{k=0}^{n} a_k \cdot t^k = a_n \cdot t^n + a_{n-1} \cdot t^{n-1} + \ldots + a_1 \cdot t^1 + a_0 \cdot t^0$$

mit $a_0, \ldots, a_n \in K$ und $n \in \mathbb{N}$ ein *Polynom* in der Unbestimmten t mit Koeffizienten in K.

Für zwei Polynome $f = \sum_{k=0}^{n} a_k \cdot t^k$ und $g = \sum_{k=0}^{m} b_k \cdot t^k$ soll gelten

$$f = g \quad \Longleftrightarrow \quad a_k = b_k \quad \forall\, k = 0, \ldots, \max\{m, n\}, \tag{70}$$

wobei $a_k = 0$ für $k > n$ und $b_k = 0$ für $k > m$. Wir sagen, zwei Polynome sind gleich, wenn ihre Koeffizienten übereinstimmen und sprechen dabei vom *Koeffizientenvergleich.*

Die Menge

$$K[t] = \left\{ \sum_{k=0}^{n} a_k \cdot t^k \;\middle|\; a_k \in K, n \in \mathbb{N} \right\}$$

aller Polynome in der Unbestimmten t mit Koeffizienten in K heißt der *Polynomring* in der Unbestimmten t über dem Körper K.

Für zwei Polynome $f = \sum_{k=0}^{n} a_k \cdot t^k \in K[t]$ und $g = \sum_{k=0}^{m} b_k \cdot t^k \in K[t]$ sowie ein Skalar $\lambda \in K$ definieren wir

$$\lambda \cdot f = \sum_{k=0}^{n} (\lambda \cdot a_k) \cdot t^k$$

und

$$f + g = \sum_{k=0}^{\max\{m,n\}} (a_k + b_k) \cdot t^k$$

mit $a_k = 0$ für $k > n$ und $b_k = 0$ für $k > m$ sowie

$$f \cdot g = \sum_{k=0}^{m+n} \left(\sum_{i=0}^{k} a_i \cdot b_{k-i} \right) \cdot t^k = \sum_{k=0}^{m+n} \left(\sum_{i+j=k} a_i \cdot b_j \right) \cdot t^k$$

wobei $a_i = 0$ für $i > n$ und $b_j = 0$ für $j > m$.

Eine nicht-leere Teilmenge $I \subseteq K[t]$ heißt ein *Ideal* von $K[t]$, wenn für $f, g \in I$ und $h \in K[t]$ stets $f + g \in I$ und $h \cdot f \in I$ gilt.

Wir erinnern uns, daß eine K-Algebra ein K-Vektorraum mit einer Multiplikation ist, bezüglich derer der Vektorraum ein Ring mit Eins ist und mit der die Skalarmultiplikation verträglich ist (siehe Bemerkung 5.9).

Proposition B7.2 (Der Polynomring als K-Algebra.)
Der Polynomring $\mathsf{K}[\mathsf{t}]$ *ist eine kommutative* K*-Algebra mit Basis* $\mathsf{B} = \big(\mathsf{t}^k \mid k \in \mathbb{N}\big)$.

Beweis: In weiterführenden Vorlesungen wird gezeigt, daß $\mathsf{K}[\mathsf{t}]$ ein kommutativer Ring mit Eins ist (siehe [**Mar08**, Satz 6.15]). Beachtet man, daß die Skalarmultiplikation $\lambda \cdot f = (\lambda \cdot t^0) \cdot f$ erfüllt, so folgen damit automatisch auch die fehlenden Vektorraumgesetze (siehe auch Bemerkung B7.22) sowie die Verträglichkeit von Multiplikation und Skalarmultiplikation. $\mathsf{K}[\mathsf{t}]$ ist also eine K-Algebra.

Da jedes Polynom eine endliche Linearkombination der t^k ist, ist die Familie $\big(t^k \mid k \in \mathbb{N}\big)$ ein Erzeugendensystem des K-Vektorraumes $\mathsf{K}[t]$, und wegen (70) ist sie zudem linear unabhängig, da $\sum_{k=0}^n a_k \cdot t^k = 0$ genau dann gilt wenn $a_k = 0$ für alle $k = 0, \ldots, n$. □

Bemerkung B7.3
Die Abbildung $i : \mathsf{K} \hookrightarrow \mathsf{K}[t] : a \mapsto a \cdot t^0$ ist ein K-Algebrenmonomorphismus, und wir identifizieren die *konstanten Polynome* deshalb mit den Elementen aus K. Das paßt damit zusammen, daß das Polynom t^0 die Eins der K-Algebra $\mathsf{K}[t]$ ist. Wir schreiben deshalb

$$a_n \cdot t^n + \ldots + a_1 \cdot t^1 + a_0 \cdot t^0 = a_n \cdot t^n + \ldots + a_1 \cdot t + a_0.$$

Definition B7.4 (Der Grad eines Polynoms)
Sei $f = \sum_{k=0}^n a_k t^k \in \mathsf{K}[t]$ mit $a_n \neq 0$, dann heißt $\deg(f) := n$ der *Grad* von f und $\mathrm{lc}(f) := a_n$ der *Leitkoeffizient* von f. Zudem setzen wir $\deg(0) := -\infty$ und $\mathrm{lc}(0) := 0$. Ist $\mathrm{lc}(f) = 1$ oder $f = 0$, so nennen wir f *normiert*.

Beachte, ein Polynom f ist genau dann konstant, wenn $\deg(f) \leq 0$.

Lemma B7.5 (Gradformeln)
Seien $f, g \in \mathsf{K}[t] \setminus \{0\}$*. Dann gelten:*

a. $\deg(f + g) \leq \max\big\{\deg(f), \deg(g)\big\}$.
b. $\deg(f \cdot g) = \deg(f) + \deg(g)$.

Beweis: Für den Beweis verweisen wir auf [**Mar08**, Proposition 6.16]. Er ergibt sich durch Einsetzen der Definition. □

Beispiel B7.6
Sei $f = 2t + 1, g = -2t + 1 \in \mathbb{Q}[t]$, dann gilt $f + g = 2$, also $\deg(f + g) < \max\big\{\deg(f), \deg(g)\big\}$, aber $f \cdot g = -4t^2 + 1$ und somit $\deg(f \cdot g) = \deg(f) + \deg(g)$.

Definition B7.7 (Nullstellen von Polynomen)
Es sei L eine K-Algebra, $b \in \mathsf{L}$ und $f = \sum_{k=0}^n a_k \cdot t^k \in \mathsf{K}[t]$. Wir setzen

$$f(b) := \sum_{k=0}^{n} a_k \cdot b^k \in \mathsf{L}.$$

Gilt $f(b) = 0$, so heißt b eine *Nullstelle* von f in L.

Beispiel B7.8

Sei $f = t^2 - 4 = t^2 - 4 \cdot t^0 \in \mathbb{R}[t]$, $L = \mathrm{Mat}_2(\mathbb{R})$ und $b = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$, dann gilt

$$f(b) = b^2 - 4 \cdot b^0 = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}^2 - 4 \cdot \mathbb{1}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Also ist b eine Nullstelle von f in $\mathrm{Mat}_2(\mathbb{R})$.

Satz B7.9 (Division mit Rest)
Es seien $f, g \in K[t] \setminus \{0\}$ *und* $I \subseteq K[t]$ *ein Ideal.*

a. *Es gibt eindeutige Polynome* $q, r \in K[t]$ *mit* $f = q \cdot g + r$ *und* $\deg(r) < \deg(g)$.

b. *Ist* $\lambda \in K$ *eine Nullstelle von* f *in* K*, dann gibt es ein Polynom* $q \in K[t]$ *mit* $f = q \cdot (t - \lambda)$*, d.h. wir können* $t - \lambda$ *als Linearfaktor abspalten.*

c. *Ist* $\deg(f) = n$*, so hat* f *höchstens* n *Nullstellen in* K.

d. *Es gibt genau ein normiertes Polynom* $\mu \in K[t]$ *mit* $I \stackrel{!}{=} \{\mu \cdot p \mid p \in K[t]\}$.

Beweis: Für den Beweis verweisen wir auf [**Mar08**, Satz 7.27, Proposition 7.39, Satz 7.42, Korollar 7.51]. Teil a. beweist man dabei mit Induktion nach dem Grad von f und kürzt dabei den Leitterm von f mit Hilfe des Leitterms von g. Teil b. folgt unmittelbar aus Teil a. indem man f durch den Linearfaktor $t - \lambda$ teilt, und Teil c. folgt aus Teil b. mit Induktion nach n. In Teil d. wählt man μ als normiertes Nicht-Null-Polynom kleinsten Grades, wenn $I \neq 0$, und zeigt mit Hilfe von Teil a., daß jedes Polynom in I ein Vielfaches von μ ist. □

Beispiel B7.10

Sei $f = t^3 - 1 \in \mathbb{R}[t]$, dann gilt offenbar $f(1) = 1^3 - 1 = 0$. Polynomdivision liefert:

$$\begin{array}{l}
(t^3 \qquad\qquad - 1) : (t-1) = t^2 + t + 1. \\
\underline{\ t^3 - t^2\ } \\
\qquad\ \ t^2 \\
\qquad\ \underline{t^2 - t\ } \\
\qquad\qquad\ t - 1 \\
\qquad\qquad\ \underline{t - 1\ } \\
\qquad\qquad\quad\ \ -
\end{array}$$

Also gilt $f = (t^2 + t + 1) \cdot (t - 1)$.

Proposition B7.11 (Das Minimalpolynom)
Es sei L *eine* K*-Algebra und* $b \in L$.

a. *Der* Einsetzhomomorphismus

$$\phi_b : K[t] \longrightarrow L : f \mapsto f(b)$$

ist ein K*-Algebrenhomomorphismus, d.h.* $(f + g)(b) = f(b) + g(b)$*,* $(f \cdot g)(b) = f(b) \cdot g(b)$*,* $(\lambda \cdot f)(b) = \lambda \cdot f(b)$ *und* $1(b) = 1$.

b. *Es gibt ein eindeutig bestimmtes normiertes Polynom* $\mu_b \in K[t]$*, so daß*

$$\mu_b K[t] := \{\mu_b \cdot g \mid g \in K[t]\} \stackrel{!}{=} \{h \in K[t] \mid h(b) = 0\} =: \mathrm{Ker}(\phi_b).$$

μ_b *heißt das* Minimalpolynom *von* b*.*

c. *Gibt es ein* $0 \neq h \in K[t]$ *mit* $h(b) = 0$*, so ist* μ_b *das normierte Nicht-Null-Polynom kleinsten Grades, das* b *als Nullstelle hat.*

Beweis: Daß der Einsetzhomomorphismus ein K-Algebrenhomomorphismus ist, folgt unmittelbar aus den Definitionen (siehe auch [**Mar08**, Lemma 7.36]). Mithin ist der Kern

$$\mathrm{Ker}(\phi_b) = \{h \in K[t] \mid h(b) = 0\}$$

von ϕ_b ein Ideal im Ring $K[t]$, siehe [**Mar08**, Satz 6.43]. Aus der Algebra wissen wir, daß $K[t]$ ein Hauptidealring ist, siehe [**Mar08**, Satz 7.51]. Es gibt also ein normiertes Polynom $\mu_b \in K[t]$ mit

$$\mathrm{Ker}(\phi_b) = \mu_b K[t] = \{\mu_b \cdot g \mid g \in K[t]\}.$$

Die Eindeutigkeit von μ_b folgt dann leicht aus den Kürzungsregeln in $K[t]$, siehe [**Mar08**, Beispiel 7.2]. Zudem folgt aus der Gradformel unmittelbar, daß jedes Nicht-Null-Polynom in $\mu_b K[t]$ mindestens den Grad von μ_b hat. □

Bemerkung B7.12 (Polynome versus Polynomfunktionen)

a. Auch die Menge K^K aller Abbildungen von K nach K ist eine K-Algebra und die Abbildung

$$\psi : K[t] \longrightarrow K^K : f \mapsto f$$

die einem Polynom die zugehörige *Polynomfunktion* zuordnet, ist ein K-Algebrenhomomorphismus.

b. Zwei verschiedene Polynome können dieselbe Polynomfunktion liefern!
Die beiden Polynome $f = t^2 - t \in \mathbb{F}_2[t]$ und $g = 0 \in \mathbb{F}_2[t]$ induzieren beide als Polynomfunktion die Funktion konstant Null, da

$$f(0) = 0^2 - 0 = 0 \quad \text{und} \quad f(1) = 1^2 - 1 = 0.$$

Die Polynome f und g sind aber verschieden. In diesem Fall ist die Abbildung ψ aus Teil a. nicht injektiv.

c. Enthält K *unendlich* viele Elemente, so ist die Abbildung ψ aus Teil a. injektiv, d.h. zwei Polynome sind genau dann verschieden, wenn die zugehörigen Polynomfunktionen verschieden sind. Liefern nämlich zwei Polynome f und g die gleichen Polynomfunktionen, so hat die Differenz $f - g$ unendlich viele Nullstellen und muß wegen Satz B7.9 somit das Nullpolynom sein.

Definition B7.13 (Vielfachheiten von Nullstellen)

a. Ist $\lambda \in K$ und $f = (t - \lambda)^m \cdot g \in K[t]$ mit $m \geq 1$ und $g(\lambda) \neq 0$, so nennen wir λ eine Nullstelle von f mit *Vielfachheit* $\mathrm{mult}(f, \lambda) = m$.

b. Es sei $K \subseteq L$ ein Teilkörper des Körpers L und $f \in K[t]$ mit $n = \deg(f) > 0$. Gibt es $b_1, \ldots, b_n \in L$ und ein $0 \neq c \in L$ mit $f = c \cdot (t - b_1) \cdot \ldots \cdot (t - b_n)$ so sagen wir, daß f über L in *Linearfaktoren* zerfällt.

c. Wir nennen einen Körper K *algebraisch abgeschlossen*, wenn jedes nicht-konstante Polynom in $K[t]$ über K in Linearfaktoren zerfällt.

Beispiel B7.14
Betrachte das Polynom $f = t^4 - 2t^3 + 2t^2 - 2t + 1 = (t-1)^2 \cdot (t^2+1) \in \mathbb{R}[t]$. Dann ist $\lambda = 1$ keine Nullstelle von $t^2 + 1$. Mithin ist $\lambda = 1$ eine Nullstelle von f mit Vielfachheit $\mathrm{mult}(f, 1) = 2$. Man beachte, daß f über $\mathbb{R}$ nicht in Linearfaktoren zerfällt, da t^2+1 keine Nullstelle besitzt. Über $\mathbb{C}$ zerfällt f hingegen in Linearfaktoren

$$f = (t-1)^2 \cdot (t-i) \cdot (t+i).$$

Satz B7.15 (Fundamentalsatz der Algebra)

a. *Der Köper $\mathbb{C}$ der komplexen Zahlen ist algebraisch abgeschlossen.*

b. *Jeder Körper K ist Teilkörper eines algebraisch abgeschlossenen Körpers. Der kleinste solche Oberkörper von K ist bis auf Isomorphie eindeutig bestimmt und wird der* algebraische Abschluß $\overline{K}$ *von K genannt.*

Beweis: Die als Fundamentalsatz der Algebra bekannte Aussage in Teil a. wird in der Vorlesung Einführung in die Funktionentheorie mit Mitteln der Analysis bewiesen und unter Umständen auch in der Vorlesung Einführung in die Algebra mit algebraischen Mitteln. Die Aussage in Teil b. ist Bestandteil der Vorlesung Einführung in die Algebra. □

Ist der Körper K nicht algebraisch abgeschlossen, so zerfällt nicht jedes Polynom in Linearfaktoren. Zumindest aber läßt sich jedes Polynom als Produkt von nicht mehr weiter zerlegbaren Polynomen schreiben.

Definition B7.16 (Irreduzible Polynome)
Ein nicht-konstantes Polynom $f \in K[t] \setminus K$ heißt *irreduzibel*, wenn aus $f = g \cdot h$ mit $g, h \in K[t]$ stets $\deg(g) = 0$ oder $\deg(h) = 0$ folgt.

Beispiel B7.17
Aus $f = g \cdot h$ folgt mit der Gradformel

$$\deg(f) = \deg(g) + \deg(h).$$

Ist $\deg(f) = 1$, so folgt unmittelbar, daß f irreduzibel ist. Ist $\deg(f) \in \{2, 3\}$, so ist f genau dann irreduzibel, wenn man von f keinen Faktor vom Grad 1 abspalten kann, d.h. wenn f keine Nullstelle in K hat.

Satz B7.18 (Primfaktorzerlegung im Polynomring)
Jedes nicht-konstante normierte Polynom in $K[t]$ läßt sich als Produkt von endlich vielen normierten irreduziblen Polynomen schreiben, und diese Faktoren sind bis auf die Reihenfolge eindeutig bestimmt.

Beweis: Für den Beweis verweisen wir auf [**Mar08**, Satz 7.65]. □

Beispiel B7.19
Das Polynom $f = t^4 - 2t^3 + 2t^2 - 2t + 1$ aus Beispiel B7.14 hat in $\mathbb{R}[t]$ die Primfaktorzerlegung

$$f = (t-1)^2 \cdot \left(t^2+1\right)$$

und in $\mathbb{C}[t]$ die Primfaktorzerlegung

$$f = (t-1)^2 \cdot (t-i) \cdot (t+i).$$

Satz B7.20 (Bézout-Identität)
Seien $f, g \in K[t]$ *zwei normierte teilerfremde Polynome, d.h. sie haben keinen Primfaktor gemeinsam, so gibt es Polynome* $p, q \in K[t]$ *mit*

$$1 = p \cdot f + q \cdot g.$$

Allgemeiner gilt, sind $q_1, \ldots, q_r \in K[t]$ *normierte Polynome und gibt es keinen Primfaktor, den alle gemeinsam haben, dann gibt es Polynome* $p_1, \ldots, p_r \in K[t]$ *mit*

$$1 = p_1 \cdot q_1 + \ldots + p_r \cdot q_r.$$

Beweis: Für den Beweis verweisen wir auf [**Mar08**, Korollar 7.51,Satz 7.54].

Wir wollen die Aussage hier im Spezialfall $g = t - \lambda$ beweisen. Teilen wir f durch $t - \lambda$ mit Rest, so finden wir Polynome $q, r \in K[t]$ mit

$$f = q \cdot (t - \lambda) + r$$

und $\deg(r) < \deg(t-\lambda) = 1$. Damit ist r eine Konstante. Wäre $r = 0$, so wäre $t-\lambda$ ein gemeinsamer Primfaktor von f und g, also ist $r \neq 0$. Dann ist

$$1 = \frac{1}{r} \cdot f - \frac{q}{r} \cdot g$$

die gesuchte Darstellung. □

Bemerkung B7.21 (Rationale Funktionen)
Der Polynomring $K[t]$ über einem Körper hat sehr viele Eigenschaften mit dem Ring $\mathbb{Z}$ der ganzen Zahlen gemeinsam – in beiden Ringen gibt es eine Division mit Rest (d.h. sie sind Euklidische Ringe im Sinne der Algebra, siehe [**Mar08**, Beispiel 7.26, Korollar 7.28]) und in beiden Ringen hat man eine eindeutige Primfaktorzerlegung (siehe [**Mar08**, Korollar 7.63, Korollar 7.65]). Deshalb kann man bei den Polynomen das Problem der fehlenden multiplikativen Inversen genauso lösen wie im Fall der ganzen Zahlen, man führt Brüche ein. Dies führt zum Körper

$$K(t) = \left\{ \frac{f}{g} \;\middle|\; f, g \in K[t], g \neq 0 \right\}$$

der *rationalen Funktionen*. Das Kürzen von Brüchen funktioniert wie in den rationalen Zahlen, und gleiches gilt für die Addition und die Multiplikation.

Bemerkung B7.22 (Abbrechende Folgen)
Wir bezeichnen mit $e_k = (\delta_{ik} \mid i \in \mathbb{N}) \in K^{\mathbb{N}}$ die Folge in K, die an der Stelle k den Eintrag 1 und sonst stets den Eintrag 0 hat, und wir betrachten den K-Vektorraum

$$V = \operatorname{Lin}(e_k \mid k \in \mathbb{N}) = \operatorname{Lin}(e_0, e_1, e_2, \ldots)$$

der sogenannten *abbrechenden Folgen* — beachte, daß eine Folge in V nur endlich viele Glieder ungleich Null haben kann!

$B = (e_k \mid k \in \mathbb{N})$ ist eine Basis von V, und es gibt mithin genau eine K-lineare Abbildung

$$\varphi : V \longrightarrow K[t]$$

mit

$$\varphi(e_k) = t^k$$

für alle $k \in \mathbb{N}$. Da φ eine Basis auf eine Basis abbildet, ist φ ein Isomorphismus. Wir können die Vektorräume V und $K[t]$ mittels φ miteinander identifizieren.

Zudem existiert wegen Aufgabe B7.23 auf V genau eine Multiplikation, die bilinear ist und für die $e_i \cdot e_j = e_{i+j}$ gilt. Mit dieser Multiplikation wird V eine K-Algebra und φ wird zum K-Algebrenisomorphismus.

In vielen Büchern zur Linearen Algebra ist es üblich, den Polynomring als den Vektorraum der abbrechenden Folgen einzuführen und auf diesem dann wie oben eine Multiplikation einzuführen. Dies hat den Vorteil, daß Polynome klar definierte Ojekte, nämlich abbrechende Folgen, sind und man Polynome nicht als *Ausdrücke* einer bestimmten Form einführen muß, ohne genau zu sagen, was das eigentlich heißen soll. Ich habe auf diesen Zugang verzichtet, da es Studienanfänger erfahrungsgemäß eher verwirrt, wenn Polynome Folgen sein sollen, während sie meist ohne Probleme hinnehmen, nicht wirklich gesagt zu bekommen, was ein Ausdruck der Form $\sum_{k=0}^{n} a_k \cdot t^k$ eigentlich sein soll. Die Identifikation von V mit $K[t]$ rechtfertigt unser Vorgehen nun im Nachhinein und uns reicht im weiteren Verlauf, daß wir wissen, wie wir mit Polynomen zu rechnen haben.

Aufgaben

Aufgabe B7.23 (Existenz- und Eindeutigkeitssatz für bilineare Abbildungen)
Sei V ein K-Vektorraum mit Basis $B = (x_i \mid i \in I)$ und sei $F = \big(y_{ij} \mid (i,j) \in I \times I\big)$ eine Familie von Vektoren im K-Vektorraum W. Dann gibt es genau eine multilineare Abbildung $f : V \times V \to W$ mit $f(x_i, x_j) = y_{ij}$ für alle $(i,j) \in I \times I$.

Sind $x = \sum_{\substack{i \in I \\ \text{endlich}}} a_i x_i, y = \sum_{\substack{i \in I \\ \text{endlich}}} b_i x_i \in V$, so gilt

$$f(x,y) = \sum_{\substack{i \in I \\ \text{endlich}}} \sum_{\substack{j \in I \\ \text{endlich}}} a_i b_j y_{ij}. \tag{71}$$

Aufgabe B7.24 (Polynominterpolation)
Es seien $b_0, \ldots, b_n \in K$ paarweise verschieden und $c_0, \ldots, c_n \in K$ beliebig. Zeige, es gibt genau ein Polynom $f \in K[t]$ vom Grad $\deg(f) \leq n$ mit $f(b_i) = c_i$ für alle $i = 0, \ldots, n$.

Aufgabe B7.25
Zerlege das Polynom $f = t^4 + t^3 + 2t - 4 \in \mathbb{C}[t]$ in Linearfaktoren.

Aufgabe B7.26
Bestimme das Minimalpolynom $\mu_b \in \mathbb{Q}[t]$ von $b = \sqrt{2} \in \mathbb{R}$.

Aufgabe B7.27
Zeige, ist $f \in \mathbb{R}[t]$ irreduzibel, so ist $\deg(f) \in \{1, 2\}$.

Hinweis: Betrachte für eine komplexe Nullstelle λ von f die Fälle $\lambda \in \mathbb{R}$ und $\lambda \in \mathbb{C} \setminus \mathbb{R}$. In letzterem Fall zeige, daß auch das konjugiert Komplexe von λ eine Nullstelle von f ist und betrachte dann das Polynom $g = (t - \lambda) \cdot (t - \overline{\lambda})$.

Aufgabe B7.28
Es sei $f \in \mathrm{End}_K(V)$ und $x \in V$.

a. Zeige, daß die Menge
$$I_{f,x} := \{p \in K[t] \mid p(f)(x) = 0\}$$
ein Ideal in $K[t]$ ist.

b. Zeige, daß die Menge
$$U_{f,x} := \{p(f)(x) \mid p \in K[t]\}$$
ein Unterraum von V ist.

c. Zeige, ist $m \in \mathbb{N}$ minimal mit $f^m(x) = 0$, so gilt
$$I_{f,x} = \{t^m \cdot p \mid p \in K[t]\}$$
und
$$U_{f,x} = \mathrm{Lin}\left(f^{m-1}(x), f^{m-2}(x), \ldots, f(x), x\right)$$
ist ein zyklischer Unterraum wie in Aufgabe 5.38.

ANHANG C

Einige Ergänzungen zur linearen Algebra

§ C1 Bilinearformen und Sesquilinearformen

A) Bilinearformen

In Definition 8.8 haben wir den Begriff einer multlinearen Abbildung auf Vektorräumen eingeführt. In diesem Abschnitt wollen wir einen besonders wichtigen Spezialfall dieser Begriffsbildung untersuchen, die Bilinearformen.

Definition C1.1 (Bilinearformen)
Es sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung

$$b : V \times V \to \mathbb{K},$$

die linear in beiden Argumenten ist, nennen wir *bilinear* oder eine *Bilinearform*, d. h. für $x, y, z \in V$ und $\lambda, \mu \in \mathbb{K}$ gilt (vgl. Definition 8.8):

$$b(\lambda x + \mu y, z) = \lambda b(x, z) + \mu b(y, z)$$

und

$$b(z, \lambda x + \mu y) = \lambda b(z, x) + \mu b(z, y).$$

Die Menge aller Bilinearformen auf V bezeichnen wir mit

$$\mathrm{Bil}_{\mathbb{K}}(V) = \{b : V \times V \to \mathbb{K} \mid b \text{ ist bilinear}\}.$$

Beispiel C1.2 (Bilinearformen)

a. Die Determinante definiert eine Bilinearform

$$\det : \mathbb{K}^2 \times \mathbb{K}^2 \longrightarrow \mathbb{K} : \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right) \mapsto \det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} = x_1 y_2 - x_2 y_1.$$

b. Ist $A \in \mathrm{Mat}_n(\mathbb{K})$ eine quadratische Matrix, dann wird durch

$$b_A : \mathbb{K}^n \times \mathbb{K}^n \longrightarrow \mathbb{K} : (x, y) \mapsto b_A(x, y) = x^t \circ A \circ y = x^t A y$$

eine Bilinearform auf $\mathbb{K}^n$ definiert, wie unmittelbar aus der Distributivität des Matrixproduktes folgt.

c. Wählen wir in Teil b. die Matrix $A = \mathbb{1}_n$, so erhalten wir die Bilinearform

$$b_A : \mathbb{K}^n \times \mathbb{K}^n \longrightarrow \mathbb{K} : (x, y) \mapsto \sum_{i=1}^{n} x_i \cdot y_i.$$

d. Ist $n = 2$ und ist

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{K}),$$

so ist $b_A = \det$.

Bemerkung C1.3

Die Menge $\mathrm{Bil}_{\mathbb{K}}(V)$ ist offenbar ein Unterraum des $\mathbb{K}$-Vektorraums $\mathbb{K}^{(V\times V)}$ aller Abbildungen von $V \times V$ nach $\mathbb{K}$, d.h. die Summe zweier Bilinearformen sowie das skalare Vielfache einer Bilinearform sind wieder bilinear.

Definition C1.4 (Matrixdarstellung von Bilinearformen)

Es sei V ein $\mathbb{K}$-Vektorraum mit Basis $B = (x_1, \ldots, x_n)$ und $b : V \times V \to \mathbb{K}$ eine Bilinearform auf V. Wir nennen die Matrix

$$M_B(b) = \big(b(x_i, x_j)\big)_{1\leq i,j\leq n} \in \mathrm{Mat}_n(\mathbb{K})$$

die *Matrixdarstellung* von b bezüglich der Basis B.

Beispiel C1.5

Wir betrachten die Bilinearform $b = b_{\mathbb{1}_2}$ auf $\mathbb{R}^2$ sowie die Basis $B = (x_1, x_2) = \big((1,1)^t, (1,2)^t\big)$. Dann ist

$$M_B(b) = \begin{pmatrix} b(x_1,x_1) & b(x_1,x_2) \\ b(x_2,x_1) & b(x_2,x_2) \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}.$$

Bemerkung C1.6 (Matrixdarstellung von Vektoren)

Wir sollten darauf hinweisen, daß wir die Bezeichnung M_B bereits einmal verwendet haben, nämlich bei der Matrixdarstellung von Vektoren bezüglich einer Basis $B = (x_1, \ldots, x_n)$. Ist $x = \sum_{i=1}^n \lambda_i x_i$, so bezeichnet

$$M_B(x) = (\lambda_1, \ldots, \lambda_n)^t \in \mathbb{K}^n$$

die Matrixdarstellung des Vektors $x \in V$ bezüglich der Basis B. Wir werden diese Bezeichnung auch im folgenden wieder benötigen, es wird aus dem Kontext aber stets unmittelbar ersichtlich sein, in welcher Bedeutung M_B verwendet wird.

Proposition C1.7 (Matrixdarstellung von Bilinearformen)

Es sei V ein $\mathbb{K}$-Vektorraum mit Basis $B = (x_1, \ldots, x_n)$. Dann ist die Abbildung

$$M_B : \mathrm{Bil}_{\mathbb{K}}(V) \longrightarrow \mathrm{Mat}_n(\mathbb{K}) : b \mapsto M_B(b)$$

ein Isomorphismus von $\mathbb{K}$-Vektorräumen.

Insbesondere ist eine Bilinearform durch ihre Matrixdarstellung eindeutig bestimmt, und für $x, y \in V$ gilt

$$b(x,y) = M_B(x)^t \circ M_B(b) \circ M_B(y).$$

Beweis: Sind $b, b' \in \mathrm{Bil}_{\mathbb{K}}(V)$ und $\lambda, \lambda' \in \mathbb{K}$, so gilt

$$\begin{aligned} M_B(\lambda \cdot b + \lambda' \cdot b') =& \big((\lambda b + \lambda' b')(x_i, x_j)\big)_{i,j} = \big(\lambda \cdot b(x_i, x_j) + \lambda' \cdot b'(x_i, x_j)\big)_{i,j} \\ =& \lambda \cdot \big(b(x_i, x_j)\big)_{i,j} + \lambda' \cdot \big(b'(x_i, x_j)\big)_{i,j} = \lambda \cdot M_B(b) + \lambda' \cdot M_B(b'). \end{aligned}$$

Also ist die Abbildung M_B linear.

Ferner gilt für $x = \sum_{i=1}^n \lambda_i x_i$ und $y = \sum_{j=1}^n \mu_j x_j$

$$b(x, y) = \sum_{i=1}^n \lambda_i b(x_i, y) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j b(x_i, x_j) = M_B(x)^t \cdot M_B(b) \cdot M_B(y).$$

Die Matrixdarstellung legt die Bilinearform also eindeutig fest, und somit ist M_B injektiv. Für die Surjektivität von M_B sei eine $A = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{K})$ beliebig gegeben. Wir definieren nun eine Abbildung

$$b : V \times V \longrightarrow \mathbb{K}$$

durch

$$b(x, y) = M_B(x)^t \circ A \circ M_B(y) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \cdot \mu_j \cdot a_{ij},$$

wenn $x = \sum_{i=1}^n \lambda_i x_i$ und $y = \sum_{j=1}^n \mu_j x_j$. Da die Matrixdarstellung $M_B : V \longrightarrow \mathbb{K}^n$ linear ist und da die Matrixmultiplikation zudem distributiv ist, ist $b \in \mathrm{Bil}_{\mathbb{K}}(V)$ eine bilineare Abbildung, d.h. für $\lambda, \mu \in \mathbb{K}$ und $x, y, z \in V$ gilt

$$\begin{aligned} b(\lambda x + \mu y, z) =& M_B(\lambda x + \mu y)^t \circ A \circ M_B(z) = \big(\lambda M_B(x) + \mu M_B(y)\big)^t \circ A \circ M_B(z) \\ =& \lambda M_B(x)^t \circ A \circ M_B(z) + \mu M_B(y)^t \circ A \circ M_B(z) = \lambda b(x, z) + \mu b(y, z) \end{aligned}$$

und

$$\begin{aligned} b(z, \lambda x + \mu y) =& M_B(z)^t \circ A \circ M_B(\lambda x + \mu y) = M_B(z)^t \circ A \circ \big(\lambda M_B(x) + \mu M_B(y)\big) \\ =& M_B(z)^t \circ A \circ \lambda M_B(x) + M_B(z)^t \circ A \circ \mu M_B(y) = \lambda b(z, x) + \mu b(z, y). \end{aligned}$$

Außerdem gilt

$$b(x_i, x_j) = M_B(x_i)^t \circ A \circ M_B(x_j) = e_i^t \circ A \circ e_j = a_{ij}$$

für $1 \leq i, j \leq n$, so daß $M_B(b) = A$. Somit ist b auch surjektiv. □

Beispiel C1.8

Wir betrachten die Bilinearform aus Beispiel C1.5 sowie die Vektoren $x = (0, 1)^t = -x_1 + x_2$ und $y = (3, 4)^t = 2x_1 + x_2$. Dann gilt

$$M_B(x) = (-1, 1)^t \quad \text{und} \quad M_B(y) = (2, 1)^t$$

sowie

$$4 = b(x, y) = M_B(x)^t \circ M_B(b) \circ M_B(y) = \begin{pmatrix} -1 & 1 \end{pmatrix} \circ \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} \circ \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Korollar C1.9 (Alle Bilinearformen auf $\mathbb{K}^n$ sind von der Form b_A.)
Die Abbildung $\mathrm{Mat}_n(\mathbb{K}) \longrightarrow \mathrm{Bil}_{\mathbb{K}}(\mathbb{K}^n) : A \mapsto b_A$ *ist ein Isomorphismus. Insbesondere ist jede Bilinearform auf* $\mathbb{K}^n$ *von der Gestalt* b_A *für eine eindeutig bestimmte Matrix* $A \in \mathrm{Mat}_n(\mathbb{K})$.

Beweis: Die Abbildung $A \mapsto b_A$ ist die Umkehrabbildung der Matrixdarstellung

$$M_E : \mathrm{Bil}_{\mathbb{K}}\left(\mathbb{K}^n\right) \longrightarrow \mathrm{Mat}_n(\mathbb{K})$$

bezüglich der kanonischen Basis $E = (e_1, \ldots, e_n)$ von $\mathbb{K}^n$, da

$$b_A(e_i, e_j) = e_i^t \circ A \circ e_j = a_{ij}$$

und somit $M_E(b_A) = A$. □

Bemerkung C1.10
Die Aussage in Korollar C1.9 bedeutet insbesondere, daß

$$b = b_{M_E(b)} \quad \text{und} \quad A = M_E(b_A).$$

für $b \in \mathrm{Bil}_{\mathbb{K}}(\mathbb{K}^n)$ und $A \in \mathrm{Mat}_n(\mathbb{K})$, wenn E die kanonische Basis von $\mathbb{K}^n$ ist.

Satz C1.11 (Basiswechsel bei Bilinearformen)
Sei V *ein* $\mathbb{K}$*-Vektorraum mit Basen* $B = (x_1, \ldots, x_n)$ *und* $D = (y_1, \ldots, y_n)$ *und sei* $b \in \mathrm{Bil}_{\mathbb{K}}(V)$, *dann gilt*

$$M_B(b) = \left(T_D^B\right)^t \circ M_D(b) \circ T_D^B.$$

Beweis: Wir setzen $M_B(b) = (a_{ij})$ und $\left(T_D^B\right)^t \circ M_D(b) \circ T_D^B = C = (c_{ij}) \in \mathrm{Mat}_n(\mathbb{K})$. Man beachte, daß für einen beliebigen Vektor $z \in V$ nach Proposition 5.4 die Gleichung

$$M_D(z) = M_D\big(\mathrm{id}_V(z)\big) = M_D^B(\mathrm{id}_V) \circ M_B(z) = T_D^B \circ M_B(z) \tag{72}$$

gilt. Damit erhalten wir für $1 \leq i, j \leq n$ dann

$$\begin{aligned} a_{ij} &\overset{\text{Def.}}{=} b(x_i, x_j) \overset{\text{C1.7}}{=} M_D(x_i)^t \circ M_D(b) \circ M_D(x_j) \\ &\overset{(72)}{=} \left(T_D^B \circ M_B(x_i)\right)^t \circ M_D(b) \circ \left(T_D^B \circ M_B(x_j)\right) \\ &= M_B(x_i)^t \circ \left(T_D^B\right)^t \circ M_D(b) \circ T_D^B \circ M_B(x_j) = e_i^t \circ C \circ e_j = c_{ij}. \end{aligned}$$

Also gilt $M_B(b) = (a_{ij}) = (c_{ij}) = \left(T_D^B\right)^t \circ M_D(b) \circ T_D^B$, was zu zeigen war. □

Beispiel C1.12
Die Matrixdarstellung der Bilinearform $b = b_{\mathbb{1}_2}$ aus Beispiel C1.5 bezüglich der kanonischen Basis E ist die Einheitsmatrix $\mathbb{1}_2$. Zugleich ist

$$T_E^B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

die Transformationsmatrix für die Basis $B = \big((1,1)^t, (1,2)^t\big)$. Dann gilt aber

$$\begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} = M_B(b) = \big(T_E^B\big)^t \circ M_E(b) \circ T_E^B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \circ \mathbb{1}_2 \circ \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Bemerkung C1.13
Es bleibt festzuhalten, daß sowohl Endomorphismen $f : V \to V$ als auch Bilinearformen $b : V \times V \to \mathbb{K}$ sich nach Wahl einer Basis B durch Matrizen $M_B^B(f)$ bzw. $M_B(b)$ beschreiben lassen. Bei Basiswechsel, der durch die Matrix $T = T_D^B$ beschrieben wird, haben Endomorphismen und Bilinearformen aber ein unterschiedliches Transformationsverhalten. Es gilt:

$$\boxed{M_B^B(f) = T^{-1} \circ M_D^D(f) \circ T \qquad \text{und} \qquad M_B(b) = T^t \circ M_D(b) \circ T.}$$

B) Normalform symmetrischer Bilinearformen

Definition C1.14 (Symmetrische Bilinearformen)
Es sei V ein $\mathbb{K}$-Vektorraum.

a. Eine Bilinearform $b \in \mathrm{Bil}_{\mathbb{K}}(V)$ heißt *symmetrisch*, falls für $x, y \in V$ stets gilt

$$b(x,y) = b(y,x).$$

b. Eine Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ heißt *symmetrisch*, falls $A = A^t$.

Beispiel C1.15

a. Die Bilinearform det aus Beispiel C1.2 ist für $\mathbb{K} = \mathbb{R}$ nicht symmetrisch, da

$$\det(e_1, e_2) = 1 \neq -1 = \det(e_2, e_1).$$

b. Die Matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{Q})$$

ist symmetrisch.

Proposition C1.16 (Symmetrische Bilinearformen)
Sei V ein $\mathbb{K}$-Vektorraum mit Basis $B = (x_1, \ldots, x_n)$, $b \in \mathrm{Bil}_{\mathbb{K}}(V)$ und $A \in \mathrm{Mat}_n(\mathbb{K})$.

a. *Die Bilinearform b ist genau dann symmetrisch, wenn $M_B(b)$ symmetrisch ist.*

b. *Die Bilinearform b_A ist genau dann symmetrisch, wenn A symmetrisch ist.*

Beweis: Es sei $M_B(b) = A = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{K})$.

Ist A symmetrisch, so ist auch b symmetrisch, da für $x, y \in V$ gilt

$$\begin{aligned} b(x,y) =& M_B(x)^t \circ A \circ M_B(y) = \big(M_B(x)^t \circ A \circ M_B(y)\big)^t \\ =& M_B(y)^t \circ A^t \circ M_B(x) = M_B(y)^t \circ A \circ M_B(x) = b(y,x). \end{aligned}$$

Ist umgekehrt b symmetrisch, dann folgt für $1 \leq i, j \leq n$

$$a_{ij} = b(x_i, x_j) = b(x_j, x_i) = a_{ji},$$

so daß auch $A = A^t$ symmetrisch ist. Damit ist a. gezeigt, und b. folgt aus a., da $A = M_E(b_A)$ für die kanonische Basis E von $\mathbb{K}^n$ gilt. □

Definition C1.17 (Quadratische Form)
Ist $b \in \mathrm{Bil}_{\mathbb{K}}(V)$ eine symmetrische Bilinearform auf V, dann nennen wir

$$q_b : V \to \mathbb{K} : x \mapsto b(x, x)$$

die *quadratische Form* zu b. Für $A \in \mathrm{Mat}_n(\mathbb{K})$ schreiben wir auch q_A statt q_{b_A}.

Beispiel C1.18
Ist $A = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{K})$ symmetrisch, dann gilt für $x = (t_1, \ldots, t_n)^t \in \mathbb{K}^n$

$$q_A(x) = x^t A x = \sum_{i,j} a_{ij} t_i t_j = \sum_{i=1}^{n} a_{ii} t_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} t_i t_j.$$

Damit können wir q_A als ein homogenes quadratisches Polynom in den Unbestimmten $t_1, \ldots, t_n$ auffassen.

Z.B., ist

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R}),$$

so gilt

$$q_A(t_1, t_2) = t_1^2 + 4t_1 t_2 + 5t_2^2.$$

Auf den ersten Blick scheint es, daß die quadratische Form q_b weit weniger Information enthält, als die symmetrische Bilinearform b. Erstaunlicherweise kann man b jedoch aus q_b zurückgewinnen.

Proposition C1.19 (Polarisierung einer Bilinearform)
Sei $b \in \mathrm{Bil}_{\mathbb{K}}(V)$ eine symmetrische Bilinearform. Dann gilt für $x, y \in V$

$$b(x, y) = \tfrac{1}{2}\big(q_b(x + y) - q_b(x) - q_b(y)\big).$$

Beweis: Die Aussage folgt durch einfaches Einsetzen der Definition von q_b in die rechte Seite. □

Satz C1.20 (Existenz einer Orthogonalbasis)
Sei V ein $\mathbb{K}$-Vektorraum mit $1 \leq \dim_{\mathbb{K}}(V) < \infty$. Ist $b \in \mathrm{Bil}_{\mathbb{K}}(V)$ eine symmetrische Bilinearform, dann besitzt V eine Basis $B = (x_1, \ldots, x_n)$, so daß $M_B(b)$ eine Diagonalmatrix ist, d.h.

$$b(x_i, x_j) = 0 \quad \forall\, i \neq j.$$

Wir nennen eine solche Basis B eine Orthogonalbasis *bezüglich der Bilinearform b.*

Beweis: Wir führen den Beweis durch Induktion über $n = \dim_{\mathbb{K}}(V)$, wobei im Fall $n = 1$ nichts zu zeigen ist.

Sei also $n > 1$. Wir bezeichnen mit $q_b : V \to \mathbb{K} : x \mapsto b(x, x)$ die zu b gehörende quadratische Form. Ist q_b identisch Null, so ist nach Proposition C1.19 auch b identisch Null, und jede Basis ist eine Orthogonalbasis bezüglich b.

Wir können also annehmen, daß es ein $x \in V$ gibt mit $b(x, x) = q_b(x) \neq 0$. Setze $U := \mathrm{Lin}\,(x)$ und

$$U^\perp := \{y \in V \mid b(x, y) = 0\}.$$

Aus der Bilinearität von b folgt, daß $U^\perp$ ein Unterraum von V ist. Wir wollen nun zeigen, daß in der Tat $V = U \oplus U^\perp$ gilt.

Sei dazu zunächst $y \in V$ beliebig. Dann setzen wir

$$x' := \tfrac{b(x,y)}{b(x,x)} \cdot x \in U$$

und erhalten

$$b\big(y - x', x\big) = b(y, x) - \tfrac{b(x,y)}{b(x,x)} \cdot b(x, x) = 0$$

Deshalb ist $y - x' \in U^\perp$ und

$$y = x' + (y - x') \in U + U^\perp,$$

womit $V = U + U^\perp$ gezeigt ist.

Sei nun $y \in U \cap U^\perp$, dann gibt es ein $\lambda \in \mathbb{K}$ mit $y = \lambda x$ und damit

$$\lambda \cdot q_b(x) = b(x, \lambda x) = b(x, y) = 0.$$

Da aber $q_b(x) \neq 0$ gilt, ist $\lambda = 0$ und damit $y = 0$. Also gilt $U \cap U^\perp = \{0\}$. Insgesamt haben wir damit $V = U \oplus U^\perp$ gezeigt.

Schränken wir nun b auf $U^\perp$ ein, so erhalten wir per Induktion eine Orthogonalbasis $(x_2, \ldots, x_n)$ von $U^\perp$ bezüglich b und $B = (x, x_2, \ldots, x_n)$ ist dann die gesuchte Orthogonalbasis von V. □

Korollar C1.21 (Diagonalisierbarkeit symmetrischer Matrizen)
Sei $A \in \mathrm{Mat}_n(\mathbb{K})$ *eine symmetrische Matrix. Dann gibt es eine Matrix* $T \in \mathrm{Gl}_n(\mathbb{K})$ *mit*

$$T^t \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & \lambda_n \end{pmatrix},$$

wobei $\lambda_1, \ldots, \lambda_r \in \mathbb{K} \setminus \{0\}$, $\lambda_{r+1} = \ldots = \lambda_n = 0$ *und* $r = \mathrm{rang}(A)$.

Beweis: Nach Satz C1.20 besitzt $\mathbb{K}^n$ eine Orthogonalbasis $B = (x_1, \ldots, x_n)$ bezüglich der Bilinearform b_A. Dabei können wir durch Umnumerieren o. E. annehmen, daß $b_A(x_i, x_i) \neq 0$ für $i = 1, \ldots, r$ und $b_A(x_i, x_i) = 0$ für $i = r+1, \ldots, n$ für ein geeignetes r gilt. Wähle nun T als die Matrix, deren Spalten die Vektoren in B sind, dann ist $M_B(b_A) = T^t \circ A \circ T$ und hat die gewünschte Gestalt.

Es bleibt zu zeigen, daß $r = \operatorname{rang}(A)$. Aber, da T invertierbar ist, gilt

$$r = \operatorname{rang}\left(T^t \circ A \circ T\right) = \operatorname{rang}(A).$$

□

Bemerkung C1.22

a. Man beachte, daß die λ_i i. a. nicht nur von A abhängen und auch *nicht* die Eigenwerte von A sind (siehe Beispiel C1.24). Die Anzahl der Diagonalelemente ungleich Null hängt jedoch stets nur von A ab.

b. Korollar C1.21 führt zu folgender Überlegung. Da $T \in \mathrm{Gl}_n(\mathbb{K})$ ist, ist T das Produkt von Elementarmatrizen $T = P_1 \circ \ldots \circ P_k$ und somit gilt

$$D := T^t \circ A \circ T = P_k^t \circ \ldots \circ P_1^t \circ A \circ P_1 \circ \ldots \circ P_k.$$

Das heißt, daß die Diagonalmatrix D aus A durch gleichzeitiges Durchführen von elementaren Zeilenoperationen und den zugehörigen Spaltenoperationen entsteht. Dabei ist es wegen $P^t \circ (A \circ P) = \left(P^t \circ A\right) \circ P$ egal, ob zuerst die Zeilenoperation oder die Spaltenoperation durchgeführt wird.
Die Überführung einer symmetrischen Matrix A in Diagonalgestalt mittels gleichzeitiger Zeilen- und Spaltenoperationen nennt man das *symmetrische Gaußsche Eliminationsverfahren* oder den *symmetrischen Gaußalgorithmus.*
Es ist klar, daß man diesen Algorithmus ebenso einfach implementieren kann, wie den Gaußschen Algorithmus. Will man zusätzlich die Transformationsmatrix T bestimmen, so startet man wie bei der Invertierung einer Matrix mit $(A \mid \mathbb{1}_n)$, führt bei A die Zeilen- und Spaltenoperationen durch, bei $\mathbb{1}_n$ aber nur die Spaltenoperationen. Ist dann A diagonalisiert, so ist $\mathbb{1}_n$ in die Transformationsmatrix überführt.

Wir formulieren den Algorithmus nun in rekursiver Form. Die Eingabe muß dann ein Schema der Form $(A \mid \mathbb{1}_n)$ sein, damit die Ausgabe den gewünschten Erfolg hat.

Algorithmus C1.23 (Symmetrischer Gaußalgorithmus)

INPUT: A, T mit $A \in \mathrm{Mat}_n(\mathbb{K})$ symmetrisch und $T \in \mathrm{Gl}_r(\mathbb{K})$, $r \geq n$.

OUTPUT: $T \in \mathrm{Gl}_r(\mathbb{K})$ so, daß $\tilde{T}^t \circ A \circ \tilde{T}$ eine Diagonalmatrix ist, wobei $\tilde{T}$ durch Streichen der ersten $r - n$ Spalten und Zeilen aus T entsteht.

1. Schritt: Setze $m = r - n$.

2. Schritt: Man suche in der ersten Spalte von A den ersten Eintrag, der nicht Null ist. Existiert ein solcher, merke man sich die Zeilennummer z, sonst gehe man zu Schritt 5.

3. Schritt: Ist $z \neq 1$, so addiere die z-te Zeile von A zur ersten und die z-te Spalte zur ersten. Addiere ferner die $z+m$-te Spalte von T zur $m+1$-ten.

4. Schritt: Für $k = 2, \ldots, n$ addiere man das $-A[1,k]/A[1,1]$-fache der ersten Zeile von A zur k-ten und das $-A[1,k]/A[1,1]$-fache der ersten Spalte zur k-ten. Sodann addiere man das $-A[1,k]/A[1,1]$-fache der $1+m$-ten Spalte von T zur $k+m$-ten.

5. Schritt: Falls $n > 1$, dann erzeuge man eine Matrix B, indem man aus A die erste Zeile und die erste Spalte streicht. Sodann rufe man die Prozedur mit den Parametern B und T auf und speichere das Ergebnis in T.

6. Schritt: Man gebe T zurück.

Beispiel C1.24

Wir betrachten die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \in \mathrm{Mat}_n(\mathbb{R}).$$

Sodann bilden wir das Schema $(A \mid \mathbb{1}_2)$ und wenden den symmetrischen Gaußalgorithmus an:

$$\left(\begin{array}{cc|cc} 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right) \xmapsto{\text{I=I+II}} \left(\begin{array}{cc|cc} 3 & 2 & 1 & 0 \\ 2 & 1 & 1 & 1 \end{array}\right) \xmapsto{\text{II=II-}\frac{2}{3}\text{I}} \left(\begin{array}{cc|cc} 3 & 0 & 1 & -\frac{2}{3} \\ 0 & -\frac{1}{3} & 1 & \frac{1}{3} \end{array}\right).$$

Für $T = \begin{pmatrix} 1 & -\frac{2}{3} \\ 1 & \frac{1}{3} \end{pmatrix} \in \mathrm{Gl}_2(\mathbb{R})$ gilt also

$$T^t \circ A \circ T = \begin{pmatrix} 3 & 0 \\ 0 & -\frac{1}{3} \end{pmatrix}.$$

Alternativ können wir auch wie folgt vorgehen:

$$\left(\begin{array}{cc|cc} 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right) \xmapsto{\text{I}\leftrightarrow\text{II}} \left(\begin{array}{cc|cc} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{array}\right) \xmapsto{\text{II=II-I}} \left(\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & -1 \end{array}\right).$$

Dann gilt für $S = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \in \mathrm{Gl}_2(\mathbb{R})$

$$S^t \circ A \circ S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Man beachte, daß die beiden Diagonalmatrizen nicht die gleichen Diagonaleinträge besitzen!

C) Der Sylvestersche Trägheitssatz

Über dem Körper $\mathbb{K} = \mathbb{R}$ der reellen Zahlen können wir noch eine etwas schönere Normalform für Bilinearformen herleiten, bei der nicht nur die Zahl der Nicht-Null-Einträge auf der Diagonalen invariant ist.

Korollar C1.25 (Sylvesterscher Trägheitssatz für Bilinearformen über $\mathbb{R}$)
Sei V *ein* $\mathbb{R}$*-Vektorraum der Dimension* $n > 0$ *und* $b \in \mathrm{Bil}_{\mathbb{R}}(V)$ *eine symmetrische Bilinearform. Dann besitzt* V *eine Orthogonalbasis* B*, so daß*

$$M_B(b) = \left(\begin{array}{c|c|c} \mathbb{1}_k & 0 & 0 \\ \hline 0 & -\mathbb{1}_l & 0 \\ \hline 0 & 0 & 0 \end{array}\right).$$

Dabei hängen die Zahlen k *und* l *nur von* b *ab, nicht von der Orthogonalbasis* B*. Wir nennen* k *den* Trägheitsindex*,* l *den* Morseindex *und* $k-l$ *die* Signatur *von* b*.*

Beweis: Wir wählen zunächst eine Basis $D = (y_1, \ldots, y_n)$ wie in Satz C1.20, d.h. $M_D(b)$ ist eine Diagonalmatrix mit Diagonaleinträgen $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Dabei können wir nach Umnumerieren der y_i ohne Einschränkung annehmen, daß $\lambda_1, \ldots, \lambda_k > 0$, $\lambda_{k+1}, \ldots, \lambda_{k+l} < 0$ und $\lambda_{k+l+1}, \ldots, \lambda_n = 0$ für geeignete $k, l \in \mathbb{N}$ gilt. Setzen wir nun

$$x_i := \begin{cases} \frac{1}{\sqrt{|\lambda_i|}} \cdot y_i, & i = 1, \ldots, k+l, \\ y_i, & i = k+l+1, \ldots, n \end{cases}$$

und $B = (x_1, \ldots, x_n)$, dann hat $M_B(b)$ die gewünschte Gestalt, da nach wie vor $b(x_i, x_j) = 0$ für $i \neq j$ und

$$b(x_i, x_i) = \begin{cases} 1, & i = 1, \ldots, k, \\ -1, & i = k+1, \ldots, k+l, \\ 0, & i = k+l+1, \ldots, n. \end{cases}$$

Es bleibt, zu zeigen, daß die Zahlen k und l unabhängig von der Wahl von B sind. Dazu zeigen wir zunächst, daß

$$k = \max\left\{\dim_{\mathbb{R}}(U) \;\middle|\; U \leq V, q_b(x) > 0 \; \forall \; 0 \neq x \in U\right\}. \tag{73}$$

Ist $0 \neq x = \sum_{i=1}^k \lambda_i x_i \in \mathrm{Lin}\,(x_1, \ldots, x_k)$, so gilt

$$q_b(x) = b(x, x) = \sum_{i,j=1}^{k} \lambda_i \cdot \lambda_j \cdot b(x_i, x_j) = \sum_{i=1}^{k} \lambda_i^2 > 0. \tag{74}$$

Mithin ist $\mathrm{Lin}\,(x_1, \ldots, x_k)$ einer der Unterräume von V, die auf der rechten Seite betrachtet werden, und mithin ist das Maximum mindestens k.

Sei nun $U \leq V$ irgendein Unterraum von V mit $q_b(x) > 0$ für alle $0 \neq x \in U$. Für ein beliebiges $x \in W := \mathrm{Lin}\,(x_{k+1}, \ldots, x_n)$ zeigt man wie in (74), daß $q_b(x) \leq 0$ gilt. Daraus folgt unmittelbar, daß

$$U \cap W = \{0\}$$

gelten muß, und dann folgt

$$\dim_{\mathbb{R}}(U) = \dim_{\mathbb{R}}(U + W) - \dim_{\mathbb{R}}(W) + \dim_{\mathbb{R}}(U \cap W) \leq n - (n-k) = k.$$

Damit ist (73) gezeigt und k hängt allein von b ab.

Zudem gilt für jede Basis C von V

$$k + l = \operatorname{rang}\left(M_B(b)\right) = \operatorname{rang}\left(\left(T_C^B\right)^t \circ M_C(b) \circ T_C^B\right) = \operatorname{rang}\left(M_C(b)\right),$$

so daß auch $k + l$ nicht von der Wahl der Orthogonalbasis B abhängt, aber dann trifft dies auch auf die Differenz $l = (k + l) - k$ zu. □

Korollar C1.26 (Sylvesterscher Trägheitssatz für symmetrische Matrizen über $\mathbb{R}$)
Sei $A \in \operatorname{Mat}_n(\mathbb{R})$ *eine symmetrische Matrix. Dann existiert eine invertierbare Matrix* $T \in \operatorname{Gl}_n(\mathbb{R})$*, so daß*

$$T^t \circ A \circ T = \left(\begin{array}{c|c|c} \mathbb{1}_k & 0 & 0 \\ \hline 0 & -\mathbb{1}_l & 0 \\ \hline 0 & 0 & 0 \end{array}\right).$$

Dabei hängen die Zahlen k *und* l *nur von* A *ab und nicht von* T.
Wir nennen k *den* Trägheitsindex*,* l *den* Morseindex *und* $k - l$ *die* Signatur *von* A.

Beweis: Dies folgt, indem wir Korollar C1.25 auf $b = b_A$ anwenden und $T = T_E^B$ setzen, wobei B die Orthogonalbasis aus Korollar C1.25 ist und E die kanonische Basis von $\mathbb{R}^n$. □

Beispiel C1.27
In Beispiel C1.24 haben wir die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \in \operatorname{Mat}_n(\mathbb{R})$$

auf zwei mögliche Weisen als Bilinearform zu einer Diagonalmatrix transformiert. Bei der ersten Möglichkeit haben wir die Matrix

$$T = \begin{pmatrix} 1 & -\frac{2}{3} \\ 1 & \frac{1}{3} \end{pmatrix} \in \operatorname{Gl}_2(\mathbb{R})$$

verwendet und

$$T^t \circ A \circ T = \begin{pmatrix} 3 & 0 \\ 0 & -\frac{1}{3} \end{pmatrix}$$

erhalten. Unser Beweis legt nahe, die erste Spalte von T durch $\sqrt{3}$ zu dividieren und die zweite mit $\sqrt{3}$ zu multiplizieren. Wir erhalten dann

$$\tilde{T} = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}$$

und

$$\tilde{T}^t \circ A \circ \tilde{T} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

A hat also den Trägheitsindex 1, den Morseindex 1 und die Signatur 0.

Die zweite Transformation von A mittels der Matrix

$$S = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \in \mathrm{Gl}_2(\mathbb{R})$$

hat gleich zur Normalform

$$S^t \circ A \circ S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

geführt. Wir sehen, daß die Transformationsmatrix nicht eindeutig bestimmt ist.

D) Sesquilinearformen

Wenn der Grundkörper der Körper der komplexen Zahlen ist, dann kann man die Bedingung der Linearität einer Bilinearform in der ersten Komponente verändern und kommt zum Begriff der Sesquilinearform. Diese sind im Zusammenhang mit geometrischen Begriffen wie Länge und Winkel wichtig.

Definition C1.28 (Sesquilinearformen)
Es sei V ein $\mathbb{C}$-Vektorraum und $\bar{\cdot} : \mathbb{C} \longrightarrow \mathbb{C}$ bezeichne die komplexe Konjugation.

a. Eine Abbildung

$$b : V \times V \to \mathbb{C}$$

heißt *sesquilinear*[1] oder eine *Sesquilinearform*, falls für $x, y, z \in V$ und $\lambda, \mu \in \mathbb{C}$

$$b(\lambda x + \mu y, z) = \overline{\lambda} \cdot b(x, z) + \overline{\mu} \cdot b(y, z)$$

gilt sowie

$$b(z, \lambda x + \mu y) = \lambda \cdot b(z, x) + \mu \cdot b(z, y).$$

Sesq(V) bezeichnet den $\mathbb{C}$-Vektorraum aller Sesquilinearformen auf V.

b. Ist $B = (x_1, \ldots, x_n)$ eine Basis und b eine Sesquilinearform von V, so heißt

$$M_B(b) = \big(b(x_i, x_j)\big)_{1 \leq i,j \leq n} \in \mathrm{Mat}_n(\mathbb{C})$$

die *Matrixdarstellung* von b bezüglich der Basis B.

c. Für eine Matrix $A = (a_{ij})_{i,j} \in \mathrm{Mat}_n(\mathbb{C})$ heißt die Matrix

$$A^* = \overline{A}^t = \big(\overline{a_{ji}}\big)_{i,j} \in \mathrm{Mat}_n(\mathbb{C})$$

die *adjungierte Matrix* zu A.

Beispiel C1.29 (Sesquilinearformen)
Ist $A \in \mathrm{Mat}_n(\mathbb{C})$, so wird durch

$$b_A^s : \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C} : (x, y) \mapsto \overline{x}^t \circ A \circ y = x^* \circ A \circ y$$

eine Sesquilinearform auf $\mathbb{C}^n$ definiert.

[1] Sesquilinear bedeutet ein-einhalb-fach linear. Dies bezieht sich darauf, daß die Abbildung in der ersten Komponente nur die eine Hälfte der Linearitätsbedingung erfüllt.

Proposition C1.30 (Matrixdarstellung einer Sesquilinearform)
Es sei V *ein* $\mathbb{C}$*-Vektorraum mit Basen* $B = (x_1, \ldots, x_n)$ *und* $D = (y_1, \ldots, y_n)$.

a. *Zu jedem* $A \in \mathrm{Mat}_n(\mathbb{C})$ *gibt es genau ein* $b \in \mathrm{Sesq}(V)$ *mit* $M_B(b) = A$.

b. *Ist* b *eine Sesquilinearform auf* V *und sind* $x, y \in V$*, so gilt*

$$b(x, y) = M_B(x)^* \circ M_B(b) \circ M_B(y).$$

c. *Ist* b *eine Sequilinearform auf* V*, so gilt*

$$M_B(b) = \left(T_D^B\right)^* \circ M_D(b) \circ T_D^B.$$

Beweis: Der Beweis geht analog zu den entsprechenden Beweisen für Bilinearformen, siehe Proposition C1.7 und Satz C1.11. □

Definition C1.31 (Hermitesche Sesquilinearformen)

a. Eine Sesquilinearform b auf einem $\mathbb{C}$-Vektorraum V heißt *hermitesch*, falls

$$b(x, y) = \overline{b(y, x)}$$

für alle $x, y \in V$ gilt.

b. Eine Matrix $A \in \mathrm{Mat}_n(\mathbb{C})$ heißt *hermitesch* oder *selbstadjungiert*, falls $A = A^*$.

Beispiel C1.32
Die Matrix

$$A = \begin{pmatrix} 1 & i \\ -i & 0 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{C})$$

ist hermitesch, da

$$A^* = \overline{A}^t = \begin{pmatrix} \overline{1} & \overline{i} \\ \overline{-i} & \overline{0} \end{pmatrix}^t = \begin{pmatrix} 1 & -i \\ i & 0 \end{pmatrix}^t = A.$$

Proposition C1.33 (Hermitesche Sesquilinearformen)
Sei V *ein* $\mathbb{C}$*-Vektorraum mit Basis* $B = (x_1, \ldots, x_n)$*,* $b \in \mathrm{Sesq}(V)$ *und* $A \in \mathrm{Mat}_n(\mathbb{C})$.

a. b *ist genau dann hermitesch, wenn* $M_B(b)$ *hermitesch ist.*

b. *Die Sesquilinearform* b_A^s *ist genau dann hermitesch, wenn* A *hermitesch ist.*

Beweis: Der Beweis geht analog zur entsprechenden Aussage für Bilinearformen, siehe Proposition C1.16. □

Bemerkung C1.34

a. Ist $A \in \mathrm{Mat}_n(\mathbb{R}) \subseteq \mathrm{Mat}_n(\mathbb{C})$, dann gilt

$$A^* = \overline{A}^t = A^t.$$

Insbesondere, A ist genau dann hermitesch, wenn A symmetrisch ist.

b. Ist b eine hermitesche Sesquilinearform auf dem $\mathbb{C}$-Vektorraum V, so gilt

$$b(x,x) = \overline{b(x,x)}$$

für alle $x \in V$. Das geht aber nur, wenn

$$b(x,x) \in \mathbb{R}$$

stets eine reelle Zahl ist! Man bezeichnet die Abbildung

$$q_b : V \to \mathbb{R} : x \mapsto b(x,x)$$

dann als die zu b gehörende *quadratische Form*, und man prüft leicht nach, daß für $x, y \in V$ stets

$$b(x,y) = \tfrac{1}{4}\big(q_b(x+y) - q_b(x-y) + iq_b(x+iy) - iq_b(x-iy)\big)$$

gilt, so daß die quadratische Form die Sesquilinearform b eindeutig bestimmt.

E) Definite Bilinearformen und Sesquilinearformen

Im folgenden beschränken wir uns auf $\mathbb{R}$- und $\mathbb{C}$-Vektorräume, da wir für die Werte $b(x,x)$ einer Bilinearform bzw. einer Sesquilinearform entscheiden müssen, ob sie positiv oder negativ sind.

Definition C1.35 (Definitheit)

a. Eine symmetrische Bilinearform $b \in \mathrm{Bil}_{\mathbb{R}}(V)$ heißt *positiv definit*, falls

$$b(x,x) > 0$$

für alle $0 \neq x \in V$ gilt. Sie heißt *negativ definit*, falls stattdessen

$$b(x,x) < 0$$

für alle $0 \neq x \in V$ gilt. Und sie heißt schließlich *indefinit*, falls es $x, y \in V$ gibt mit

$$b(x,x) > 0 > b(y,y).$$

b. Eine hermitesche Sesquilinearform $b \in \mathrm{Sesq}(V)$ heißt *positiv definit*, falls

$$b(x,x) > 0$$

für alle $0 \neq x \in V$ gilt. Sie heißt *negativ definit*, falls stattdessen

$$b(x,x) < 0$$

für alle $0 \neq x \in V$ gilt. Und sie heißt schließlich *indefinit*, falls es $x, y \in V$ gibt mit

$$b(x,x) > 0 > b(y,y).$$

Beispiel C1.36

a. Die hermitesche Sesquilinearform

$$b^s_{\mathbb{1}_n} : \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C}^n : (x, y) \mapsto \overline{x}^t \circ y = \sum_{i=1}^n \overline{x_i} \cdot y_i$$

ist positiv definit, da

$$b^s_{\mathbb{1}_n}(x, x) = \sum_{i=1}^n \overline{x_i} \cdot x_i = \sum_{i=1}^n |x_i|^2 > 0$$

für alle $0 \neq x = (x_1, \ldots, x_n) \in \mathbb{C}^n$.

b. Für

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \in \operatorname{Mat}_2(\mathbb{R})$$

ist die Bilinearform b_A auf $\mathbb{R}^2$ symmetrisch. Da für $x = (x_1, x_2)^t \in \mathbb{R}^2$ mit $x \neq (0, 0)$ ferner gilt

$$b_A(x, x) = x_1^2 + 2x_1x_2 + 4x_2^2 = (x_1 + x_2)^2 + 3x_2^2 > 0$$

ist b_A zudem positiv definit.

Bemerkung C1.37
Im allgemeinen ist die Bedingung der positiven Definitheit durchaus nicht einfach nachzuprüfen, da man meist nicht alle Vektoren $0 \neq x \in V$ überprüfen kann. Man beachte auch, daß es nicht reicht, etwa für eine Basis $B = (x_1, \ldots, x_n)$ von V nachzuprüfen, daß $b(x_i, x_i) > 0$ für alle $i = 1, \ldots, n$ gilt.

Betrachte dazu die folgende symmetrische Bilinearform auf $\mathbb{R}^2$

$$b : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R} : \big((a_1, a_2)^t, (b_1, b_2)^t\big) \mapsto a_1b_1 - a_2b_2$$

sowie die Basis $(x_1, x_2) = \big((1, 0)^t, (2, 1)^t\big)$. Dann gilt $b(x_1, x_1) = 1 > 0$ und $b(x_2, x_2) = 3 > 0$, aber $b(e_2, e_2) = -1 < 0$.

Wir werden in Satz 13.36 Kriterien kennenlernen, die es uns erlauben, positive Definitheit zu entscheiden. Diese formuliert man dann für symmetrische und hermitesche Matrizen (siehe auch Definition 13.33).

Aufgaben

Aufgabe C1.38
Es sei $b : \mathbb{K}^2 \times \mathbb{K}^2 \to \mathbb{K} : \big((x_1, x_2)^t, (y_1, y_2)^t\big) \mapsto 2 \cdot x_1 \cdot y_1 + x_1 \cdot y_2 + y_1 \cdot x_2 - x_2 \cdot y_2$. Ferner bezeichne $E = (e_1, e_2)$ die kanonische Basis des $\mathbb{K}^2$ und $B = (x_1, x_2)$ mit $x_1 = (1, 1)^t$ und $x_2 = (1, -1)^t$ sei eine weitere Basis.
Zeige, daß b eine symmetrische Bilinearform ist und berechne die Matrixdarstellungen $M_E(b)$ und $M_B(b)$ sowie die Transformationsmatrix T^B_E mit

$$M_B(b) = \big(T^B_E\big)^t \cdot M_E(b) \cdot T^B_E.$$

Aufgabe C1.39

Bestimme für die folgende symmetrische Matrix $A \in \mathrm{Mat}_4(\mathbb{R})$ eine Transformationsmatrix $T \in \mathrm{Gl}_4(\mathbb{R})$, so daß $T^t \circ A \circ T$ eine Diagonalmatrix wie im Sylvesterschen Trägheitssatz ist, und bestimme den Trägheitsindex, den Morseindex, die Signatur und den Rang von A:

$$A = \begin{pmatrix} -3 & 2 & 0 & 0 \\ 2 & 5 & 2 & 5 \\ 0 & 2 & 4 & 2 \\ 0 & 5 & 2 & 4 \end{pmatrix}.$$

Aufgabe C1.40

Sei $b \in \mathrm{Bil}_{\mathbb{K}}(V)$.

a. Zeige, es gibt eine symmetrische Bilinearform $b' \in \mathrm{Bil}_{\mathbb{K}}(V)$ und eine schiefsymmetrische Bilinearform $b'' \in \mathrm{Bil}_{\mathbb{K}}(V)$, so daß $b = b' + b''$.

b. Zeige, die Bilinearformen b' und b'' in a. sind eindeutig bestimmt.

Anmerkung, b heißt *schiefsymmetrisch*, wenn $b(x, y) = -b(y, x)$ für alle $x, y \in V$.

Aufgabe C1.41

Bestimme für die folgende symmetrische Matrix $A \in \mathrm{Mat}_5(\mathbb{R})$ eine Transformationsmatrix $T \in \mathrm{Gl}_5(\mathbb{R})$, so daß $T^t \circ A \circ T$ eine Diagonalmatrix wie im Sylvesterschen Trägheitssatz ist, und bestimme den Trägheitsindex, den Morseindex, die Signatur und den Rang von A:

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

Aufgabe C1.42

Zeige, sind $A, B \in \mathrm{Mat}_n(\mathbb{R})$ zwei symmetrische Matrizen mit $x^t A x \leq x^t B x$ für alle $x \in \mathbb{R}^n$, dann ist der Trägheitsindex von A kleiner oder gleich dem Trägheitsindex von B.

§ C2 Lineare Algebra mit Singular

Im vorliegenden Abschnitt wollen wir zeigen, wie ein Computeralgebrasystem eingesetzt werden kann, um Rechnungen in der linearen Algebra durchzuführen. Wir verwenden hierzu das am Fachbereich entwickelte System Singular. Es ist frei erhältlich für die Betriebssysteme Linux, Windows und MacOS von der Webseite:

`http://www.singular.uni-kl.de`

Auf den Linuxrechnern des Fachbereichs startet man Singular einfach durch den Befehl `Singular` von einer einfachen Textkonsole aus. Man erhält dann zunächst einige Informationen zum Programm sowie einen Eingabeprompt `>`:

```
                     SINGULAR                             /
 A Computer Algebra System for Polynomial Computations   /   version 3-1-1
                                                       0<
     by: G.-M. Greuel, G. Pfister, H. Schoenemann        \   Feb 2010
FB Mathematik der Universitaet, D-67653 Kaiserslautern    \
>
```

Der Eingabeprompt `>` fordert zur Eingabe von Singular-Befehlen auf. Wir wollen hier nur einige kurze Anmerkungen zur allgemeinen Syntax machen und hoffen, daß sich alles weitere aus den im folgenden besprochenen Beispielen erschließt. Unsere Konvention dabei ist, daß Singular-Ein- und Ausgaben im Gegensatz zu begleitenden Erläuterungen stets im `Typewriter`-Stil geschrieben werden.

a. Jede Singular-Sitzung sollte mit dem Befehl

   `ring r=0,t,lp;`

   beginnen. Dadurch wird der Polynomring $\mathbb{Q}[t]$ als Grundring festgelegt und erhält den Namen `r`. Selbst, wenn man nicht vor hat, Polynome zu verwenden, ist dies nötig, um mit den rationalen Zahlen rechnen zu können. Ersetzt man die Zahl `0` in der Definition von `r` durch eine Primzahl `p`, so verwendet man statt der rationalen Zahlen den Körper $\mathbb{Z}/p\mathbb{Z}$; ersetzt man sie durch `real` oder `complex`, so rechnet man mit reellen oder komplexen Dezimalzahlen, was aber tunlichst vermieden werden sollte, da dann Rundungsfehler auftreten können.

b. Man kann Ergebnisse von Rechnungen sowie Eingaben auch in Variablen speichern. Ein Beispiel dafür ist die Variable `r` in Teil a., in der der Polynomring $\mathbb{Q}[t]$ abgespeichert wurde. Jede Variable in Singular hat einen Namen und einen festgelegten Typen, der sagt, ob es sich um einen Ring (`ring`), ein Polynom (`poly`), ein Körperelement (`number`), eine ganze Zahl (`int`), eine Matrix (`matrix`) oder eine Liste (`list`) von Objekten handelt.

c. Nicht alle in Singular im Prinzip verfügbaren Befehle sind schon unmittelbar mit dem Programmstart geladen, viele liegen in sogenannten Bibliotheken vor.

Sie sind erst verfügbar, wenn man die entsprechende Bibliothek mit dem Befehl `LIB` eingebunden hat. Wie dies geschieht, werden wir in Beispielen sehen.

d. Jede Singular-Eingabe schließt mit einem Semikolon `;` und dem anschließenden Drücken der Return-Taste ab. Das Semikolon fordert den Singular-Interpreter dazu auf, die Eingabe zu übersetzen und auszuführen. Will man eine Eingabe über mehrere Zeilen strecken, so läßt man das Semikolon am Zeilenende weg und drückt die Return-Taste. Man erhält statt des üblichen Promptzeichens `>` dann einen Punkt `.` als Prompt. Dieser zeigt an, daß die Eingabe noch nicht beendet ist und sich über mehrere Zeilen erstreckt.

e. In den folgenden Beispielen ist alles, was auf einen Prompt `>` oder `.` am Zeilenanfang folgt, eine Eingabe, und jede Zeile, die ohne eines dieser Zeichen beginnt, enthält Singular-Ausgaben. Text, der auf `//` folgt, enthält Kommentare, die beim Ausführen des Kommandos nicht beachtet werden. Will man das Beispiel selbst in Singular nachprüfen, kann man sie getrost weglassen. Sie dienen nur der Erläuterung für den Leser. Ausgaben, die beim Laden von Bibliotheken auftreten, werden wir in den Beispielen weglassen.

f. Man beendet Singular mit dem Befehlt `exit`. Hilfe zur Syntax von Singular findet man im Manual auf der Singular-Webseite oder durch den Befehl `help`.

Beispiel C2.1 (Reduzierte Zeilen-Stufen-Form und Rang einer Matrix)
Wir wollen eine reduzierte Zeilen-Stufenform und damit den Rang der folgenden Matrix berechnen:

$$A = \begin{pmatrix} 1 & 2 & 1 & 3 & 1 \\ 2 & 4 & 7 & 3 & 0 \\ 4 & 8 & 9 & 9 & 2 \\ 3 & 6 & 0 & 2 & 1 \end{pmatrix} \in \mathrm{Mat}(4 \times 5, \mathbb{Q})$$

Dazu benutzen wir die Singular-Befehle `rowred`.

```
> LIB "matrix.lib";
> ring r=0,t,dp;
> matrix A[4][5]=1,2,1,3,1,2,4,7,3,0,4,8,9,9,2,3,6,0,2,1;
> print(rowred(A));  // zeigt die rZSF von A
1,2,0,0,1/11,
0,0,1,0,-2/11,
0,0,0,1,4/11,
0,0,0,0,0
```

Beispiel C2.2 (Kern einer Matrix)
Mit dem Befehl `syz` können wir eine Basis des Kerns der Matrix in Beispiel C2.1 berechnen. Man bezeichnet die Relationen zwischen den Spalten der Matrix, die durch die Vektoren im Kern beschrieben werden, auch als *Syzygien*, und `syz` ist die Abkürzung dieses Begriffs.

```
> LIB "matrix.lib";
> ring r=0,t,dp;
> matrix A[4][5]=1,2,1,3,1,2,4,7,3,0,4,8,9,9,2,3,6,0,2,1;
> matrix B=syz(A);
> print(B);
-2,0,
1, -1,
0, 4,
0, -8,
0, 22
```

Der Kern von A hat also die Basisvektoren $(-2,1,0,0,0)^t$ und $(0,-1,4,-8,22)^t$.

Beispiel C2.3 (Lösung eines linearen Gleichungssystems)
Wir setzen nun $b = (2,2,6,4)^t$ und wollen das lineare Gleichungssystem $Ax = b$ lösen. Der Befehl `concat` hängt zwei Matrizen hintereinander.

```
> matrix b[4][1]=2,2,6,4;
> matrix Ab=concat(A,b); // Bilde die erweiterte Koeffizientenmatrix.
> print(Ab);
1,2,1,3,1,2,
2,4,7,3,0,2,
4,8,9,9,2,6,
3,6,0,2,1,4
> print(syz(Ab));  // Berechne eine Basis des Kerns von Ab.
-2,0, 0,
1, -1,0,
0, 4, -2,
0, -8,4,
0, 22,-12,
0, 0, 1
```

Wir haben nun eine Basis des Kerns der erweiterten Koeffizientenmatrix berechnet. Der Algorithmus stellt sicher, daß es genau dann einen Vektor mit letzter Komponente ungleich null gibt, wenn das Gleichungssystem lösbar ist. Es gibt dann auch nur einen solchen Vektor und das ist der letzte Basisvektor. Dividiert man die ersten fünf Einträge des Vektors durch das Negative des letzten Eintrags, so erhält man eine spezielle Lösung, hier

$$c = (0,0,2,-4,12)^t.$$

Vergißt man bei den übrigen Vektoren in der berechneten Basis die letzte Komponente, so erhält man eine Basis des homogenen Lösungsraums, hier

$$\text{Lös}(A,0) = \text{Lin}\left((-2,1,0,0,0)^t, (0,-1,4,-8,22)^t\right),$$

wie wir schon aus Beispiel C2.2 wissen.

Beispiel C2.4 (Eigenwerte einer Matrix)
Wir wollen die Eigenwerte der folgenden Matrix bestimmen

$$A = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{Q}).$$

Dazu verwenden wir unter anderem die Singular-Befehle `det` zum Berechnen der Determinante, `unitmat` für die Einheitsmatrix und `factorize` zum Berechnen der Primfaktorzerlegung eines Polynoms.

```
> LIB "matrix.lib";
> ring r=0,t,dp;
> matrix A[3][3]=0,1,1,-1,2,1,-1,1,2;
> poly p=det(t*unitmat(3)-A);
> p;
t3-4t2+5t-2
> short=0;
> p;
t^3-4*t^2+5*t-2
> factorize(p);
[1]:
   _[1]=1
   _[2]=t-1
   _[3]=t-2
[2]:
   1,2,1
```

Das charakteristische Polynom von A ist also

$$\chi_A = t^3 - 4t^2 + 5t - 2 = (t-1)^2 \cdot (t-2),$$

so daß die Eigenwerte $\lambda = 1$ mit algebraischer Vielfachheit 2 sowie $\lambda = 2$ mit algebraischer Vielfachheit 1 sind.

Beispiel C2.5 (Minimalpolynom einer Matrix)
Als nächstes wollen wir das Minimalpolynom der Matrix A in Beispiel C2.4 berechnen. Dazu verwenden wir den Algorithmus 10.15 sowie einige Singular-Befehle. `transpose` transponiert eine Matrix, `flatten` schreibt die Einträge einer Matrix in einen Zeilenvektor und `power` potenziert eine Matrix.

```
> matrix C=transpose(flatten(power(A,0)));
> for (int i=1;i<=3;i++)
. {
.   C=concat(C,transpose(flatten(power(A,i))));
. }
```

```
> matrix D=syz(C);
> print(D);
2, 0,
-3,2,
1, -3,
0, 1
> poly mu;
> for (i=1;i<=4;i++){mu=mu+D[1][i]*t^(i-1);}
> mu;
t^2-3*t+2
> factorize(mu);
[1]:
   _[1]=1
   _[2]=t-1
   _[3]=t-2
[2]:
   1,1,1
```

Das Minimalpolynom von A ist also

$$\mu_A = t^2 - 3t + 2 = (t-1) \cdot (t-2),$$

und die Matrix A ist somit diagonalisierbar, da das Minimalpolynom in paarweise verschiedene Linearfaktoren zerfällt.

Die obige Befehlssequenz ist recht lang. Falls man bereits weiß, daß das Minimalpolynom über dem Grundkörper in Linearfaktoren zerfällt, so kann man auch den Singular-Befehl `minipoly` aus der Bibliothek `linalg.lib` verwenden, aber nur dann! Um sicherzustellen, daß das Minimalpolynom zerfällt, kann man zunächst das charakteristische Polynom berechnen und faktorisieren, denn nur wenn dieses zerfällt, zerfällt auch das Minimalpolynom. Für die Matrix A aus unserem Beispiel wissen wir bereits, daß es zerfällt. Wir können also den Befehl `minipoly` anwenden.

```
> LIB "linalg.lib";
> minipoly(A);
[1]:          // das Minimalpolynom hat die zwei Nullstellen 1 und 2
   _[1]=1
   _[2]=2
[2]:          // beide kommen mit Vielfachheit 1 vor
   1,1
```

Beispiel C2.6 (Diagonalisierung einer Matrix)
Wir haben in Beispiel C2.5 gesehen, daß die Matrix A aus Beispiel C2.4 diagonalisierbar ist. Nun wollen wir die zugehörige Transformationsmatrix T bestimmen.

Dazu erinnern wir uns, daß A genau die Eigenwerte 1 und 2 besitzt. Zu diesen müssen wir Basen der Eigenräume bestimmen.

```
> matrix T1=syz(A-unitmat(3));
> print(T1);
1,0,
1,-1,
0,1
> matrix T2=syz(A-2*unitmat(3));
> print(T2);
1,
1,
1
> matrix T=concat(T1,T2);
> print(T);
1,0, 1,
1,-1,1,
0,1, 1
> print(inverse(T)*A*T);
1,0,0,
0,1,0,
0,0,2
```

Beispiel C2.7 (Jordansche Normalform)
In diesem Beispiel wollen wir die Jordansche Normalform und die Transformationsmatrix T für die Matrix

$$A = \begin{pmatrix} 21 & 5 & 1 & 21 & 5 \\ -23 & 4 & 8 & -31 & 1 \\ -2 & -1 & -2 & -1 & -1 \\ -17 & -4 & -1 & -17 & -4 \\ 22 & -2 & -8 & 30 & 1 \end{pmatrix} \in \mathrm{Mat}_5(\mathbb{Q})$$

berechnen. Dazu bestimmen wir zunächst das charakteristische Polynom und das Minimalpolynom von A und faktorisieren diese.

```
> LIB "matrix.lib";
> LIB "linalg.lib";
> ring R=0,t,dp;
> matrix A[5][5]=21, 5, 1, 21, 5,
. -23,4, 8, -31,1,
. -2, -1,-2,-1, -1,
. -17,-4,-1,-17,-4,
. 22, -2,-8,30, 1;
```

```
> print(A);
21, 5, 1, 21, 5,
-23,4, 8, -31,1,
-2, -1,-2,-1, -1,
-17,-4,-1,-17,-4,
22, -2,-8,30, 1
> short=0;
> poly chi=det(t*unitmat(5)-A);
> chi;
t^5-7*t^4+10*t^3+18*t^2-27*t-27
> factorize(chi);
[1]:
   _[1]=1
   _[2]=t-3
   _[3]=t+1
[2]:
   1,3,2
> minipoly(A);
[1]:
   _[1]=-1
   _[2]=3
[2]:
   2,2
```

Wir sehen also, daß

$$\chi_A = (t-3)^3 \cdot (t+1)^2$$

und

$$\mu_A = (t-3)^2 \cdot (t+1)^2.$$

Damit ist die Jordansche Normalform von A festgelegt. Sie muß zu den beiden Eigenwerten 3 und -1 je mindestens einen Jordanblock der Größe 2 enthalten, weil sie im Minimalpolynom beide mit Vielfachheit zwei vorkommen. Zudem muß sie den Eigenwert 3 noch ein drittes Mal auf der Diagonalen haben, so daß ein weiterer Jordanblock der Größe eins zum Eigenwert 3 nötig ist. Also gilt

$$J_A = \left(\begin{array}{cc|c|cc} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ \hline 0 & 0 & 3 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{array}\right).$$

Da die Jordansche Normalform drei Jordanblöcke besitzt, müssen wir letztlich drei Basisvektoren finden, die die zyklischen A-invarianten Unterräume definieren, zu denen die Blöcke gehören. Dabei wenden wir den Algorithmus 11.15 an.

```
> matrix B=syz(A-3*unitmat(5)); // Basis von Ker(A-3*id)
> print(B);
-5,0,
1, -1,
1, 0,
4, 0,
0, 1
> matrix C=syz(power(A-3*unitmat(5),2)); // Basis von Ker((A-3*id)^2)
> print(C);
0,-5,0,
1,0, 0,
0,1, 0,
0,4, 0,
0,0, 1
```

Ein kurzer Blick genügt, um zu sehen, daß der erste und der dritte Basisvektor von $\operatorname{Ker}\big((A-3\cdot\mathbb{1}_5)^2\big)$ nicht im $\operatorname{Ker}(A-3\cdot\mathbb{1}_5)$ liegt. Wir können also jeden der beiden wählen, um den zyklischen Unterraum der Größe zwei zum Eigenwert 3 zu bilden. Wählen wir

$$x_1 = (0,1,0,0,0)^t.$$

```
> // berechne (A-3*unitmat(5)) * erste Spalte von C
. matrix X1[5][1]=C[1..5,1];
> print((A-3*unitmat(5))*X1);
5,
1,
-1,
-4,
-2
```

Damit hat der zyklische Unterraum der Größe zwei zum Eigenwert 3 die Basisvektoren

$$(A-3\cdot\mathbb{1}_5)x_1 = (5,1,-1,-4,-2)^t \quad \text{und} \quad x_1 = (0,1,0,0,0)^t,$$

und diese sind die ersten beiden Spalten der Matrix T.

Nun müssen wir noch den Vektor $(A-3\cdot\mathbb{1}_5)x_1$ zu einer Basis von $\operatorname{Ker}(A-3\cdot\mathbb{1}_5)$ ergänzen. Ein Blick auf die Basis von $\operatorname{Ker}(A-3\cdot\mathbb{1}_5)$ zeigt, daß jeder der beiden Vektoren es tut.

```
> matrix X2[5][1]=B[1..5,2];   // waehle X2
```

Wir wählen deshalb

$$x_2 = (0,-1,0,0,1)^t,$$

und dieser ist die dritte Spalte von T.

```
> print(syz(A+unitmat(5))); // Basis von Ker(A+id)
-1,
0,
1,
1,
0
> matrix D=syz(power(A+unitmat(5),2)); // Basis von Ker((A+id)^2)
> print(D);
-1,0,
0, -2,
1, 1,
1, 0,
0, 2
> matrix X3[5][1]=D[1..5,2];
> print((A+unitmat(5))*X3);   // (A+unitmat(5)) * 2. Spalte von D
1,
0,
-1,
-1,
0
```

Daraus folgt, daß der Vektor

$$x_3 = (0, -2, 1, 0, 2)^t \in \operatorname{Ker}\left(A + \mathbb{1}_5^2\right) \setminus \operatorname{Ker}(A + \mathbb{1}_5)$$

liegt, und daß die letzten beiden Spalten von T die Vektoren

$$(A + \mathbb{1}_5)x_3 = (1, 0, -1, 1-, 0)^t \quad \text{und} \quad x_3 = (0, -2, 1, 0, 2)^t$$

sind.

```
> // bestuecke die Matrix T
. matrix T=(A-3*unitmat(5))*X1;
> T=concat(T,X1);
> T=concat(T,X2);
> T=concat(T,(A+unitmat(5))*X3);
> T=concat(T,X3);
> print(T);
5, 0,0, 1, 0,
1, 1,-1,0, -2,
-1,0,0, -1,1,
-4,0,0, -1,0,
-2,0,1, 0, 2
> // invertiere die Matrix T
. matrix S=inverse(T);
```

```
> print(S);
1, 0,0, 1, 0,
1, 1,0, 1, 1,
8, 0,-2,10,1,
-4,0,0, -5,0,
-3,0,1, -4,0
> print(inverse(T)*A*T);
3,1,0,0, 0,
0,3,0,0, 0,
0,0,3,0, 0,
0,0,0,-1,1,
0,0,0,0, -1
```

Wir erhalten also

$$T = \begin{pmatrix} 5 & 0 & 0 & 1 & 0 \\ 1 & 1 & -1 & 0 & -2 \\ -1 & 0 & 0 & -1 & 1 \\ -4 & 0 & 0 & -1 & 0 \\ -2 & 0 & 1 & 0 & 2 \end{pmatrix} \quad \text{und} \quad T^{-1} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 8 & 0 & -2 & 10 & 1 \\ -4 & 0 & 0 & -5 & 0 \\ -3 & 0 & 1 & -4 & 0 \end{pmatrix}$$

sowie

$$J_A = T^{-1} \circ A \circ T = \left(\begin{array}{cc|c|cc} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ \hline 0 & 0 & 3 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{array}\right).$$

Es gibt in SINGULAR auch einen schnelleren Weg zur Jordanschen Normalform und der Transformationsmatrix, wenn man den Befehl `jordanbasis` verwendet, was in den SINGULAR-Übungsaufgaben aber nicht gemacht werden soll!

```
> matrix E=jordanbasis(A)[1];
> matrix Z[5][5];
> for (int j=1;j<=5;j++) { Z[1..5,j]=E[1..5,6-j]; }
> print(Z);
-5,5, 0,1, 0,
1, 1, 1,0, -2,
1, -1,0,-1,1,
4, -4,0,-1,0,
0, -2,0,0, 2
> print(inverse(Z)*A*Z);
3,0,0,0, 0,
0,3,1,0, 0,
0,0,3,0, 0,
0,0,0,-1,1,
0,0,0,0, -1
```

Die `for`-Schleife oben kehrt die Reihenfolge der Spalten in der Matrix E um. Die neue Matrix Z ist dann eine zulässige Transformationsmatrix T, wobei die Reihenfolge der Jordanblöcke sich geändert hat. Die Vertauschung der Spalten ist nötig, da die Konvention der Jordanschen Normalform in SINGULAR nicht mit unserer Konvention übereinstimmt. Darauf möchte ich hier aber nicht näher eingehen.

Beispiel C2.8 (Näherungsweise Bestimmung von Eigenwerten)
Eine zufällig ausgewählte Matrix in $\mathrm{Mat}_n(\mathbb{Q})$ wird keine rationalen Eigenwerte haben. Betrachen wir sie aber als Matrix in $\mathrm{Mat}_n(\mathbb{C})$, so zerfällt sie in Linearfaktoren und mit Wahrscheinlichkeit 1 sind diese paarweise verschieden. Exakt berechnen können wir sie aber nicht, da die Zerlegung eines Polynoms in $\mathbb{C}[t]$ in seine Primfaktoren im allgemeinen nicht möglich ist. Wir können die Eigenwerte aber näherungsweise berechnen, und dies reicht häufit aus, um zu sehen, daß sie paarweise verschieden sind.

```
> LIB "matrix.lib";
> ring S=complex,t,lp;
> matrix M[3][3];
> int i,j;
> for (i=1;i<=3;i++){for (j=1;j<=3;j++){M[i,j]=random(-9,9);}}
> print(M);
8, 5, 0,
-6,-2,3,
9, -9,7
> poly f=det(t*unitmat(3)-M);
> short=0;
> f;
t^3-13*t^2+83*t-449
> LIB "solve.lib";
> solve(f);
[1]:
   9.27119961
[2]:
   (1.86440019-i*6.70474155)
[3]:
   (1.86440019+i*6.70474155)
```

Das charakteristische Polynom der 3×3-Matrix ist ein Polynom vom Grad drei mit reellen Koeffizienten. Wegen des Zwischenwertsatzes muß es eine reelle Nullstelle haben. Wenn es keine weitere reelle Nullstelle besitzt, so müssen die übrigen beiden Nullstellen komplex konjugiert zueinander sein. Unsere Rechnung oben approximiert die Nullstellen mit dem Befehl `solve` aus der Bibliothek `solve.lib`, und wir sehen an den approximierten Nullstellen das geschilderte Phänomen.

Aufgaben

Aufgabe C2.9

Bestimme mit Hilfe von SINGULAR eine Basis B von $\mathbb{Q}^5$, bezüglich derer die Matrixdarstellung der Abbildung $f : \mathbb{Q}^5 \to \mathbb{Q}^5$ Jordansche Normalform hat, wo:

$$f(x_1, x_2, x_3, x_4, x_5) = (x_1 - x_2, x_1 + 2x_2 - x_3, -x_1 + 3x_3, -x_1 - 2x_2 - 2x_3 + 2x_4 - x_5, x_1 - x_3 + 2x_5)^t.$$

§ C3 Klassifikation der Kegelschnitte in der euklidischen Ebene

Man kann die Überlegungen in 13.42 verallgemeinern, was wir hier im Fall $n = 2$ tun wollen. Dazu betrachten wir die Lösungsmenge einer allgemeinen quadratischen Gleichung in zwei Unbekannten. Dies führt zur Klassifikation der Kegelschnitte in der euklidischen Ebene.

Definition C3.1
Es sei $\big(V, \langle\cdot,\cdot\rangle\big)$ ein euklidischer Raum.

a. Eine Abbildung $f : V \to V$ heißt eine *affine Abbildung* auf V, falls es ein $y \in V$ gibt und ein $g \in \mathrm{End}_{\mathbb{R}}(V)$ mit $f(x) = y + g(x)$ für alle $x \in V$.

b. Für $y \in V$ nennen wir die affine Abbildung

$$\tau_y : V \to V : x \mapsto x + y$$

die *Translation* um den Vektor y.

c. Eine Abbildung $f : V \to V$ heißt eine *Ähnlichkeit*, wenn es einen Vektor $y \in V$ gibt und eine orthogonale Abbildung $g \in \mathrm{O}(V)$ mit $f = \tau_y \circ g$, d. h.

$$f(x) = \tau_y\big(g(x)\big) = y + g(x) \quad \forall\, x \in V.$$

d. Ist $V = \mathbb{R}^n$ und sei $f = \tau_y \circ g$ mit $g \in \mathrm{End}_{\mathbb{R}}(V)$ eine bijektive affine Abbildung auf V, dann nennen wir die induzierte Abbildung

$$\mathbb{R}[t_1, \ldots, t_n] \to \mathbb{R}[t_1, \ldots, t_n] : p \mapsto p\big(f(t_1, \ldots, t_n)\big)$$

auf der Menge der Polynome in den Veränderlichen $t_1, \ldots, t_n$ einen *affinen Koordinatenwechsel* von $\mathbb{R}[t_1, \ldots, t_n]$.

Bemerkung C3.2

a. Jede affine Abbildung $f : \mathbb{R}^n \to \mathbb{R}^n$ läßt sich offenbar in eindeutiger Weise schreiben, als $f = \tau_y \circ g$ mit $y = f(0) \in V$ und $g \in \mathrm{End}_{\mathbb{R}}(V)$.

b. Ist $f = \tau_y \circ g : \mathbb{R}^n \to \mathbb{R}^n$ eine affine Abbildung mit $y \in \mathbb{R}^n$ und $g \in \mathrm{End}_{\mathbb{R}}(\mathbb{R}^n)$ bijektiv, dann gibt es eine eindeutig bestimmte Matrix $T \in \mathrm{Gl}_n(\mathbb{R})$ mit $g = f_T$. Damit gilt für $p \in \mathbb{R}[t_1, \ldots, t_n]$ und $t = (t_1, \ldots, t_n)^t$

$$p\big(f(t_1, \ldots, t_n)\big) = p(Tt + y).$$

Ist beispielsweise $p = t_1^2 + 3t_2 - 1 \in \mathbb{R}[t_1, t_2]$, $T = T\big(\frac{\pi}{2}\big)$ die Drehung um $90°$ und $y = (2, -2)$, dann ist für $f = \tau_y \circ f_T$

$$p\big(f(t_1, t_2)\big) = p(-t_2 + 2, t_1 - 2) = (-t_2 + 2)^2 + 3(t_1 - 2) - 1.$$

Definition C3.3
Es sei $p \in \mathbb{R}[t_1, \ldots, t_n]$ dann nennen wir die Menge

$$N(p) = \big\{\lambda = (\lambda_1, \ldots, \lambda_n)^t \in \mathbb{R}^n \;\big|\; p(\lambda) = 0\big\}$$

eine *algebraische Hyperfläche* von $\mathbb{R}^n$. Ist $\deg(p) = d$, so nennen wir d auch den *Grad* der Hyperfläche. Ist $n = 2$, so sprechen wir auch von *algebraischen Kurven* statt von algebraischen Hyperflächen.

Definition C3.4
Wir definieren auf $\mathbb{R}[t_1, \ldots, t_n]$ eine Relation durch

$$p \equiv q \quad :\Leftrightarrow \quad \exists\, c \in \mathbb{R} \setminus \{0\} : p = c \cdot q$$

für $p, q \in \mathbb{R}[t_1, \ldots, t_n]$. Wir nennen p und q mit $p \equiv q$ auch *äquivalent.*

Bemerkung C3.5
Man sieht sofort, daß $\equiv$ eine Äquivalenzrelation auf $\mathbb{R}[t_1, \ldots, t_n]$ definiert.

Ferner gilt offensichtlich, daß für zwei äquivalente Polynome $p, q \in \mathbb{R}[t_1, \ldots, t_n]$ auch $N(p) = N(q)$ gilt. Interessiert man sich also nur für das Nullstellengebilde von p, so kann man p getrost durch ein äquivalentes Polynom ersetzen und somit erreichen, daß der konstante Anteil von p entweder 0 oder -1 ist.

Im Folgenden interessieren wir uns nur noch für algebraische Kurven vom Grad zwei.

Bemerkung C3.6
Ist $p \in \mathbb{R}[t_1, t_2]$ ein allgemeines Polynom zweiten Grades, dann gibt es reelle Zahlen $\alpha_{11}, \alpha_{12} = \alpha_{21}, \alpha_{22}, \alpha_1, \alpha_2, \alpha \in \mathbb{R}$ so, daß

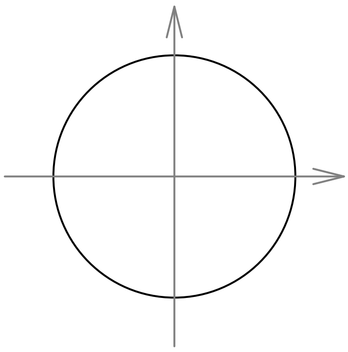
$$p = \alpha_{11} t_1^2 + 2\alpha_{12} t_1 t_2 + \alpha_{22} t_2^2 + \alpha_1 t_1 + \alpha_2 t_2 + \alpha = \langle t, St \rangle + \langle a, t \rangle + \alpha,$$

wobei $t = (t_1, t_2)^t$, $0 \neq S = (\alpha_{ij})_{i,j \in \{1,2\}} \in \mathrm{Mat}_2(\mathbb{R})$ und $a = (\alpha_1, \alpha_2)^t$.

Beispiel C3.7
Für $S = \mathbb{1}_2$, $a = (0, 0)^t$ und $\alpha = -1$ erhalten wir $p = t_1^2 + t_2^2 - 1$, und die Nullstellenmenge davon,

$$N\big(t_1^2 + t_2^2 - 1\big) = \big\{\lambda = (\lambda_1, \lambda_2)^t \in \mathbb{R}^2 \;\big|\; \lambda_1^2 + \lambda_2^2 = 1\big\},$$

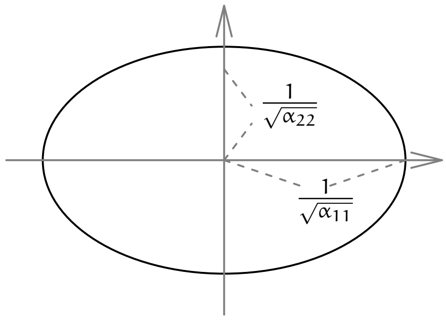
ist offenbar der Einheitskreis.



Ist S eine Diagonalmatrix mit positiven Diagonaleinträgen, d. h. $\alpha_{11}, \alpha_{22} > 0$ und $\alpha_{12} = \alpha_{21} = 0$, und ist ferner $a = (0, 0)^t$ und $\alpha = -1$, dann erhalten wir als Nullstellengebilde von p

$$N\Big(\big(\sqrt{\alpha_{11}}t_1\big)^2 + \big(\sqrt{\alpha_{22}}t_2\big)^2 - 1\Big) = \Big\{(\lambda_1, \lambda_2)^t \in \mathbb{R}^2 \;\Big|\; \big(\sqrt{\alpha_{11}}\lambda_1\big)^2 + \big(\sqrt{\alpha_{22}}\lambda_2\big)^2 = 1\Big\}$$

eine Ellipse.



Satz C3.8

Es sei

$$p = \langle t, St\rangle + \langle a, t\rangle + \alpha \in \mathbb{R}[t_1, t_2] \tag{75}$$

ein Polynom zweiten Grades mit symmetrischer Matrix $0 \neq S = (\alpha_{ij}) \in \mathrm{Mat}_2(\mathbb{R})$. *Dann gibt es eine affine Koordinatentransformation mittels einer Ähnlichkeit* $f = \tau_y \circ f_T$ *von* $\mathbb{R}^2$ *mit* $T \in \mathrm{SO}(2)$*, so daß* $q := p\big(f(t_1, t_2)\big)$ *äquivalent zu einer der folgenden Normalformen ist:*

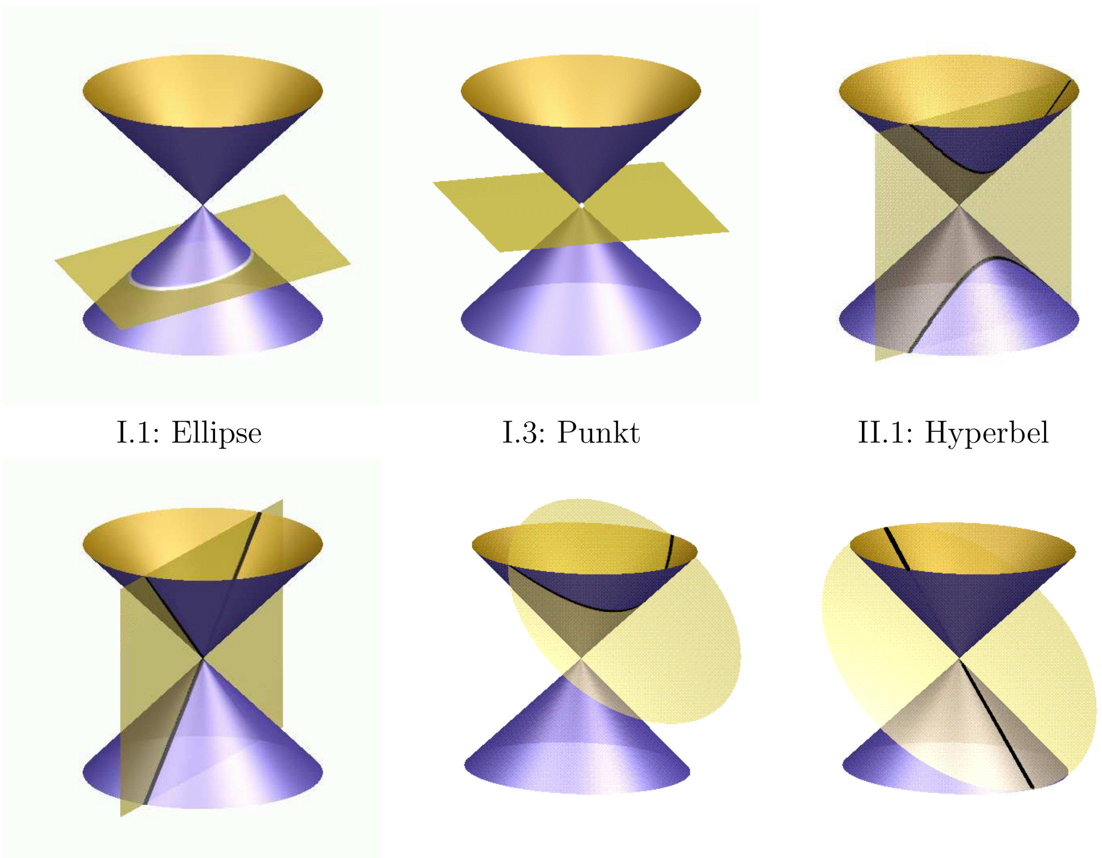
- I: $\underline{\det(S) > 0.}$
  - I.1: $\underline{\alpha \neq 0 \text{ und } \alpha_{11} > 0.}$ *Dann ist* $q \equiv (\lambda_1 t_1)^2 + (\lambda_2 t_2)^2 - 1$ *und* $N(q)$ *ist eine* Ellipse.
  - I.2: $\underline{\alpha \neq 0 \text{ und } \alpha_{11} < 0.}$ *Dann ist* $q \equiv (\lambda_1 t_1)^2 + (\lambda_2 t_2)^2 + 1$ *und* $N(q)$ *ist die leere Menge.*
  - I.3: $\underline{\alpha = 0.}$ *Dann ist* $q \equiv (\lambda_1 t_1)^2 + (\lambda_2 t_2)^2$ *und* $N(q)$ *ist ein Punkt.*
- II: $\underline{\det(S) < 0.}$
  - II.1: $\underline{\alpha \neq 0.}$ *Dann ist* $q \equiv (\lambda_1 t_1)^2 - (\lambda_2 t_2)^2 - 1$ *und* $N(q)$ *ist eine* Hyperbel.
  - II.2: $\underline{\alpha = 0.}$ *Dann ist* $q \equiv (\lambda_1 t_1)^2 - (\lambda_2 t_2)^2$ *und* $N(q)$ *besteht aus zwei verschiedenen Geraden durch den Ursprung.*
- III: $\underline{\det(S) = 0,\ a \neq (0,0)^t.}$ *Dann ist* $q \equiv t_1^2 - \lambda t_2$ *und* $N(q)$ *ist eine* Parabel.
- IV: $\underline{\det(S) = 0,\ a = (0,0)^t.}$
  - IV.1: $\underline{\alpha \neq 0 \text{ und } S \text{ hat einen positiven Eigenwert.}}$ *Dann ist* $q \equiv t_1^2 - \lambda$, $\lambda > 0$, *und* $N(q)$ *besteht aus zwei parallelen Geraden.*
  - IV.2: $\underline{\alpha \neq 0 \text{ und } S \text{ hat einen negativen Eigenwert.}}$ *Dann ist* $q \equiv t_1^2 + \lambda$, $\lambda > 0$, *und* $N(q)$ *ist die leere Menge.*
  - IV.3: $\underline{\alpha = 0.}$ *Dann ist* $q \equiv t_1^2$ *und* $N(q)$ *besteht aus einer* Doppelgraden, *d. h. einer Geraden, die man doppelt zählt.*

Bemerkung C3.9

Dies ist die vollständige Klassifikation der Kurven zweiten Grades. Sie heißen auch *Kegelschnitte*, da alle, bis auf die Fälle I.2, IV.1 und IV.2 als Schnitt des Kreiskegels

$$N\big(t_1^2 + t_2^2 - t_3^2\big) \subset \mathbb{R}^3$$

mit einer geeigneten Ebene im $\mathbb{R}^3$ realisierbar sind (siehe Abbildung 1).



I.1: Ellipse I.3: Punkt II.1: Hyperbel

II.2: Zwei Geraden mit Schnitt III: Parabel IV.3: Doppelgerade

Abbildung 1. Kegelschnitte

I.1 besagt, daß sich jede Ellipse durch Translation und Drehung so bewegen läßt, daß die Hauptachsen der Ellipse mit den Koordinatenachsen übereinstimmen. Daher kommt der Name Hauptachsentransformation.

Beweis von Satz C3.8:

1. Fall: $a = (0,0)^t$: Wir betrachten zunächst den Fall $a = (0,0)^t$.

Nach dem Satz über die Hauptachsentransformation 13.31 existiert ein $T \in SO(2)$, so daß

$$T^t \circ S \circ T = T^{-1} \circ S \circ T = \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix}.$$

Man beachte noch, daß nicht beide Eigenwerte μ_1 und μ_2 null sein können, da $S \neq 0$. Also können wir o. E. annehmen, daß $\mu_1 \neq 0$ und daß $\mu_1 \geq \mu_2$ gilt, falls $\mu_2 \neq 0$.

Die lineare Abbildung $f_T : \mathbb{R}^2 \to \mathbb{R}^2 : x \mapsto Tx$ ist eine Drehung und es gilt

$$\begin{aligned} p(Tt) &= \big\langle Tt, (S \circ T)t \big\rangle + \alpha \\ &= \big\langle t, \big(T^t \circ S \circ T\big)t \big\rangle + \alpha \\ &= \mu_1 t_1^2 + \mu_2 t_2^2 + \alpha. \end{aligned}$$

Da wir p ohnehin nur bis auf Äquivalenz klassifizieren wollen, können wir o. E. annehmen, daß $\alpha = 0$ oder $\alpha = -1$ gilt. Setzen wir nun noch $\lambda_i = \sqrt{|\mu_i|}$, dann erhalten wir folgende Fälle.

Fall 1.1: $\mu_1, \mu_2 > 0$: Dies ist gleichbedeutend dazu, daß S positiv definit ist, und nach dem Hauptminorenkriterium dazu, daß $\det(S) > 0$ und $\alpha_{11} > 0$. Ist $\alpha = -1$, so sind wir im Fall I.1, und ist $\alpha = 0$, so sind wir Fall I.3.

Fall 1.2: $\mu_1, \mu_2 < 0$: Dies ist gleichbedeutend dazu, daß $-S$ positiv definit ist, daß also $\det(S) = \det(-S) > 0$ und $-\alpha_{11} > 0$. Ist $\alpha = -1$, so sind wir im Fall I.2, und für $\alpha = 0$ wieder im Fall I.3, da wir dann das Polynom nochmals mit -1 multiplizieren können, um ein äquivalentes der gesuchten Form zu erhalten.

Fall 1.3: $\mu_1 > 0, \mu_2 < 0$: Dies ist gleichbedeutend dazu, daß $\mu_1 \cdot \mu_2 = \det(S) < 0$ ist. Im Fall $\alpha = -1$ führt dies zu Fall II.1, und im Fall $\alpha = 0$ führt es zu Fall II.2.

Fall 1.4: $\mu_1 > 0, \mu_2 = 0$ oder $\mu_1 < 0, \mu_2 = 0$: Das ist dann gleichbedeutend dazu, daß $\det(S) = 0$ ist. Für $\mu_1 > 0$ und $\alpha = -1$ erhalten wir Fall IV.1, für $\mu_1 < 0$ und $\alpha = -1$ den Fall IV.2, und für $\alpha = 0$ in den Fall IV.3.

2. Fall: $a \neq (0,0)^t$: Sind wir im Fall $a = (0,0)^t$ noch ohne Translation ausgekommen, so werden wir jetzt doch Translationen betrachten müssen.

Für $c \in \mathbb{R}^2$ bewirkt die Translation $\tau_c : \mathbb{R}^2 \to \mathbb{R}^2 : x \mapsto x + c$ folgende Koordinatentransformation für p

$$\begin{aligned} p(t+c) &= \langle t+c, St+Sc \rangle + 2\langle a, t+c \rangle + \alpha \\ &= \langle t, St \rangle + 2\langle a+Sc, t \rangle + \langle c, Sc \rangle + 2\langle a, c \rangle + \alpha \qquad (76) \\ &= \langle t, St \rangle + 2\langle b, t \rangle + \beta, \end{aligned}$$

wenn wir $b = a + Sc$ und $\beta = \langle c, Sc \rangle + 2\langle a, c \rangle + \alpha$ setzen.

Fall 2.1: $\exists\, c \in \mathbb{R}^2 : b = a + Sc = (0,0)^t$: Dann haben wir p durch $p\big(\tau_c(t)\big)$ auf den ersten Fall "$a = (0,0)^t$" zurückgeführt. Es gibt also ein $T \in$ SO(2), so daß $q = p\big((\tau_c \circ f_T)(t)\big)$ äquivalent zu einem der Fälle I, II oder IV ist.

Fall 2.2: $\forall\, c \in \mathbb{R}^2 : b = a + Sc \neq (0,0)^t$: Aus Lemma C3.10 folgt, daß es ein $c \in \mathbb{R}^2$ gibt mit $Sb = S^2c + Sa = 0$. Setzen wir nun noch $\delta := -\frac{\beta}{2\langle b,b \rangle}$, dann gilt für die

Translation $\tau_{c+\delta b}$[2]

$$\begin{aligned}
p(t+c+\delta b) =& \langle t, St\rangle + 2\big\langle a + S(c+\delta b), t\big\rangle + \big\langle c+\delta b, S(c+\delta b)\big\rangle + 2\langle a, c+\delta b\rangle + \alpha \\
=& \langle t, St\rangle + 2\langle b + \delta Sb, t\rangle + \delta^2\langle b, Sb\rangle + 2\delta\langle b, b\rangle + \beta \\
=& \langle t, St\rangle + 2\langle b, t\rangle + 2\delta\langle b, b\rangle + \beta \\
=& \langle t, St\rangle + 2\langle b, t\rangle.
\end{aligned}$$

Beachtet man, daß, wegen $Sb = 0$, Null auf alle Fälle ein Eigenwert von S ist und daß $S \neq 0$, so folgt aus dem Satz über Hauptachsentransformation 13.31 die Existenz eines $T \in \mathrm{SO}(2)$, so daß

$$D := T^t \circ S \circ T = T^{-1} \circ S \circ T = \begin{pmatrix} \mu_1 & 0 \\ 0 & 0 \end{pmatrix},$$

wobei $\mu_1 \neq 0$. Insbesondere sind wir also in dem Fall $\det(S) = 0$.

Ferner gilt für $T^t b =: (\mu, \lambda)^t$ unter Berücksichtigung, daß $T^t = T^{-1}$,

$$(\mu_1\mu, 0) = \big(T^t \circ S \circ T\big) \circ (T^t b) = T^t \circ (Sb) = 0,$$

und mithin ist $T^t b = (0, \lambda)^t$, wobei $\lambda \neq 0$, da T^t invertierbar und $b \neq (0,0)^t$. Aber dann überführt $t \mapsto Tt$ das Polynom $\langle t, St\rangle + 2\langle b, t\rangle$ in das Polynom

$$\big\langle Tt, (S \circ T)t\big\rangle + 2\big\langle b, Tt\big\rangle = \big\langle t^t, Dt\big\rangle + 2\big\langle T^t b, t\big\rangle = \mu_1 t_1^2 + 2\lambda t_2.$$

D. h. dann aber, daß

$$q := p\big((\tau_{c+\delta b} \circ f_T)(t)\big) = \mu_1 t_1^2 + 2\lambda t_2,$$

und damit sind wir genau im Fall III. □

Lemma C3.10

Ist $S \in \mathrm{Mat}_n(\mathbb{R})$ *symmetrisch, so gilt für die lineare Abbildung* $f_S : \mathbb{R}^n \to \mathbb{R}^n$.

a. $\mathrm{Ker}\left(f_S^2\right) = \mathrm{Ker}(f_S)$ *und* $\mathrm{Im}\left(f_S^2\right) = \mathrm{Im}(f_S)$.

b. *Zu jedem* $a \in \mathbb{R}^n$ *existiert ein* $c \in \mathbb{R}^n$, *so daß* $S^2c + Sa = 0$.

Beweis: a. Für $x \in \mathrm{Ker}\left(f_S^2\right)$ ergibt sich aus

$$0 = \big\langle x, S^2 x\big\rangle = \langle Sx, Sx\rangle,$$

also $f_S(x) = Sx = 0$ und $x \in \mathrm{Ker}(f_S)$. Die umgekehrte Inklusion ist klar. Wir wissen bereits, daß $\mathrm{Im}(f_S) \supseteq \mathrm{Im}\left(f_S^2\right)$ gilt. Da nun ferner

$$\begin{aligned}
\dim_{\mathbb{R}}\big(\mathrm{Im}(f_S)\big) &= n - \dim_{\mathbb{R}}\big(\mathrm{Ker}(f_S)\big) \\
&= n - \dim_{\mathbb{R}}\Big(\mathrm{Ker}\left(f_S^2\right)\Big) = \dim_{\mathbb{R}}\Big(\mathrm{Im}\left(f_S^2\right)\Big)
\end{aligned}$$

gilt, folgt also die Gleichheit.

[2]Man setze zunächst in der Gleichung (76) für c den Wert $c + \delta b$ ein. Dann ziehe man die Skalarprodukte auseinander und gruppiere sie neu, so daß man $b = a + Sc$, $Sb = 0$ sowie die Definition von β verwenden kann. Man beachte auch, daß S symmetrisch, also selbstadjungiert, ist.

b. Es gilt für $a \in \mathbb{R}^n$, daß $S(-a) = f_S(-a) \in \mathrm{Im}(f_S) = \mathrm{Im}\left(f_S^2\right)$, also gibt es nach a. ein $c \in \mathbb{R}^n$ mit $S^2c + Sa = f_S^2(c) - f_S(-a) = 0$.

$\square$

Literaturverzeichnis

[BF87] Martin Barner and Friedrich Flohr, *Analysis I*, 3. Auflage ed., Walter de Gruyter, 1987.

[Coh96] Henri Cohen, *A course in computational algebraic number theory*, 3 ed., Graduate Texts in Mathematics, no. 138, Springer, 1996.

[Dec10] Wolfram Decker, *Grundlagen der Mathematik I*, Vorlesungsskript, TU Kaiserslautern, 2010.

[Ebb92] Heinz-Dieter Ebbinghaus (ed.), *Zahlen*, 3 ed., Springer, 1992.

[Gat08] Andreas Gathmann, *Grundlagen der Mathematik*, Vorlesungsskript WS2007/08, TU Kaiserslautern, 2008.

[GK00] Gert-Martin Greuel and Thomas Keilen, *Lineare Algebra I & II*, Vorlesungsskript, FB Mathematik, Universität Kaiserslautern, 2000.

[Mar08] Thomas Markwig, *Algebraische Strukturen*, Vorlesungsskript, TU Kaiserslautern, 2008.

[Moo82] Gregory H. Moore, *Zermelo's axiom of choice: Its origins, development and influence*, Studies in the History of Mathematics and Physical Sciences, no. 8, Springer, 1982.

[Mül07] Rainer Müller, *Aufgaben zur vollständigen induktion*, http://www.emath.de, 2007.

[Sze50] Tibor Szele, *On Zorn's lemma*, Publicationes Mathematicae Debreccen **1** (1949/50), 254–57.