Skript zur Vorlesung

Mathematik für Ingenieure I

WS 2003/04

Prof. Dr. H. Kurzweil

Mathematisches Institut
Bismarckstr. 1 1/2
Zi. 204, Tel. 85 22505

6. Februar 2004

Organisation des Übungsbetriebs: Dr. S. Kräutle, Institut für Angewandte Mathematik, Martensstr. 3, Zi. 11.130, kraeutle@am.uni-erlangen.de

Literatur für die Vorlesung (u. a.)

- K. Kiyek, F. Schwarz: Mathematik für Informatiker 1 und 2. Teubner Verlag, 1991.
- W. Dörfler, W. Peschek: Einführung in die Mathematik für Informatiker. Hanser Verlag, 1988.
- G. Berendt: Mathematische Grundlagen für Informatiker. Band 1: Diskrete Mathematik. Band 2: Lineare Algebra, Analysis und Elemente der Wahrscheinlichkeitstheorie. BI Hochschultaschenbücher Band 641 und 642. Wissenschaftsverlag, 1990.
- K. Burg, H. Haf, F. Wille: Höhere Mathematik für Ingenieure. Band II: Lineare Algebra. Teubner Verlag, 1992.
- I. N. Bronstein, K. A. Semedjajew: Taschenbuch der Mathematik. Teubner, n-te Auflage, $n \geq 23$.
- K. Jänich: Lineare Algebra, Springer 1996

Teil 1: Lineare Algebra

Zur Notation

Es sei

$\mathbb{N}$ die Menge der natürlichen Zahlen $= \{1, 2, 3, \dots\}$

$\mathbb{N}_0 := \mathbb{N} \cup \{0\} = \{0, 1, 2, 3, \dots\}$

$\mathbb{Z}$ die Menge der ganzen Zahlen $= \{\dots, -2, -1, 0, 1, 2, \dots\}$

$\mathbb{Q}$ die Menge der rationalen Zahlen $= \{\frac{a}{b} \mid a, b \in \mathbb{Z},\ \ b \neq 0\}$

$\mathbb{R}$ die Menge der reellen Zahlen

Die Definition der reellen Zahlen liegt nicht auf der Hand; wir kommen später darauf zurück und begnügen uns zunächst mit der geometrischen Veranschaulichung von $\mathbb{R}$ als reelle *Zahlengerade* $\mathbb{R}^1$: Hier interpretieren wir jede reelle Zahl r als einen Punkt von $\mathbb{R}^1$, der rechts von 0 im Abstand r liegt, falls r positiv ist, und der links von 0 im Abstand $-r$ liegt, falls r negativ ist. Jedem Punkt von $\mathbb{R}^1$ entspricht genau eine reelle Zahl; z. B. entspricht $r = \sqrt{2}$ einem Punkt, der zwischen 1 und 2, genauer zwischen 1,4 und 1,5 liegt, usw. ...

Zunächst ist es für uns viel wichtiger, dass wir im $\mathbb{R}$ über die vier *Grundrechnungsarten* (*addieren, subtrahieren, multiplizieren, dividieren*) verfügen [1]. Insbesondere gilt: Seien $a, b \in \mathbb{R}$.

- Es existiert genau ein $x \in \mathbb{R}$ mit $a + x = b$, nämlich $x = b - a$.
- Ist $a \neq 0$, so existiert genau ein $x \in \mathbb{R}$ mit $ax = b$, nämlich $x = \frac{b}{a}$.

Da wir im Folgenden nur an den Grundrechnungsarten interessiert sind, und diese auch noch in anderen Zahlbereichen als $\mathbb{R}$, etwa in $\mathbb{Q}$, gelten, zu denen wir später kommen, sei stellvertretend für diese im Folgenden

$K := \mathbb{R}$

gesetzt.

[1] Diese gelten auch in $\mathbb{Q}$, zum Teil auch in $\mathbb{Z}, \mathbb{N}_0$ oder $\mathbb{N}$.

1 Lineare Gleichungssysteme

1. Seien $a, b \in K$ und L die Menge aller $x \in K$, für die gilt [2]

$(*1) \qquad ax = b.$

Dann gilt:

- Im Falle $a \neq 0$ ist $L = \{\frac{b}{a}\}$. Die Gleichung $(*1)$ besitzt also genau eine Lösung x.
- Im Falle $a = b = 0$ ist $L = K$. Jede Zahl $x \in K$ ist Lösung.
- Im Falle $a = 0, b \neq 0$ ist $L = \emptyset$. Es existiert keine Lösung.

2. Es sei K^2 die Menge aller Paare (x, y) von reellen Zahlen x, y.

$$K^2 := \{\ (x, y)\ \mid\ x, y \in K\}.$$

Dabei ist:

$$(x, y) = (x', y') \quad \Leftrightarrow \quad x = x' \quad \text{und} \quad y = y'\ .$$

Seien $a, b, c \in K$, und sei L die Menge aller Paare $(x, y) \in K^2$ für die gilt:

$(*2) \qquad ax + by = c\ .$

Ist $a = b = 0$, so ist $L = K^2$, wenn auch $c = 0$, und $L = \emptyset$, wenn $c \neq 0$.

Sei im Folgenden

$$(a, b) \neq (0, 0)\ [3]$$

- Ist $b \neq 0$ so liefert $(*2)$

  $y = -\frac{a}{b}x + \frac{c}{b}$, also

  $L = \{(x, y) \in K^2 \ \Big|\ y = -\frac{a}{b}x + \frac{c}{b}\} = \{(x, -\frac{a}{b}x + \frac{c}{b})\ \mid\ x \in K\}\ .$

- Ist $b = 0$ (also $a \neq 0$) so liefert $(*2)$

  $x = \frac{c}{a}$, also

  $L = \{(x, y) \in K^2\ \mid\ x = \frac{c}{a}\} = \{(\frac{c}{a}, y)\ \mid\ y \in K\}.$

[2] L ist also die Lösungsmenge der Gleichung $(*1)$

[3] es gelte also $a \neq 0$ oder $b \neq 0$

In beiden Fällen gibt es demnach mehrere Lösungen.

Am Beispiel der Gleichung $(*2)$ formulieren wir nun Zusammenhänge, die später in einem allgemeinen Rahmen auftreten.

Notation: Ein Element $(x_1, x_2) \in K^2$ schreiben wir als

$$\underline{x} := (x_1, x_2).$$

Seien $\underline{a}, \underline{b} \in K^2$, also $\underline{a} = (a_1, a_2),\ \underline{b} = (b_1, b_2)$

$$\begin{aligned} \underline{a} + \underline{b} &:= (a_1 + a_2,\ b_1 + b_2) \quad (\in K^2) \\ -\underline{b} &:= (-b_1,\ -b_2) \qquad\qquad (\in K^2) \\ \underline{a} - \underline{b} &:= \underline{a} + (-\underline{b}). \end{aligned}$$

Für $\lambda \in K$ sei

$$\begin{aligned} \lambda\underline{a} &:= (\lambda a_1, \lambda a_2) \qquad (\in K^2) \text{ *(skalare Multiplikation mit } \lambda)* \\ K\underline{a} &:= \{\lambda\underline{a} \mid \lambda \in K\} \end{aligned}$$

Zum Beispiel ist für $\lambda\underline{a} = -\underline{a}$ für $\lambda = -1$ und $\lambda\underline{a} = \underline{0}$ für $\lambda = 0$.

In $(*2)$ setzen wir nun $c = 0$ und bezeichnen die Lösungsmenge L mit U. Also ist

$$U = \{\underline{u} \in K^2 \mid au_1 + bu_2 = 0\}.$$

1.1 $\quad \underline{u} \in U,\ \lambda \in K \ \Rightarrow\ \lambda\underline{u} \in U$
$\quad\quad\ \underline{u}, \underline{v} \in U \ \Rightarrow\ \underline{u} + \underline{v} \in U$

Beweis: Multiplizieren wir die Gleichung $au_1 + bu_2 = 0$ mit $\lambda \in K$, so folgt $a(\lambda u_1) + b(\lambda u_2) = 0$. Mit $\underline{u}$ liegt also auch $\lambda\underline{u}$ in U. Seien $\underline{u}, \underline{v} \in U$. Dann gilt

$$0 = 0 + 0 = (au_1 + bu_2) + (av_1 + bv_2) = a(u_1 + v_1) + b(u_2 + v_2),$$

es folgt $\underline{u} + \underline{v} \in U$. □

1.2 *Ist* $\underline{0} \neq \underline{u} \in U$ *und* $(a, b) \neq 0$, *so ist* $U = K\underline{u}$.

Beweis: Sei $b \neq 0$. Dann ist $U = \left\{\left(x, -\frac{a}{b}x\right) \mid x \in K\right\}$, insbesondere hat $\underline{u}$ die Form $u = \left(u_1, -\frac{a}{b}u_1\right),\ u_1 \neq 0$. Mit $\lambda := \frac{x}{u_1}$ folgt

$$\lambda u = \big(\lambda u_1, -\tfrac{a}{b}\lambda u_1\big) = \big(x, -\tfrac{a}{b}x\big).$$

Mit $x \in K$ durchläuft auch λ ganz K; es folgt die Behauptung. Der Fall $a \neq 0$ ist analog. □

Notation: Sei $\underline{p} \in K^2$ und $U \subseteq K^2$.

$$U + \underline{p} := \{\underline{u} + \underline{p} \mid \underline{u} \in U\}.$$

Ist hier $U = \emptyset$, so sei $U + \underline{p} = \emptyset$.

Man beachte:

- $U = K^2 \Rightarrow U + \underline{p} = K^2$
- $U = \{(0,0)\} \Rightarrow U + \underline{p} = \{\underline{p}\}$
- Für $\underline{x} \in K^2$ gilt

$$\underline{x} \in U + \underline{p} \;\Leftrightarrow\; \exists\ \underline{u} \in U \quad \text{mit} \quad \underline{x} = \underline{u} + \underline{p} \;\Leftrightarrow\; \underline{x} - \underline{p} \in U.$$

1.3 *Seien $a, b, c, \in K$. Sei U die Lösungsmenge von $ax + by = 0$ und L die Lösungsmenge von $ax + by = c$. Sei $\underline{p} \in L$. Dann gilt $L = U + \underline{p}$.*

Beweis: $\underline{p} \in L$ bedeutet $ap_1 + ap_2 = c$. Für $\underline{x} = (x_1, x_2) \in K^2$ folgt

$$\begin{aligned}
\underline{x} \in L &\Leftrightarrow ax_1 + bx_2 = c \Leftrightarrow (ax_1 + bx_2) - (ap_1 + bp_2) = 0 \\
&\Leftrightarrow a(x_1 - p_1) + b(x_2 - p_2) = 0 \Leftrightarrow (x_1 - p_1, x_2 - p_2) \in U \\
&\Leftrightarrow \underline{x} - \underline{p} \in U \\
&\Leftrightarrow \underline{x} \in U + \underline{p}
\end{aligned}$$

Also ist $L = U + \underline{p}$. □

Beispiel:

$$x + 2y = 3 \quad (a = 1, b = 2, c = 3)$$

$\underline{u} = (1, -\frac{1}{2})$ ist eine Lösung von

$$x + 2y = 0.$$

Also ist

$$U \overset{1.2}{=} K(1, -\tfrac{1}{2}) = \{\lambda(1, -\tfrac{1}{2}) \mid \lambda \in K\} = \{(\lambda, -\tfrac{\lambda}{2}) \mid \lambda \in K\}.$$

Weiter ist $\underline{p} = (1,1)$ in L. Also ist

$$\begin{aligned} L &= U + p = K(1, -\tfrac{1}{2}) + (1,1) \\ &= \{\lambda, -\tfrac{\lambda}{2}) + (1,1) \mid \lambda \in K\} \\ &= (\lambda + 1, -\tfrac{\lambda}{2} + 1) \mid \lambda \in K\}. \end{aligned}$$

Geometrische Interpretation: Analog wie die Vorstellung von K als Zahlengerade K^1 stellen wir uns K^2 als (Tafel-) Ebene vor mit K^1 als x-Achse und K^1 als y-Achse. Die Elemente $\underline{p}$ von K^2 interpretieren wir dann als Punkte der Ebene. Der Punkt $\underline{p}$ hat die x-Koordinate p_1 und die y-Koordinate p_2. Für $u \in K^2, u \neq (0,0)$ ist $K\underline{u}$ die Gerade, die durch die Punkte $(0,0)$ und $\underline{u}$ geht. [4] Ist weiter $\underline{p} \in K^2$ so ist $K\underline{u} + \underline{p}$ die Gerade, die durch $\underline{p}$ geht, und parallel zu der Geraden $K\underline{u}$ ist.

Merke 1: Ist L die Lösungsmenge von $(*2)$, so gilt einer der Fälle

- $L = \emptyset$
- L ist Gerade in der Ebene K^2
- $L = K^2$

Unabhängig von dieser geometrischen Interpretation, also unabhängig von der Anschauung *definieren* wir: Eine Teilmenge $G \subseteq K^2$ heißt *Gerade*, wenn G die Lösungsmenge L einer Gleichung $(*2)$ mit $(a,b) \neq (0,0)$ ist.

Eine Gerade G hat also die Form

$$G = K\underline{u} + \underline{p} \quad , \quad \underline{u} \neq \underline{0}.$$

Mit Hilfe dieser abstrakten Definition zeigen wir:

1.4 *Seien $\underline{u}, \underline{v}$ Punkte im K^2 und $\underline{v}$ liege nicht auf der Geraden $K\underline{u}$. Dann gilt*

$$(K\underline{u} + \underline{v}) \cap (K\underline{v} + \underline{u}) = \{\underline{u} + \underline{v}\}.$$

Beweis: Es ist $\underline{u}+\underline{v} \in (K\underline{u}+\underline{v}) \cap (K\underline{v}+\underline{u})$. Sei $\underline{d} \in (K\underline{u}+\underline{v}) \cap (K\underline{v}+\underline{u})$. Dann existieren Zahlen $\lambda, \mu \in K$ mit

$$\underline{d} = \mu\underline{u} + \underline{v} = \lambda\underline{v} + \underline{u}.$$

Es folgt

$$(\mu-1)\underline{u} = \mu\underline{u} - \underline{u} = \lambda\underline{v} - \underline{v} = (\lambda-1)\underline{v}.$$

Damit liegt $(\lambda-1)\underline{v}$ auf der Geraden $K\underline{u}$. Nach Voraussetzung sind die Geraden $K\underline{u}$ und $K\underline{v}$ verschieden, also

$$(\lambda-1)\underline{v} \in K\underline{u} \cap K\underline{v} = \{(0,0)\}.$$

Wegen $\underline{v} \neq 0$ folgt $\lambda - 1 = 0$, also $\lambda = 1$, d.h. $\underline{d} = \underline{u} + \underline{v}$. □

3. Seien $a, b, c, d, e, f, \in K$ und L die Menge aller $(x, y) \in K^2$, für die gilt

$$(*3) \qquad \begin{aligned} ax + by &= c \\ dx + ey &= f \end{aligned}$$

Sei L_1 die Lösungsmenge von $ax + by = c$ und L_2 die von $dx + ey = f$. Dann ist

$$L = L_1 \cap L_2,$$

und es gilt einer der folgenden Fälle:

- $L_1 = L_2 = K^2$, also $L = K^2$.
- $L_1 = \emptyset$ oder $L_2 = \emptyset$, also $L = \emptyset$.
- L_1 ist Gerade, $L_2 = K^2$, also $L = L_1$ ist Gerade.
- L_2 ist Gerade, $L_1 = K^2$, also $L = L_2$ ist Gerade.
- L_1 und L_2 sind Geraden. Hier gibt es drei Unterfälle:
  - $L_1 = L_2$, also $L = L_1$ ist Gerade.
  - $L_1 \neq L_2$ und $L = \emptyset$, also sind L_1 und L_2 *parallele* Geraden.
  - $L_1 \neq L_2$ und $L_1 \cap L_2 \neq \emptyset$. Dann ist $L_1 \cap L_2 = \{\underline{p}\}$, und $\underline{p}$ ist der Schnittpunkt der Geraden L_1, L_2.

$(*3)$ besitzt nur im letzten Fall genau eine Lösung.

4. Es sei K^3 die Menge aller Tripel (x, y, z) von Zahlen aus K. Dabei gelte

$$(x, y, z) = (x', y', z') \quad \Leftrightarrow \quad x = x',\ y = y' \quad , \quad z = z'.$$

Seien $a, b, c, d \in K$. Wir bezeichnen nun mit L die Menge aller $(x, y, z) \in K^3$, für die gilt:

$(*)$ $\qquad ax + by + cz = d.$

Im Falle $(a, b, c) = (0, 0, 0)$ ist $L = K^3$, wenn auch $d = 0$, und $L = \emptyset$, wenn $d \neq 0$. Sei

$$(a, b, c) \quad \neq \quad (0, 0, 0).$$

Die Menge K^3 stellen wir uns als 3-dimensionalen Raum vor, und ein Element

$$\underline{p} = (p_1, p_2, p_3) \in K^3$$

als einen Punkt mit der x-Koordinate p_1, der y-Koordinate p_2 und der z-Koordinate p_2 .

Wir machen plausibel, dass die Lösungsmenge L von $(*4)$ eine Ebene im K^3 ist: [5] Ist $a = b = 0$, so hat man die Gleichung

$$cz = d$$

und L ist die Ebene parallel zur x, y-Ebene, die durch den Punkt $(0, 0, \frac{d}{z})$ geht. Analog beschreibt die Gleichung

$$ax = d \quad \text{bzw.} \quad by = d$$

eine Ebene, die parallel zur y, z-Ebene bzw. x, z-Ebene ist und durch den Punkt $(\frac{d}{a}, 0, 0)$ bzw. $(0, \frac{d}{b}, 0)$ geht.

Wir können also annehmen, dass mindestens 2 Koeffizienten a, b, c verschieden von 0 sind. Seien etwa $a \neq 0$ und $b \neq 0$. Dann schneidet L die x, y-Ebene in der Geraden

$$ax + by = d$$

und die y, z-Ebene in der Geraden

[5] Wir können dies hier nicht beweisen, weil wir den Begriff der Ebene nicht definiert haben. Um diese Schwierigkeit zu vermeiden, liegt es nahe – wie bei den Geraden – zu definieren: Eine Teilmenge L von K^3 heißt *Ebene*, falls L die Lösungsmenge einer Gleichung $(*3)$ mit $(a, b, c) \neq (0, 0, 0)$ ist

$$by + cz = d.$$

L ist nun die eindeutig bestimmte Ebene im K^3, in der diese beiden Geraden liegen.

5. Seien $a, b, c, d, e, f, g, h, i, j, k, l \in K$ und sei L die Menge aller $(x, y, z) \in K^3$, für die gilt:

$$\begin{aligned} & ax + by + cz = d \\ (*5) \quad & ex + fy + gz = h \\ & ix + jy + kz = l. \end{aligned}$$

Sind L_1 bzw. L_2 bzw. L_3 die Lösungsmenge der 1. bzw. 2. bzw. 3. Gleichung, so ist

$$L = L_1 \cap L_2 \cap L_3.$$

Wir wissen, dass jede Lösungsmenge L_1, L_2 oder L_3 entweder die leere Menge oder eine Ebene oder ganz K^3 ist. Der Durchschnitt zweier verschiedener Ebenen ist eine Gerade (oder leer) und der Durchschnitt zweier verschiedener Geraden ist ein Punkt (oder leer).

Merke 2: Die Lösungsmenge L von $(*5)$ ist entweder leer oder K^3 oder eine Ebene oder eine Gerade oder ein Punkt.

Nur im letzten Fall besitzt $(*5)$ genau eine Lösung.

Definitionen und Bezeichnungen:

Für $n \in \mathbb{N}$ sei K^n die Menge aller n-Tupel *(Vektoren)*

$$\underline{a} := (a_1, \ldots, a_n) \quad , \quad a_1, \ldots, a_n \in K.$$

Dabei ist

$$\underline{a} = \underline{b} \quad \Leftrightarrow \quad a_i = b_i \quad \text{für} \quad i = 1, \ldots, n.$$

Also ist $\underline{0} := (0, \ldots, 0) \in K^n$; für $i = 1, \ldots, n$ sei

$$e_i := (0, \ldots, 0, \underset{i}{1}, \ldots 0) \in K^n$$

Definition der *Addition* in K^n:

$$\begin{aligned} \underline{a}+\underline{b} &:= (a_1+b_1,\ldots,a_n+b_n) \\ -\underline{a} &:= (-a_1,\ldots,-a_n) \\ \underline{a}-\underline{b} &:= \underline{a}+(-\underline{b}). \end{aligned}$$

Definition der *skalaren Multiplikation in* K^n:

Für $\lambda \in K, \quad \underline{a} \in K^n$ sei

$$\lambda\underline{a} := (\lambda a_1,\ldots,\lambda a_n).$$

Z.B. ist

$$\underline{a} = a_1e_1+\cdots+a_ne_n = \textstyle\sum_{i=1}^n a_ie_i.$$

Seien $n,m \in \mathbb{N}$ und seien $a_{ij},\ b_i$ Zahlen aus K für $i = \{1,\ldots,m\}$ und $j \in \{1,\ldots,n\}$. Es sei L die Menge aller $\underline{x} \in K^n$, sodass gilt

$$(*) \qquad \begin{array}{ccccc} a_{11}x_1 & +\cdots+ & a_{1n}x_n & = & b_1 \\ \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & +\cdots+ & a_{mn}x_n & = & b_n \end{array}$$

oder kurz

$$(*) \quad \textstyle\sum_{j=1}^n a_{ij}x_j = b_i\ ,\ i = 1,\ldots,m.$$

Das System $(*)$ heißt *lineares Gleichungssystem* über K, es besteht aus m linearen Gleichungen in n Variablen. $(*)$ heißt *homogen*, wenn $b_1 = \cdots = b_m = 0$; andernfalls ist es *inhomogen*.

Die Zahlen a_{ij}, b_i definieren Vektoren aus K^m bzw. K^n

$$\begin{aligned} a_{*j} &:= (a_{1j},\ldots,a_{mj}) \in K^m, \quad j = 1,\ldots,n \qquad (Spaltenvektoren) \\ a_{i*} &:= (a_{i1},\ldots,a_{in}) \in K^n, \quad i = 1,\ldots,m \qquad (Zeilenvektoren) \\ \underline{b} &:= (b_1,\ldots,b_m) \in K^m, \end{aligned}$$

und die *Koeffizientenmatrix*

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

sowie die *erweiterte Koeffizientenmatrix*

$$(A, \underline{b}) := \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix}.$$

Dabei sind $a_{*1}, \dots, a_{*n}$ die *Spalten* und $a_{1*}, \dots, a_{m*}$ die *Zeilen* von A. [6] Verwendet man die Spalten von A, so wird $(*)$ eine Gleichung in K^m:

$$(*) \quad \begin{array}{l} x_1 a_{*1} + x_2 a_{*2} + \dots + x_{n-1} a_{*n-1} + x_n a_{*n} = \underline{b}, \quad \text{oder} \\ \sum_{j=1}^{n} x_j a_{*j} = \underline{b}. \end{array}$$

Zu folgenden zwei Aussagen vergleiche man 1.1.

1.5 *Sei $\underline{b} = \underline{0}$ in $(*)$. Dann gilt 1.1 für $U := L$.*

Beweis: Multiplizieren wir die Gleichung $\sum_{j=1}^{n} u_j a_{*j} = 0$ mit $\lambda \in K$, so folgt $\sum_j (\lambda u_j) a_{*j} = 0$. Mit $\underline{u}$ liegt also auch $\lambda \underline{u}$ in U. Seien $\underline{u}, \underline{v} \in U$. Dann gilt

$$0 \;=\; 0 + 0 \;=\; \left(\sum\nolimits_j u_j a_{*j}\right) + \left(\sum\nolimits_j v_j a_{*j}\right) \;\overset{!}{=}\; \sum\nolimits_j (u_j a_{*j} + v_j a_{*j}) \;\overset{!}{=}$$

$\sum_j (u_j + v_j) a_{*j}$.

Es folgt $\underline{u} + \underline{v} \in U$. □

[6] z. B. ist $\underline{b}$ die letzte Spalte von $(A, \underline{b})$

1.6 *Sei $\underline{p} \in K^n$ eine Lösung von $(*)$, und U die Menge aller $\underline{x} \in K^n$, die das zu $(*)$ gehörende homogene Gleichungssystsem lösen (setze $b_1 = \cdots = b_m = 0$). Dann gilt*

$$L = U + \underline{p} := \{\underline{u} + \underline{p} \mid \underline{u} \in U\}.$$

Beweis: $\underline{p} \in L$ bedeute $\sum_{j=1}^{n} p_j a_{*j} = \underline{b}$. Für $\underline{x} \in K^n$ folgt

$$\begin{aligned}
\underline{x} \in L \quad &\Leftrightarrow \quad \sum_{j=1}^{n} x_j a_{*j} = \underline{b} \quad \Leftrightarrow \quad \sum_{j=1}^{n} x_j a_{*j} - \sum_{j=1}^{n} p_j a_{*j} = 0 \\
&\Leftrightarrow \quad \sum_{j=1}^{n} (x_j - p_j) a_{*j} = \underline{0} \\
&\Leftrightarrow \quad \underline{x} - \underline{p} \in U \\
&\Leftrightarrow \quad \underline{x} \in U + \underline{p}.
\end{aligned}$$

□

1.7 *Ein lineares Gleichungssystem $(*)$ geht durch Umformungen der Form*

- *Addiere das Vielfache einer Gleichung zu einer anderen.*
- *Multipliziere eine Gleichung mit einem Skalar $r \in K, \quad r \neq 0$.*
- *Vertausche zwei Gleichungen*

in ein lineares Gleichungssystem über, das **dieselbe** *Lösungsmenge wie $(*)$ besitzt.* □

Da man dies – mehr oder weniger – direkt einsehen kann, verzichten wir hier auf einen Beweis, kommen allerdings später in einem allgemeineren Kontext darauf zurück.

Notation: Wir bezeichnen mit $K^{m \times n}$ die Menge aller Matrizen mit Einträgen aus K, welche n Spalten und m Zeilen besitzen (also die Menge aller $m \times n$-Matrizen über K). Speziell sei

$$\mathbf{0} := \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix} \in K^{m \times n}$$

die *Nullmatrix* und

$$\mathbf{1} := \mathbf{1}_n := \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \in K^{n\times n}$$

die *Einheitsmatrix*.

Definition: Sei $A \in K^{m\times n}$. *Zeilenumformungen* von A sind:

- Z_1 Addiere ein Vielfaches einer Zeile zu einer anderen.
- Z_2 Multipliziere eine Zeile mit einem Skalar $\lambda \in K,\ 0 \neq \lambda$.
- Z_3 Vertausche zwei Zeilen von A.

Definition: Eine Matrix $A \in K^{m\times n}$ heißt **Stufenmatrix**, wenn $A = \mathbf{0}$ oder Folgendes gilt:

- Es gibt eine Zahl $r \in \{1, \ldots, m\}$, sodass $a_{i*} \neq \underline{0}$ für $i = 1, \ldots, r$ ist. Im Fall $r < m$ ist $a_{i*} = 0$ für $i = r+1, \ldots, m$.
- Ist $a_{i*} \neq \underline{0}$, so ist der erste von 0 verschiedene Eintrag von a_{i*} (*Pivot* der Zeile a_{i*}) eine 1.
- Alle Einträge in der Spalte oberhalb eines Pivot sind 0.
- Für $i = 2, \ldots, r$ gilt: Der Pivot der i-ten Zeile steht rechts vom Pivot der $(i-1)$-ten Zeile.

1.8 *Sei $A \in K^{m\times n}$. Man kann A durch Zeilenumformungen Z_1, Z_2 und Z_3 auf Stufenform bringen.*

Beispiel: $\begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \overset{Z_3}{\to} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix} \overset{Z_1}{\longrightarrow} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix} \overset{Z_1,Z_1}{\longrightarrow} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$.

Die **Gaußelimination** ist ein Verfahren *(Algorithmus)*, um A in Stufenform zu bringen. Wir erklären sie, indem wir ein kleines *Programm* andeuten, und beweisen damit 2.4.

Ausgangspunkt sind die Zeilenvektoren $a_{i*},\ i = 1, \ldots, m$ und Spaltenvektoren

$a_{*j},\ j = 1, \ldots, n$. Die Gaußelimination verändert sie nach folgendem Rezept:

$k := 1$

for $j = 1, \ldots, n$ repeat

if $(a_{kj}, \ldots, a_{mj}) = \underline{0}$ then $j := j + 1$

if $(a_{kj}, \ldots, a_{mj}) \neq \underline{0}$ bestimme ein $i \in \{k, \ldots, m\}$ mit $a_{ij} \neq 0$

und setze sukzessive

$$a_{k*} := a_{i*},\ a_{i*} := a_{k*} \qquad (Z_3)$$

$$\lambda := a_{kj}$$

$$a_{k*} := \tfrac{1}{\lambda} a_{k*} \qquad (Z_2)$$

für $i = 1, \ldots, k-1, k+1, \ldots, m$ do

$$\mu := -a_{ij}$$

$$a_{i*} := \mu a_{k*} + a_{i*} \qquad (Z_1)$$

$k := k + 1$

until $k \geq m$ oder $k \geq n$.

In den folgenden Aussagen 1.9 und 1.10 setzen wir voraus:

- $(A, \underline{b})$ ist die erweiterte Koeffizientenmatrix des linearen Gleichungssystems $(*)$
- L ist die Lösungsmenge von $(*)$
- A hat **Stufenform**
- r ist die Anzahl der von $\underline{0}$ verschiedenen Zeilen von A und $d := n - r$.

Man mache sich klar, dass d die Anzahl der Spalten von A ist, in denen kein Pivot vorkommt.

1.9 *Genau dann besitzt $(*)$ eine Lösung (also $L \neq \emptyset$), wenn*

$$b_{r+1} = \cdots = b_m = 0. \qquad \square$$

1.10 *Sei $b_{r+1} = \cdots = b_m = 0$.*

(a) *Ist* $d = 0$, *so ist* $n = r$,

$$A = \left(\begin{array}{c|ccc} & \mathbf{1}_n & \\ \hline 0 & \dots & 0 \\ \vdots & & \\ 0 & \dots & 0 \end{array}\right)$$

und $L = \{(b_1, \dots, b_n)\}$.

(b) *Sei* $d \neq 0$ *und seien* $a_{*j_1}, \dots, a_{*j_d}$ *die Spalten von* A, *in denen kein Pivot vorkommt. Dann existiert zu jeder Wahl von* $(y_1, \dots, y_d) \in K^d$ *genau ein* $\underline{x} \in L$, *sodass gilt:*

$$x_{i_k} = y_k \quad \textit{für} \quad k = 1, \dots, d. \qquad \square$$

Man nennt $x_{i_1}, \dots, x_{i_d}$ *freie Variable* und d die "Dimension" des "Lösungsraums" L (oder Freiheitsgrad), und interpretiert L "geometrisch":

$$\begin{array}{ll} d = 0: & L \text{ ist Punkt} \\ d = 1: & L \text{ ist Gerade} \\ d = 2: & L \text{ ist Ebene} \\ \vdots & \vdots \end{array}$$

REZEPT *für das Lösen des linearen Gleichungssystems* $(*)$:

Forme die erweiterte Koeffizientenmatrix $(A, \underline{b})$ durch Zeilenumformumgen Z_1, Z_2, Z_3 in eine Matrix $(A', \underline{b}')$ um, sodass A' eine Stufenmatrix ist. Nach 1.7 sind die Lösungen von $(*)$ genau die Lösungen des zu dem $(A', \underline{b}')$ gehörenden linearen Gleichungssystems. Letztere beschreibt 1.9 und 1.10.

Beispiel:

$$A := \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} \in K^{3\times 3},\ \underline{b} \in K^3 \quad \text{beliebig}$$

$$\begin{pmatrix} 0 & 1 & 2 & b_1 \\ 1 & 2 & 1 & b_2 \\ 1 & 1 & 1 & b_3 \end{pmatrix} \xrightarrow{Z_3} \begin{pmatrix} 1 & 2 & 1 & b_2 \\ 0 & 1 & 2 & b_1 \\ 1 & 1 & 1 & b_3 \end{pmatrix} \xrightarrow{Z_1} \begin{pmatrix} 1 & 2 & 1 & b_2 \\ 0 & 1 & 2 & b_1 \\ 0 & -1 & 0 & b_3 - b_2 \end{pmatrix}$$

$$\xrightarrow{Z_1} \begin{pmatrix} 1 & 2 & 1 & b_2 \\ 0 & 1 & 2 & b_1 \\ 0 & 0 & 2 & b_3 - b_2 + b_1 \end{pmatrix}$$

$$\xrightarrow{Z_1} \begin{pmatrix} 1 & 0 & -3 & b_2 - 2b_1 \\ 0 & 1 & 2 & b_1 \\ 0 & 0 & 2 & b_3 - b_2 + b_1 \end{pmatrix}$$

$$\xrightarrow{Z_2} \begin{pmatrix} 1 & 0 & -3 & b_2 - 2b_1 \\ 0 & 1 & 2 & b_1 \\ 0 & 0 & 1 & \frac{1}{2}(b_3 - b_2 + b_1) \end{pmatrix}$$

$$\xrightarrow{Z_1, Z_1} \begin{pmatrix} 1 & 0 & 0 & b_2 - 2b_1 + \frac{3}{2}(b_3 - b_2 + b_1) \\ 0 & 1 & 0 & b_1 - (b_3 - b_2 - b_1) \\ 0 & 0 & 1 & \frac{1}{2}(b_3 - b_2 + b_1) \end{pmatrix}$$

Wir haben folgendes Gleichungssystem gewonnen:

$$\begin{aligned} x_1 \qquad\quad &= -\tfrac{1}{2}b_1 - \tfrac{1}{2}b_2 + \tfrac{3}{2}b_3 \\ x_2 \qquad &= b_2 - b_3 \\ x_3 &= \tfrac{1}{2}b_1 - \tfrac{1}{2}b_2 + \tfrac{1}{2}b_3, \end{aligned}$$

und stellen fest: Für jedes $\underline{b} \in K^3$ besitzt das Gleichungssystem

$$\begin{aligned} x_2 + 2x_3 &= b_1 \\ x_1 + 2x_2 + x_3 &= b_2 \\ x_1 + x_2 + x_3 &= b_3 \end{aligned}$$

genau eine Lösung (nämlich die obige).

Anhand dieses Beispiels mache man sich allgemein klar:

1.11 *Sei $A \in K^{m\times n}$. Folgende Aussagen sind äquivalent:*

(i) *Das lineare Gleichungssystem $(*)$ besitzt für* **jedes** $\underline{b} \in K^m$ **genau** *eine Lösung.*

(ii) *Es ist $n = m$ und es gilt: Ist A' eine Stufenmatrix, die aus A durch Zeilenumformungen Z_1, Z_2, Z_3 hervorgeht, so ist $A' = \mathbf{1}_n$.* □

Um das Gleichungssystem $(*)$ zu lösen, kann man auch die Spalten von A umformen. Erlaubte **Spaltenumformungen** sind

S_1 Addiere ein Vielfaches einer Spalte zu einer anderen.

S_2 Multipliziere eine Spalte mit einem Skalar $\lambda \in K, \lambda \neq 0$.

S_3 Vertausche zwei Spalten.

Hierbei ändern sich allerdings die Variablen $x_1, \ldots, x_n$. Vertauscht man z. B. zwei Spalten a_{*i} und a_{*j}, so hat man auch x_i und x_j zu vertauschen. Wir werden in einem allgemeineren Zusammenhang darauf zurück kommen, und begnügen uns hier mit einem einfachen Beispiel. Sei die Gleichung

$(+)$ $\qquad ax + by = c \quad$ mit $\quad a, b, \in K, a \neq 0$

gegeben. In diese Gleichung substituieren wir neue Variablen x', y', die durch die Gleichungen

$$(T) \qquad \begin{array}{ll} x = x' - \frac{b}{a}y' \ , & \text{also} \quad x' = x + \frac{b}{a}y' \\ y = y' & \qquad\quad y' = y \end{array}$$

definiert sind. Dies ergibt

$$ax + by \quad = \quad a\left(x' - \tfrac{b}{a}y'\right) + by' = ax'.$$

Wir haben also die Gleichung

$(++)$ $\qquad ax' + 0 \cdot y' = c$

zu lösen, um mit Hilfe der *Transformation* (T) die Lösung (x, y) zu bekommen. Dabei erhält man die Koeffizientenmatrix $(a, 0) \in K^{1\times 2}$ von $(++)$ aus der Koeffizientenmatrix $(a, b) \in K^{1\times 2}$ von $(+)$, indem man das $(-\frac{b}{a})$-fache der ersten Spalte von (a, b) zur zweiten Spalte von (a, b) addiert, also durch S_1.

Zeilenumformungen $Z_1,\ Z_2,\ Z_3$ und Spaltenumformungen $S_1,\ S_2,\ S_3$ einer Matrix A nennen wir *elementare Umformungen von A.*

Notation: **a)** Für $A \in K^{m\times n}$ sei $[A]$ die Menge aller $B \in K^{m\times n}$, die aus A durch elementare Umformungen hervorgehen.

b) Für $r = 0, 1, \ldots, \min\{n, m\}$ sei $E_r \in K^{m\times n}$ definiert durch $E_0 := \mathbf{0}$ und

$$E_r := \left(\begin{array}{c|c} \mathbf{1}_r & \mathbf{0} \\ \hline \mathbf{0} & \end{array}\right), \quad r \geq 1.$$

Offenbar ist $E_n = \mathbf{1}_n$, wenn $n = m$. Man mache sich klar:

1.12 *Sei $A \in K^{m\times n}$ eine Stufenmatrix und r die Anzahl der von $\underline{0}$ verschiedenen Zeilen von A. Dann kann A durch elementare Spaltenumformungen in die Matrix E_r umgewandelt werden.* □

Mit 1.8 folgt:

1.13 Satz: *Sei $A \in K^{m\times n}$. Dann existiert eine Zahl $r \in \mathbb{N}_0$, $r \leq \min\{n, m\}$, mit $E_r \in [A]$.* □

Die Frage, ob diese Zahl r durch A bestimmt ist, werden wir später beantworten.

Wir definieren nun – zum späteren Gebrauch – die Addition und skalare Multiplikation von Matrizen.

Die Menge $K^{m\times n}$ aller $(m\times n)$-Matrizen kann mit dem K^{mn} identifiziert werden. Für

$$A = (a_{ij}),\ B = (b_{ij}) \quad \in K^{m\times n} \quad \text{und} \quad \lambda \in K$$

sei dementsprechend *(komponentenweise)* definiert.

Matrizenaddition:

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

$$\lambda A := \begin{pmatrix} \lambda a_{11} & \dots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \dots & \lambda a_{mn} \end{pmatrix} = (\lambda a_{ij})$$

$$-A := (-a_{ij})$$

$$A - B := A + (-B) = (a_{ij} - b_{ij}).$$

Zum Schluss dieses Kapitels erklären wir eine grundlegende Sprachregelung der Mathematik.

Definition: Sei X eine nichtleere Menge und

$$X^2 := \{(x, y) \mid x, y \in X\} \qquad \textit{(kartesisches Produkt)}$$

Eine *Relation $\mathfrak{R}$ auf X* ist eine Teilmenge $\mathfrak{R}$ der Menge X^2.

Sei $\mathfrak{R}$ Relation auf X. Man schreibt

$$x \sim y \overset{\text{def.}}{\Leftrightarrow} (x, y) \in \mathfrak{R}.$$

Die Relation $\mathfrak{R}$ heißt *Äquivalenzrelation*, wenn folgende Aussagen für alle $x, y, z \in X$ gelten:

(i) $x \sim x$ *(reflexiv)*

(ii) $x \sim y \Rightarrow y \sim x$ *(symmetrisch)*

(iii) $x \sim y$ und $y \sim z \;\Rightarrow\; x \sim z$ *(transitiv)*

Beispiele: Wir beginnen mit zwei trivialen Beispielen von Äquivalenzrelationen:

1) $\mathfrak{R} = X^2$, d.h. $x \sim y$ für alle $x, y \in X$.

2) $\mathfrak{R} = \{(x, x) \mid x \in X\}$, d.h. $x \sim y \Leftrightarrow x = y$.

Das dritte Beispiel ist nicht trivial: Sei $X := K^{m\times n}$. Anstatt $\mathfrak{R} \subseteq X^2$ definieren wir $A \sim B$ für $A, B \in X$

$$A \sim B \overset{\text{def.}}{\Leftrightarrow} B \in [A],$$

d.h. B entsteht aus A durch elementare Umformungen. Wir behaupten, dass "$\sim$" eine Äquivalenzrelation ist: Die Aussage (i) ist trivial *(man tut nichts)*; (ii) ist etwas diffiziler: Ist $B \in [A]$, so kann A aus B gewonnen werden, wenn man dieselben elementaren Umformungen, welche A in B überführen, rückwärts anwendet. Man mache Beispiele und vergleiche 1.7. Die Aussage (iii) ist wieder

einfach: Entsteht B aus A und C aus B durch elementare Umformungen, so entsteht C aus A durch solche – und B ist eine *Zwischenstation*.

Sei "$\sim$" eine Äquivalenzrelation auf der Menge X. Für $x \in X$ sei

$$[x] := \{y \in X \mid x \sim y\}.$$

Die Menge $[x]$ ist eine Teilmenge von X, sie heißt *Äquivalenzklasse*. Eine Umformulierung der Gesetze (i), (ii), (iii) ist:

1.14 *a)* $x \in [x]$

b) $y \in [x] \Rightarrow [x] = [y]$

c) $[x] \cap [y] \neq \emptyset \Rightarrow [x] = [y]$

Beweis:

a) ist (i).

b) Sei $y \in [x]$ und $z \in [y]$. Nach (iii) ist $x \sim z$, also $z \in [x]$. Dies beweist $[y] \subseteq [x]$. Nach (ii) ist $x \in [y]$. Mit x, y vertauscht folgt genauso $[x] \subseteq [y]$. Also gilt $[x] = [y]$.

c) Sei $z \in [x] \cap [y]$. Aus b) folgt $[x] = [z] = [y]$. □

Notation: Die Menge aller Teilmengen der Menge X heißt *Potenzmenge* von X; wir bezeichnen sie mit 2^X. Seien $Y, Z \in 2^X$

$$Y \cup Z := \{x \in X \mid x \in Y \text{ oder } x \in Z\} \ (\in 2^X) \qquad \textit{(Vereinigung)}$$

$$Y \cap Z := \{x \in X \mid x \in Y \text{ und } x \in Z\} \ (\in 2^X) \qquad \textit{(Schnitt)}$$

Ist $Y \cap Z = \emptyset$, so heißen Y, Z *disjunkt*.

Definition: Sei $\mathfrak{P}$ eine Teilmenge von 2^X. Dann heißt $\mathfrak{P}$ eine *Partition* von X, wenn gilt:

1) Zu jedem $x \in X$ existiert ein $Y \in \mathfrak{P}$ mit $x \in Y$.

2) Seien $Y, Z \in \mathfrak{P}$ mit $Y \neq Z$. Dann ist $Y \cap Z = \emptyset$.

1.15 *a) Sei "$\sim$" eine Äquivalenzrelation auf X, und sei $\mathfrak{P} = \{[x] \mid x \in X\}$. Dann ist $\mathfrak{P}$ eine Partition von X.*

b) Sei $\mathfrak{P}$ eine Partition auf X. Definiere auf X die Relation

$$x \sim y \Leftrightarrow \textit{Es existiert } Y \in \mathfrak{P} \textit{ mit } x, y \in Y.$$

Dann ist diese Relation Äquivalenzrelation und $\mathfrak{P}$ ist die Menge der Äquivalenzklassen.

Beweis: a) folgt aus 1.14 a), b).

b) Wir haben (i), (ii) und (iii) zu zeigen. Die Aussagen (i) und (ii) sind trivial. Wir zeigen (iii). Sei $x \sim y$ und $y \sim z$, d.h. es existieren $Y, Z \in \mathfrak{P}$ mit $x, y \in Y$ und $y, z \in Z$. Also ist $y \in Y \cap Z$. Aus 2) folgt $Y = Z$. Also liegen x, z in Y, d.h. $x \sim z$. □

Zum Schluss geben wir noch ein Beispiel einer Äquivalenzrelation, welche zur Definition der rationalen Zahlen führt.

X sei die Menge aller Paare $a = (a_1, a_2) \in \mathbb{Z}^2$ mit $a_2 \neq 0$. Wir definieren für $a, b \in X$

$$a \sim b \overset{\text{def.}}{\Leftrightarrow} a_1 b_2 = b_1 a_2$$

und behaupten, dass dies eine Äquivalenzrelation auf X ist:

$a \sim a$, denn $a_1 a_2 = a_2 a_1$
$a \sim b \Rightarrow b \sim a$, denn $b_1 a_2 = a_1 b_2$

Sei $a \sim b$ und $b \sim c$, also $a_1 b_2 = b_1 a_2$ und $b_1 c_2 = c_1 b_2$. Wir haben $a \sim c$, d.h.

$(+)$ $\quad a_1 c_2 = c_1 a_2$

Zu zeigen: Ist $b_1 = 0$, so ist $a_1 b_2 = 0 = c_1 b_2$, also $a_1 = 0 = c_1$, denn $b_2 \neq 0$. Es folgt $(+)$. Sei $b_1 \neq 0$. Es ist

$$a_1(b_1 c_2) = a_1(c_1 b_2) = c_1(a_1 b_2) = c_1(b_1 a_2),$$

also $a_1 b_1 c_2 = c_1 b_1 a_2$. Daraus folgt $(+)$, denn $b_1 \neq 0$.

Also ist "$\sim$" Äquivalenzrelation. Wir bezeichnen mit $\mathbb{Q}$ die Menge aller Äquivalenzklassen $[a]$, und schreiben

$$\frac{a_1}{a_2} := [a].$$

Wir definieren auf $\mathbb{Q}$ eine *Addition* und *Multiplikation* durch

$$(++) \qquad \begin{aligned} \frac{a_1}{a_2} + \frac{b_1}{b_2} &:= \frac{a_1 b_2 + a_2 b_1}{a_2 b_2} \\ \frac{a_1}{a_2} \cdot \frac{b_1}{b_2} &:= \frac{a_1 b_1}{a_2 b_2} \end{aligned}$$

Wir müssen zeigen, dass dies *wohldefiniert*, d.h. *unabhängig* von der Wahl der Repräsentanten $a \in [a] = \frac{a_1}{a_2}$ und $b \in [b] = \frac{b_1}{b_2}$ ist. Sei $a' \in [a]$ und $b' \in [b]$, also

$$a_1 a_2' = a_1' a_2 \quad \text{und} \quad b_1 b_2' = b_1' b_2.$$

Man verifiziere:

$$(a_1 b_1 + a_2 b_1) a_2' b_2' = (a_1' b_1' + a_2' b_1') a_2 b_2$$

$$(a_1 b_1)(a_2' b_2') = (a_1' b_1')(a_2 b_2)$$

Doch dies ist $(++)$.

Also ist die Addition und Multiplikation in $\mathbb{Q}$ erklärt. Es folgen nun die üblichen Rechengesetze, die in $\mathbb{Q}$ gelten.

2 Abbildungen, lineare Abbildungen und Permutationen

Seien X, Y zwei Mengen. Eine *Abbildung von X nach Y* ist eine Vorschrift α, die jedem $x \in X$ ein eindeutig bestimmtes Element $y \in Y$ zuordnet; wir schreiben $y = \alpha(x)$ und

$$\alpha : X \to Y \quad \text{mit} \quad x \mapsto \alpha(x).$$

Es ist

$$\text{Bild}\,\alpha := \{\alpha(x) \mid x \in X\}$$

eine Teilmenge von Y. Die Abbildung α heißt *surjektiv*, falls Bild $\alpha = Y$, und α heißt *injektiv*, wenn gilt

$$a, b \in X, a \neq b \quad \Rightarrow \quad \alpha(a) \neq \alpha(b).$$

α heißt *bijektiv*, wenn α injektiv und surjektiv ist. Letzteres bedeutet, dass es zu jedem $y \in Y$ genau ein $x \in X$ gibt mit $\alpha(x) = y$. Nur in diesem Fall kann die *Umkehrabbildung* α^{-1} erklärt werden:

$$\alpha^{-1} : Y \to X \quad \text{mit} \quad y \mapsto x, \quad \text{falls} \quad \alpha(x) = y.$$

Sei α eine Abbildung von X nach Y, und β eine Abbildung von Y in eine weitere Menge Z. Dann heißt die Abbildung

$$\beta \circ \alpha : X \to Z \quad \text{mit} \quad x \mapsto \beta(\alpha(x))$$

die *Komposition* von α mit β.

Jede Menge X lässt die *identische Abbildung* $\mathrm{id} = \mathrm{id}_X$ zu

$$\mathrm{id}(x) = x \quad \text{für alle} \quad x \in X.$$

Sei $\alpha : X \to Y$ bijektiv; es existiert also die Umkehrabbildung α^{-1}. Dann ist

$$(\alpha^{-1} \circ \alpha)(x) = \alpha^{-1}(\alpha(x)) = x,$$

also $\alpha^{-1} \circ \alpha = \mathrm{id}_X$ und

$$(\alpha \circ \alpha^{-1})(y) = \alpha(\alpha^{-1}(y)) = y,$$

also $\alpha \circ \alpha^{-1} = \mathrm{id}_Y$.

Seien X_1, X_2, X_3, X_4 Mengen und $\alpha_1, \alpha_2, \alpha_3$ Abbildungen mit

$$\alpha_1 : X_1 \to X_2, \quad \alpha_2 : X_2 \to X_3, \quad \alpha_3 : X_3 \to X_4.$$

Dann gilt:

$$\begin{aligned}(\alpha_3 \circ (\alpha_2 \circ \alpha_1))(x) &= \alpha_3(\alpha_2(\alpha_1(x))) = (\alpha_3 \circ \alpha_2)(\alpha_1(x)) \\ &= ((\alpha_3 \circ \alpha_2) \circ \alpha_1)(x)\end{aligned}$$

und deswegen

$$\alpha_3 \circ (\alpha_2 \circ \alpha_1) \quad = \quad (\alpha_3 \circ \alpha_2) \circ \alpha_1.$$

Die Komposition von Abbildungen ist also *assoziativ.*

Bemerkung: Sei $\alpha : X \to Y$ eine Abbildung. Für $a, b \in X$ setze

$$a \sim_\alpha b \overset{\text{def.}}{\Leftrightarrow} \alpha(a) = \alpha(b).$$

Dann ist "$\sim_\alpha$" eine Äquivalenzrelation auf der Menge X, und die Äquivalenzklassen haben die Form

$$[a] := \{x \in X \mid \alpha(x) = \alpha(a)\}.$$

Sei $X/\sim$ die Menge der Äquivalenzklassen. Dann ist

$$\alpha_1 : X \to X/\sim \quad \text{mit} \quad a \mapsto [a] \qquad \text{(wohldefiniert!)}$$

eine surjektive Abbildung und

$$\alpha_2 : X/\sim \to \text{ Bild } \alpha \quad \text{mit} \quad [a] \to \alpha(a)$$

eine bijektive Abbildung. Es gilt $(\alpha_2 \circ \alpha_1)(a) = \alpha_2([a]) = \alpha(a)$, d.h. wir haben ein *Faktorisierung* $\alpha = \alpha_2 \circ \alpha_1$ von α.

Sei umgekehrt "$\sim$" eine Äquivalenzrelation auf der Menge X und $X/\sim$ die Menge der Äquivalenzklassen von "$\sim$". Dann ist

$$\alpha : X \to X/\sim \quad \text{mit} \quad x \mapsto [x]$$

eine (surjekte) Abbildung.

Notation: Sei $A = (a_{ij}) \in K^{m \times n}$ und $\underline{x} \in K^n$. Wir setzen

$$A\underline{x} := \left(\sum_{j=1}^{n} a_{1j}x_j, \ldots, \sum_{j=1}^{n} a_{mj}x_j \right) \overset{!}{=} \sum_{j=1}^{n} x_j a_{*j} \ \in K^m.$$

Das Gleichungssystsem $(*)$ in **1** hat nun die Form

$(*)$ $\quad A\,\underline{x} = \underline{b}.$

Merke: *Für $j = 1, \ldots, n$ gilt $A\,e_j = a_{*j}$.*

Definition: Eine Abbildung $\alpha : K^n \to K^m$ heißt *linear*, wenn für alle $\underline{x}, \underline{y} \in K^n$ und $\lambda \in K$ gilt

$L_1)$ $\quad \alpha(\underline{x} + \underline{y}) = \alpha(\underline{x}) + \alpha(\underline{y})$

$L_2)$ $\quad \alpha(\lambda \underline{x}) = \lambda\, \alpha(\underline{x})$

Man rechnet nach:

2.1 *Sei $A \in K^{m\times n}$. Dann ist die Abbildung*

$$\alpha : K^n \to K^m \quad \textit{mit} \quad \underline{x} \mapsto A\underline{x}$$

linear. □

2.2 *Sei $\alpha : K^m \to K^n$ linear. Dann ist $\alpha(\underline{0}) = \underline{0}$* [7] *und $\alpha(-\underline{x}) = -\alpha(\underline{x})$.*

Beweis: Beide Aussagen folgen aus L_2. Sie folgen aber auch aus L_1:

$$\alpha(\underline{0}) = \alpha(\underline{0} + \underline{0}) \overset{L_1}{=} \alpha(\underline{0}) + \alpha(\underline{0}) \Rightarrow \alpha(\underline{0}) = \underline{0}$$

$$0 = \alpha(\underline{0}) = \alpha(\underline{x} - \underline{x}) = \alpha(\underline{x} + (-\underline{x}) \overset{L_1}{=} \alpha(\underline{x}) + \alpha(-\underline{x}).$$

Also ist $\alpha(-\underline{x}) = -\alpha(\underline{x})$. □

Merke: *Sei $\alpha : K^n \to K^m$ linear. Existiert ein $\underline{x} \in K^n$ mit*

$$\underline{x} \neq \underline{0} \quad \textit{und} \quad \alpha(\underline{x}) = \underline{0},$$

so ist α nicht injektiv.

Geometrische Interpretation: Sei $\alpha : K^2 \to K^2$ linear und bijektiv. Ist G eine Gerade, also

$$G = K\underline{a} + \underline{p} = \{\lambda \underline{a} + \underline{p} \mid \lambda \in K\},\ \underline{a} \neq \underline{0},$$

so bildet α einen Punkt $\underline{x} = \lambda \underline{a} + \underline{p}$ von G auf den Punkt

$$\alpha(\underline{x}) = \alpha(\lambda \underline{a} + \underline{p}) \quad \overset{L_1}{=} \quad \alpha(\lambda \underline{a}) + \alpha(\underline{p}) \quad \overset{L_2}{=} \quad \lambda\alpha(\underline{a}) + \alpha(\underline{p})$$

ab. Da α bijektiv ist, ist mit $\underline{a}$ auch $\alpha(\underline{a}) \neq \underline{0}$. Also ist auch

$$\alpha(G) := \{\alpha(\underline{x}) \mid \underline{x} \in G\}$$

eine Gerade.

Merke: *Eine bijektive lineare Abbildung bildet Geraden auf Geraden ab.*

[7] Die erste $\underline{0}$ ist in K^n und die zweite in K^m.

2.3 Satz: *Zu jeder Wahl von n Elementen $a_{*1}, \ldots, a_{*n}$ aus K^m (nicht notwendig verschieden) existiert genau eine lineare Abbildung $\alpha : K^n \to K^m$ mit*

$$\alpha(e_j) = a_{*j} \quad , \quad i = 1, \ldots, n,$$

und α hat die Form

$$\alpha(\underline{x}) = A\underline{x}, \qquad \underline{x} \in K^n,$$

wobei A die $(m \times n)$-Matrix ist, welche die Spalten $a_{*1}, \ldots, a_{*n}$ besitzt.

Beweis: Die Matrix $A \in K^{m \times n}$ besitze die Spalten $a_{*1}, \ldots, a_{*n}$ und sei $\alpha : K^n \to K^m$ definiert durch $\alpha(\underline{x}) = A\underline{x}$. Dann ist α linear (2.1) und $\alpha(e_j) = a_{*j}$.
Sei $\alpha' : K^n \to K^m$ linear mit $\alpha'(e_j) = a_{*j} \quad (j = 1, \ldots, n)$. Wir zeigen $\alpha = \alpha'$.
Für $\underline{x} \in K^n$ ist

$$\underline{x} = \textstyle\sum_{j=1}^{n} x_j e_j.$$

Es folgt

$$\begin{aligned} \alpha'(\underline{x}) &= \alpha'\left(\sum_{j=1}^{n} x_j e_j\right) \overset{L_1, L_2}{=} \sum_{1=r}^{n} x_j \alpha'(e_j) \\ &= \sum_{j=1}^{n} x_j a_{*j} = \sum_{j=1}^{n} x_j \alpha(e_j) \\ &\overset{L_1, L_2}{=} \alpha(\underline{x}), \end{aligned}$$

also $\alpha = \alpha'$. □

2.4 KOROLLAR: *Die linearen Abbildungen $\alpha : K^n \to K^m$ haben die Form $\alpha(\underline{x}) = A\underline{x}$ mit $A \in K^{m \times n}$.* □

Zum Beispiel gehört zu der identischen Abbildung $\mathrm{id} : K^n \to K^n$ die Einheitsmatrix $\mathbf{1}_n$.

Beispiel: Sei $\alpha : K^3 \to K^2$ die lineare Abbildung mit

$$\alpha(e_1) = (2, 3), \quad \alpha(e_2) = (1, 0), \quad \alpha(e_3) = (1, 1).$$

Dann ist $\alpha(\underline{x}) = A\underline{x}$ mit

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 3 & 0 & 0 \end{pmatrix} \in K^{2 \times 3},$$

z. B. berechnet man $\alpha(1,1,1) = (4,3)$.

Drehung: Sei φ ein Winkel zwischen 0 und 2π und

$$\delta : K^2 \to K^2$$

die Drehung des K^2 im Punkte $\underline{0}$ um den Winkel φ Dann ist

$$\delta(e_1) = (\cos\varphi, \sin\varphi) \text{ und } \delta(e_2) = (-\sin\varphi, \cos\varphi).$$

Nun definieren wir gemäß 2.3 eine lineare Abbildung $\alpha : K^2 \to K^2$ durch

$$\alpha(e_1) := \delta(e_1) \text{ und } \alpha(e_2) := \delta(e_2),$$

setzen also $a_{*1} = (\cos\varphi, \sin\varphi)$ und $a_{*2} = (-\sin\varphi, \cos\varphi)$. Dann ist $\alpha(\underline{x}) = A\underline{x}$ mit

$$A = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \in K^{2\times 2}$$

und für $\underline{x} = (x_1, x_2)$ folgt

$$\alpha(\underline{x}) = (x_1\cos\varphi - x_2\sin\varphi\,,\ x_1\sin\varphi + x_2\cos\varphi).$$

Behauptung: $\alpha = \delta$.

Beweis: Sei ψ der Winkel, den die Gerade durch $\underline{0}$ und $\underline{x}$ mit der x_1-Achse bildet und λ der Abstand zwischen $\underline{0}$ und $\underline{x}$. Dann ist

$$x_1 = \lambda\cos\psi \quad \text{und} \quad x_2 = \lambda\sin\psi.$$

Mit dem Additionstheorem von $\sin$ und $\cos$ folgt

$$\begin{aligned} \alpha(\underline{x}) &= (\lambda\cos\psi\cos\varphi - \lambda\sin\psi\sin\varphi,\ \lambda\cos\psi\sin\varphi + \lambda\sin\psi\cos\varphi) \\ &\overset{!}{=} (\lambda\cos(\psi+\varphi),\ \lambda\sin(\varphi+\psi)). \end{aligned}$$

Also ist $\alpha = \delta$. □

Notation: Wir bezeichnen mit $\mathrm{Hom}\,(K^n, K^m)$ die Menge aller linearen Abbildungen $\alpha : K^n \to K^m$, und **definieren** auf der Menge $\mathrm{Hom}\,(K^n, K^m)$ eine *Addition* und *skalare Multiplikation*: Seien $\alpha, \beta \in \mathrm{Hom}\,(K^n, K^m),\ \lambda \in K$

$$\begin{aligned} \alpha + \beta : K^n \to K^m \quad &\text{mit} \quad \underline{x} \mapsto \alpha(\underline{x}) + \beta(\underline{x}) \\ \lambda\alpha : K^n \to K^m \quad &\text{mit} \quad \underline{x} \mapsto \lambda\alpha(x) \end{aligned}$$

Man überlege sich:

2.5 *Die Abbildungen $\alpha + \beta$ und $\lambda\alpha$ sind linear (liegen also in $\mathrm{Hom}\,(K^n, K^m)$. Gehört zu α bzw. β die Matrix A bzw. B, so gehört zu $\alpha + \beta$ die Matrix $A + B$*

und zu $\lambda\alpha$ die Matrix λA. $\square$

Im Folgenden sind drei Zahlen $n, m, l \in \mathbb{N}$ im Spiel. Sei $A = (a_{ij}) \in K^{m\times n}$, $B = (b_{ij}) \in K^{l\times m}$ und $C = (c_{ij}) \in K^{l\times n}$.

Definition: Die Matrix C heißt das *Produkt* der Matrizen B, A, wenn für $i = 1, \ldots, l$ und $j = 1, \ldots, n$

$$c_{ij} = \sum_{k=1}^{m} b_{ik}a_{kj}.$$

Wir schreiben $C = BA$.

Zu $\alpha \in \mathrm{Hom}\,(K^n, K^m)$ gehöre die Matrix A und zu $\beta \in \mathrm{Hom}\,(K^m, K^l)$ gehöre die Matrix B.

2.6 Satz: *Die Abbildung $\beta \circ \alpha : K^n \to K^l$ ist linear, und zu ihr gehört die Matrix BA.*

Beweis: Sei $\gamma := \beta \circ \alpha : K^n \to K^l$. Wir zeigen, dass γ linear ist, also L_1, L_2:

$$\begin{aligned} L_1:\ \gamma(\underline{x}+\underline{y}) &= \beta(\alpha(\underline{x}+\underline{y})) \overset{L_1}{=} \beta(\alpha(\underline{x}) + \alpha(\underline{y})) \\ &\overset{L_1}{=} \beta(\alpha(\underline{x})) + \beta(\alpha(y)) \\ &= \gamma(\underline{x}) + \gamma(\underline{y}) \end{aligned}$$

$$\begin{aligned} L_2:\ \gamma(\lambda\underline{x}) &= \beta(\alpha(\lambda\underline{x})) \overset{L_1}{=} \beta(\lambda\alpha(\underline{x})) \\ &\overset{L_1}{=} \lambda(\beta(\alpha(\underline{x}))) \\ &= \lambda\gamma(\underline{x}) \end{aligned}$$

Sei $C = (c_{ij}) \in K^{l\times n}$ die Matrix, welche zu der linearen Abbildung γ gehört. Nach unserer Merkregel ist Ce_j die j–te Spalte von C. Also ist

$$c_{*j} = \gamma(e_j) = \beta(\alpha(e_j)) = B(Ae_j) = Ba_{*j} = \textstyle\sum_{k=1}^{m} a_{kj}b_{*k}.$$

Nehmen wir die i–te Komponente dieser Vektorgleichung, so folgt $C = BA$. $\square$

Wir haben nun verschiedene Rechengesetze in $K^{m\times n}$ und in $\mathrm{Hom}\,(K^n, K^m)$; wir formulieren sie im Fall $n = m$.

2.7 *a) Für $A, B, C \in K^{n\times n}$ gilt:*

$A + (B + C) = (A + B) + C$ *und* $(AB)C = A(BC)$ Assoziativgesetz

$A(B + C) = AB + AC$ *und* $(B + C)A = BA + CA$ Distributivgesetz

b) Für $\alpha, \beta, \gamma \in$ *Hom* (K^n, K^n) *gilt:*

$\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$ *und* $(\alpha \circ \beta) \circ \gamma = \alpha(\beta \circ \gamma)$ Assoziativgesetz

$\alpha \circ (\beta + \gamma) = \alpha \circ \beta + \alpha \circ \gamma$ *und* $(\beta + \gamma) \circ \alpha = \beta \circ \alpha + \gamma \circ \alpha$ Distributivgesetz

Beweis: Wir haben anfangs bemerkt, dass die Komposition assoziativ ist, also gilt das Assoziativgesetz Hom (K^n, K^m). Mit 2.6 folgt daraus das Assoziativgesetz in a). Zum Distributivgesetz in Hom (K^n, K^m) notieren wir:

$$\begin{aligned}(\alpha \circ (\beta + \gamma))(\underline{x}) &= \alpha((\beta + \gamma)(\underline{x})) = \alpha(\beta(\underline{x}) + \gamma(\underline{x})) \\ &= \alpha(\beta(\underline{x}) + \alpha(\gamma(\underline{x})) \\ &= (\alpha \circ \beta)(\underline{x}) + (\alpha \circ \gamma)(\underline{x})\end{aligned}$$

Analog folgt das andere Distributivgesetz. Mit der Übersetzung 2.6 ergibt sich daraus auch a). □

Definition: Sei $A = (a_{ij}) \in K^{m \times n}$. Eine Matrix $B = (b_{ij}) \in K^{n \times m}$ ist die zu A *transponierte Matrix*, wenn für $i = 1, \ldots, n$ und $j = 1, \ldots, m$

$$b_{ij} = a_{ji}.$$

Wir schreiben $B = A^T$.

Ein Vektor $\underline{x} = (x_1, \ldots, x_n) \in K^n$ kann als Matrix $\in K^{1 \times n}$ aufgefasst werden. Also ist

$$\underline{x}^T = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^{n \times 1}.$$

Ist nun $\underline{y} = (y_1, \ldots, y_n) \in K^{1 \times n}$, so ist das Matrixprodukt

$$\underline{y}\ \underline{x}^T = \sum_{i=1}^{n} y_i x_i \qquad (\in K^{1 \times 1} = K).$$

erklärt. Ist $A \in K^{m \times n}$, so ist auch das Matrixprodukt $A\underline{x}^T \in K^{m \times 1}$ erklärt;

dabei ist

$$A\,\underline{x}^T = \begin{pmatrix} a_{1*}\underline{x}^T \\ \vdots \\ a_{m*}\underline{x}^T \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix} \stackrel{!}{=} \sum_{j=1}^n x_j a_{*j}.$$

Dabei fassen wir nun die Spalten a_{*j} bzw. Zeilen a_{i*} als Matrizen aus $K^{m\times 1}$ bzw. $K^{1\times n}$ auf:

$$a_{*j} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \quad , \quad a_{i*} = (a_{i1}, \ldots, a_{in}).$$

Also ist $A\underline{x}^T$ unser früher erklärte Ausdruck $A\underline{x}$.

Sei $B \in K^{l\times m}$. Ist $C = BA \in K^{l\times n}$, so ist c_{ij} nun das Matrixprodukt

$$c_{ij} = b_{i*}a_{j*}.$$

2.8 *Sei $\alpha : K^n \to K^n$ eine bijektive lineare Abbildung, und zu α gehöre die Matrix $A \in K^{n\times n}$.*

a) Die Umkehrabbildung $\alpha^{-1} : K^n \to K^n$ ist ebenfalls linear.

b) Gehört zu α^{-1} die Matrix B, so ist $BA = AB = \mathbf{1}_n$.

Beweis:

a) Zu $\underline{x}, \underline{y} \in K^n$ existieren $\underline{v}, \underline{w} \in K^n$ mit

$$\alpha(\underline{v}) = \underline{x} \quad \text{und} \quad \alpha(\underline{w}) = \underline{y}.$$

Es folgt

$$\alpha(\underline{v} + \underline{w}) = \alpha(\underline{v}) + \alpha(\underline{w}) = \underline{x} + \underline{y} \quad \text{und} \quad \alpha(\lambda\underline{v}) = \lambda\alpha(\underline{v}) = \lambda\underline{x},$$

also

$$\begin{aligned} \alpha^{-1}(\underline{x} + \underline{y}) &= \underline{v} + \underline{w} = \alpha^{-1}(\underline{x}) + \alpha^{-1}(\underline{y}) \\ \alpha^{-1}(\lambda\underline{x}) &= \lambda\underline{v} = \lambda\alpha^{-1}(\underline{x}) \end{aligned}$$

b) Es ist $\alpha^{-1} \circ \alpha = \mathrm{id}$ und $\alpha \circ \alpha^{-1} = \mathrm{id}$. Zu der Abbildung id gehört die Einheitsmatrix $\mathbf{1}_n$. Aus 2.6 folgt $BA = AB = \mathbf{1}_n$. □

Im Folgenden sei $A \in K^{n\times n}$ und $\underline{b} \in K^n$. Wir betrachten das lineare Gleichungssystem

$(*)$ $\quad A\underline{x} = \underline{b}$, oder

$\quad A\underline{x}^T = b^T$ (Matrixschreibweise)

Sei α die zu A gehörende lineare Abbildung

$$\alpha : K^n \to K^m \quad \text{mit} \quad \underline{x} \mapsto A\underline{x}.$$

Die erweiterte Koeffizientenmatrix $(A, \underline{b})$ von $(*)$ überführen wir mit Z_1, Z_2, Z_4 in eine Matrix $(A', \underline{b}')$, sodass A' eine Stufenmatrix ist. Man beachte im Folgenden, dass die Gleichungssysteme $A\underline{x} = \underline{b}$ und $A'\underline{x} = \underline{b}'$ dieselbe Lösungsmenge besitzen (1.7).

Zunächst mache man sich klar:

2.9 *Es existiert $B \in K^{n\times n}$ mit $\underline{b}' = B\underline{b}$.* □

Anstatt eines Beweises nehmen wir das Beispiel auf Seite 16: Sei

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} \in K^{3\times 3}.$$

Durch elementare Zeilenumformungen haben wir (A, b) in folgende Matrix

$$(A', b') = \begin{pmatrix} 1 & 0 & 0 & -\frac{1}{2}b_1 - \frac{1}{2}b_2 + \frac{3}{2}b_3 \\ 0 & 1 & 0 & b_2 - b_3 \\ 0 & 0 & 1 & \frac{1}{2}b_1 - \frac{1}{2}b_2 + \frac{1}{2}b_3 \end{pmatrix}$$

übergeführt. Ist also

$$B := \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{3}{2} \\ 0 & 1 & -1 \\ \frac{1}{2} & -\frac{1}{2} & +\frac{1}{2} \end{pmatrix},$$

so ist $\underline{b}' = B\underline{b}$. In diesem Beispiel ist $\underline{x} = b' = Bb$ die einzige Lösung von $(*)$, siehe Seite 16. Wir rechnen:

$$BA = \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{3}{2} \\ 0 & 1 & -1 \\ \frac{1}{2} & -\frac{1}{2} & +\frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

und $AB = \mathbf{1}_3$.

Wir verallgemeinern diese Situation und bemerken zunächst:

2.10 *Sei B wie in 2.9. Ist $A' = \mathbf{1}_n$, so ist $AB = \mathbf{1}_n$.*

Beweis: Dann ist $x = b' = B\underline{b}$ die einzige Lösung von $(*)$. Für alle $\underline{b} \in K^n$ gilt also die Matrixgleichung

$$\underline{b}^T = A\underline{x} = A(Bb^T) \stackrel{!}{=} (AB)b^T.$$

Damit ist die Abbildung $b^T \mapsto (AB)b^T$ die identische Abbildung. Es folgt $AB = \mathbf{1}_n$. □

2.11 Satz: *Folgende Aussagen sind äquivalent:*

(i) *Es existiert $B \in K^{n\times n}$ mit $BA = \mathbf{1}_n$.*

(ii) *Die Abbildung α ist injektiv, d.h. für jedes $\underline{b} \in K^n$ hat $(*)$ höchstens eine Lösung.*

(iii) *Die Abbildung α ist bijektiv, d.h. für jedes $\underline{b} \in K^n$ hat $(*)$ genau eine Lösung.*

(iv) *Die Abbildung α ist surjektiv, d.h. für jedes $\underline{b} \in K^n$ hat $(*)$ eine Lösung.*

(v) *Es existiert ein $B \in K^{n\times n}$ mit $BA = AB = \mathbf{1}_n$*

Beweis: (i) ⇒ (ii): Sei $\underline{x} \in K^n$ Lösung von $(*)$. Da die Matrixmultiplikation assoziativ ist, folgt

$$Bb^T = B(Ax^T) \stackrel{!}{=} (BA)x^T = \mathbf{1}_n x^T = x^T.$$

Also ist Bb^T die einzige Lösung von $(*)$.

(ii) $\Rightarrow$ (iii): Sei d die Anzahl der Nullzeilen der Stufenmatrix A'. Im Fall $d > 0$ gibt es Stufen der "Länge > 1", also mehrere Lösungen von $(*)$ (1.10). Also ist $d = 0$. Aus der Definition der Stufenmatrix A' folgt $A' = \mathbf{1}_n$. Nach 2.9 hat $\underline{b}'$ die Form $\underline{b}' = B\underline{b}$ mit $B \in K^{n\times n}$. Dann ist $b' = B\underline{b}$ für jedes $\underline{b} \in K^n$ die einzige Lösung von $(*)$, siehe 1.10 (a).

(iii) $\Rightarrow$ (iv): trivial

(iv) $\Rightarrow$ (iii): Sei $d > 0$. Dann existieren Vektoren $\underline{b}' \in K^n$, sodass das Gleichungssystem $A'\underline{x} = \underline{b}'$ keine Lösung hat (1.10). Da A aus A' durch Umformungen Z_1, Z_2, Z_3 hervorgeht, folgt aus 1.7, dass dann das Gleichungssystem $A\underline{x} = \underline{b}$ für gewisse Vektoren $b \in K^m$ keine Lösung besitzt, im Widerspruch zu (iv). Also ist $d = 0$ und $A' = \mathbf{1}_n$. Es folgt (iii).

(iii) $\Rightarrow$ (v): Wir wenden 2.8 an. Zu der Umkehrabbildung α^{-1} gehöre die Matrix $B \in K^{n\times n}$. Dann ist $BA = AB = \mathbf{1}_n$.

(v) $\Rightarrow$ (i): trivial □

Definition: Eine Matrix $A \in K^{n\times n}$ heißt *invertierbar*, wenn eine Matrix $B \in K^{n\times n}$ existiert, sodass $BA = AB = \mathbf{1}_n$. Die Matrix B heißt zu A *invers* und wird mit A^{-1} bezeichnet.

Die Eindeutigkeit von A folgt insbesondere aus:

2.12 *Sei $A \in K^{n\times n}$ und seien $B, C \in K^{n\times n}$ nicht $CA = AB = \mathbf{1}_n$. Dann ist A invertierbar und $B = C = A^{-1}$.*

Beweis: $B = (CA)B = C(AB) = C$. □

Sei $A \in K^{n\times n}$ invertierbar. Wir haben schon in 2.10 gelernt, wie man A^{-1} **berechnet**: Man bilde die erweiterte Koeffizientenmatrix $(A, \underline{b})$ mit einem "allgemeinen" Vektor $\underline{b} \in K^n$ und überführe diese mit Z_1, Z_2, Z_3 in eine Matrix (A', b'), sodass A' Stufenform hat. Dabei entsteht eine Matrix $B \in K^{n\times n}$ mit $b' = Bb$. Da A invertierbar ist, hat man $A' = \mathbf{1}_n$. Nach 2.10 ist also $AB = \mathbf{1}_n$. Nach 2.12 ist $B = A^{-1}$.

Wir behandeln nun Abbildungen einer *endlichen* Menge. Stellvertretend für eine solche Menge nehmen wir

$$[n] := \{1, 2, \ldots, n\} \quad , \quad n \in \mathbb{N}.$$

Sei X eine Menge und (für den Augenblick) A die Menge aller Abbildungen

$$\alpha : [n] \to X.$$

Jedes $\alpha \in A$ ist bestimmt durch das n–Tupel

$$(\alpha(1), \ldots, \alpha(n)) \in X^n$$

und jedes $(x_1, \ldots, x_n) \in X^n$ definiert eine Abbildung $\alpha \in A$ durch

$$\alpha(i) := x_i \quad , \quad i = 1, \ldots, n.$$

Man bezeichnet daher die Menge A mit $X^{[n]}$.
Sei $X = \{1, 2\}$. Jedes $\alpha \in A$ ist definiert durch die Angabe von

$$U := \{i \in [n] \mid \alpha(i) = 1\} \qquad (\subseteq [n]),$$

und umgekehrt definiert jede Teilmenge U von $[n]$ ein $\alpha \in A$ durch

$$\alpha(i) := \begin{cases} 0 & i \notin U \\ \quad \text{falls} & \\ 1 & i \in U \end{cases} .$$

Wir schreiben $\alpha = \alpha_U$. Deswegen kann man identifizieren:

- $2^{[n]}$ ist die Menge aller Teilmengen von $[n]$.
- $2^{[n]}$ ist die Menge aller Abbildungen $\alpha : [n] \to \{0, 1\}$.

Insbesondere folgt $|2^{[n]}| = 2^n$.
Sei $X = [n]$. Dann ist $X^{[n]}$ die Menge aller Abbildungen $\alpha : [n] \to [n]$. Wir haben $X^{[n]}$ mit X^n identifiziert, d.h. wir können $\alpha \in X^{[n]}$ mit dem n–Tupel

$$(\alpha(1), \alpha(2), \ldots, \alpha(n)) \quad , \quad \alpha(i) \in [n],$$

identifizieren. Man schreibt

$$\alpha = \begin{pmatrix} 1 & 2 & \ldots & n \\ \alpha(1) & \alpha(2) & \ldots & \alpha(n) \end{pmatrix}.$$

Insbesondere folgt $|X^{[n]}| = n^n$.

2.13 *Für eine Abbildung* $\alpha : [n] \to [n]$ *sind folgende Aussagen äquivalent:*

(i) α *bijektiv*

(ii) α *surjektiv*

(iii) α *injektiv*

Beweis: Die Implikationen (i) $\Rightarrow$ (ii) und (i) $\Rightarrow$ (iii) sind trivial.

(ii) $\Rightarrow$ (i): Verteilt man n Objekte auf n Plätze, sodass jeder Platz belegt ist, so liegt auf jedem Platz genau ein Objekt.

(iii) $\Rightarrow$ (i): Verteilt man n Objekte auf n Plätze, sodass auf jedem Platz höchstens ein Objekt liegt, so liegt auf jedem Platz genau ein Objekt. □

Definition: Eine Abbildung $\alpha : [n] \to [n]$ heißt *Permutation*, wenn α bijektiv ist. Mit S_n bezeichnen wir die Menge aller Permutationen von $[n]$.

Nach 2.13 ist

$$\alpha = \begin{pmatrix} 1 & \dots & n \\ \alpha(1) & \dots & \alpha(n) \end{pmatrix}$$

eine Permutation, wenn alle Zahlen $\alpha(1), \dots, \alpha(n)$ verschieden sind. Daraus folgt

2.14 $|S_n| = n!$ □

Beispiele:

$n = 1 : \ S_n = \{\mathrm{id}\}$

$n = 2 : \ S_n = \{\mathrm{id}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\}$

$n = 3 : \ S_n = \{\mathrm{id}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}\}.$

Definition: Eine Permutation $\tau \in S_n$ heißt *Transposition (Vertauschung)*, wenn $a, b \in [n]$ existieren mit $a \neq b$ und

$$\tau(a) = b, \quad \tau(b) = a, \quad \text{und}$$
$$\tau(i) = i \quad \text{für alle} \quad i \in X \quad \text{mit} \quad i \neq a, b.$$

Wir schreiben

$$\tau = (ab).$$

Offenbar gilt

$$\tau \circ \tau = \mathrm{id}, \quad \text{d.h. } \tau = \tau^{-1}.$$

2.15 Lemma: *Sei $\alpha \in S_n$. Dann existieren Transpositionen $\tau_1, \tau_2, \ldots, \tau_{m-1}, \tau_m \in S_n$ mit*

$$\alpha = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_{m-1} \circ \tau_m.$$

Beweis: Wir beweisen dies durch eine *Induktion* nach n:
Induktionsvoraussetzung:

$n = 1$: Hier ist $S_n = \{\mathrm{id}\}$.

$n = 2$: Hier ist $S_n = \{\mathrm{id}, (12)\}$.

$n = 3$: Zum Beispiel ist $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = (13) \circ (12)$.

Induktionsannahme: In S_{n-1} gelte das Lemma.
Induktionsschluss:
1. Fall $\alpha(n) = n$:
Dann kann α als Permutation der Menge $\{1, 2, \ldots, n-1\}$ aufgefasst werden. Damit folgt die Behauptung aus der Induktionsannahme.
2. Fall $\alpha(n) \neq n$: Sei $a := \alpha(n)$,

$$\tau := (an), \quad \text{und} \quad \alpha' := \tau \circ \alpha.$$

Dann ist

$$\alpha = \tau \circ \alpha', \quad \text{und} \quad \alpha'(n) = n,$$

es kann also α' wieder als Element von S_{n-1} aufgefasst werden. Die Induktionsannahme liefert Transpositionen $\tau_2, \ldots, \tau_{m'}$ mit

$$\alpha' = \tau_2 \circ \cdots \circ \tau_{m'}.$$

Es folgt

$$\alpha = \tau \circ \tau_2 \circ \cdots \circ \tau_{m'}.$$

Setze $\tau_1 := \tau$ und $m = m' + 1$ $\square$

Beispiele:

$$\begin{aligned} \mathrm{id} &= (12) \circ (12) \\ (12) &= (34) \circ (56) \circ (12) \circ (34) \circ (56) \\ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} &= (13) \circ (12) = (12) \circ (23) = (13) \circ (23) \circ (12) \circ (13). \end{aligned}$$

Definition: Sei $\alpha \in S_n$. Wir nennen α gerade, wenn es eine *gerade* Anzahl m von Transpositionen $\tau_1, \ldots, \tau_m$ gibt, sodass $\alpha = \tau_1 \circ \ldots \circ \tau_m$ in 2.15 gilt; andernfalls nennen wir α *ungerade*.

2.16 *Seien $\alpha, \beta \in S_n$.*

a) α ist gerade $\Rightarrow \alpha^{-1}$ ist gerade.

b) α und β sind gerade $\Rightarrow$ $\alpha \circ \beta$ ist gerade.

c) α und β sind ungerade $\Rightarrow \alpha \circ \beta$ ist gerade.

d) α ist gerade und β ungerade $\Rightarrow \alpha \circ \beta$ ist ungerade.

Beweis: Sei α wie in 2.15 und β ein Produkt von l Transpositionen $\tau_1', \ldots, \tau_l'$

$$\beta = \tau_1' \circ \cdots \circ \tau_l'.$$

Dann ist

$$\alpha \circ \beta = \tau_1 \circ \cdots \circ \tau_m \circ \tau_1' \circ \cdots \circ \tau_l$$

ein Produkt von $m + l$ Transpositionen,

a) $\alpha = \tau_1 \circ \tau_2 \circ \ldots \circ \tau_m \overset{!}{\Rightarrow} \tau_m \circ \ldots \circ \tau_2 \circ \tau_1 \circ \alpha = \mathrm{id} \Rightarrow \alpha^{-1} = \tau_m \circ \tau_2 \circ \ldots \circ \tau_1$

b) Sind m und l gerade Zahlen, so ist $m + l$ eine gerade Zahl.

c) Sind m, l ungerade Zahlen, so ist $m + l$ eine gerade Zahl.

d) Hier ist m gerade und l gerade. Ist die Behauptung falsch, so ist $\alpha \circ \beta$ ungerade; es existiert also eine gerade Anzahl k von Transpositionen $\varphi_1, \ldots, \varphi_k$ mit

$$\alpha \circ \beta = \varphi_1 \circ \cdots \circ \varphi_k.$$

Es folgt

$$\tau_1 \circ \cdots \circ \tau_m \circ \tau_1' \circ \cdots \circ \tau_l' = \varphi_1 \circ \cdots \circ \varphi_k,$$

also

$$\beta = \tau_1' \circ \cdots \circ \tau_l' = \tau_m \circ \cdots \circ \tau_1 \circ \varphi_1 \circ \cdots \circ \varphi_k.$$

Damit ist β ein Produkt von $m + k$ vielen Transpositionen, also eine gerade Permutation, da $m + k$ gerade ist. Dieser Widerspruch beweist d). □

Wir definieren nun das *Signum* einer Permutation α durch

$$\operatorname{sgn} \alpha = \begin{cases} +1 & \text{falls } \alpha \text{ gerade} \\ -1 & \text{falls } \alpha \text{ ungerade} \end{cases}$$

und können 2.16 nun so formulieren:

2.17 *sgn* $(\alpha \circ \beta) = (\textit{sgn}\,\alpha)(\textit{sgn}\,\beta)$.

2.18 Hilfssatz: *Sind* $\tau_1, \ldots, \tau_m$ *Transpositionen mit*

$$\tau_1 \circ \cdots \circ \tau_m = \mathrm{id},$$

so ist m *gerade.*

Wir beweisen diesen Hilfssatz am Ende dieses Abschnitts und leiten aus ihm zunächst eine wichtige Konsequenz ab.

2.19 Satz: *Sei* $\alpha \in S_n$ *und seien* $\tau_1, \ldots, \tau_m \in S_n$ *Transpositionen mit*

$(*)$ $\quad \alpha = \tau_1 \circ \cdots \circ \tau_m$.

Dann gilt:

α *ist gerade* $\Leftrightarrow$ m *ist eine gerade Zahl*

α *ist ungerade* $\Leftrightarrow$ m *ist eine ungerade Zahl.*

Also ist sgn $\alpha = (-1)^m$.

Beweis: Die „ungerade“ Behauptung folgt aus der „geraden“.

Ist m gerade, so ist α nach Definition eine gerade Permutation.

Sei α nun eine gerade Permutation. Um zu zeigen, dass m gerade ist, nehmen wir an, dass m ungerade ist und führen dies zu einem Widerspruch: Da α gerade ist, existiert eine gerade Anzahl k von Transpositionen $\varphi_1, \ldots, \varphi_k$ mit

$$\alpha = \varphi_1 \circ \cdots \circ \varphi_k.$$

Es ist also

$$\tau_1 \circ \cdots \circ \tau_m = \varphi_1 \circ \cdots \circ \varphi_k,$$

oder

$$\varphi_1 \circ \cdots \circ \varphi_k \circ \tau_m \circ \cdots \circ \tau_1 = \mathrm{id}.$$

Der Hilfssatz 2.18 besagt nun, dass $k+m$ gerade ist, ein Widerspruch, denn wir hatten angenommen, dass m ungerade ist. □

2.20 Korollar: *Eine Transposition τ ist ungerade, d.h. sgn $\tau = -1$. Für $\alpha \in S_n$ folgt sgn $(\tau \circ \alpha) = (\text{sgn}\,\tau)(\text{sgn}\,\alpha) = -\text{sgn}\,\alpha$.* □

Notation: A_n sei die Menge aller geraden Permutationen in S_n.

Also ist die Differenzmenge

$$S_n \backslash A_n := \{\alpha \in S_n \mid \alpha \notin A_n\}$$

die Menge aller ungeraden Permutationen.

2.21 *Sei $\tau \in S_n$ eine Transposition. Dann ist die Abbildung*

$$\varphi : \ S_n \to S_n \quad \textit{mit} \quad \alpha \mapsto \tau \circ \alpha$$

eine Bijektion, die die geraden auf die ungeraden Permutationen und ungeraden auf die geraden Permutationen abbildet. Insbesondere existieren gleich viele gerade wie ungerade Permutationen in S_n.

Beweis: Wegen 2.20 ist nur zu zeigen, dass φ bijektiv ist.
φ ist injektiv:

$$\tau \circ \alpha = \tau \circ \alpha' \ \rightarrow \ \alpha = \alpha'.$$

φ ist surjektiv: Sei $\beta \in S_n \backslash A_n$ und $\alpha := \tau \circ \beta$. Dann ist $\alpha \in A_n$ und

$$\tau \circ \alpha = \tau \circ \tau \circ \beta = \beta.$$

□

Beweis von 2.16: Wir nehmen an, die Behauptung sei falsch und leiten daraus einen Widerspruch ab. Unsere Annahme lautet also

Es existieren Transpositionen $\tau_1, \ldots, \tau_m$, sodass gilt:

$(+)$ $\quad \mathrm{id} = \tau_1 \circ \cdots \circ \tau_m$ und m ist ungerade.

I. Unter allen Produkte wie in $(+)$ seien $\tau_1, \ldots, \tau_m$ so gewählt, dass m minimal ist. Sei $a \in \{1, \ldots, n\}$ mit $\tau_m(a) \neq a$, also etwa

$$\tau_m = (ax), \quad x \in \{1, \ldots, n\}.$$

II. Unter der Vorgabe I und $a \in \{1, \ldots, n\}$ mit $\tau_m(a) \neq a$ wähle $\tau_1, \ldots, \tau_m$ so, dass die Anzahl der $\tau \in \{\tau_1, \ldots, \tau_{m-1}\}$ mit der Eigenschaft $\tau(a) \neq a$ minimal ist.
Wegen $\mathrm{id}(a) = a$, aber $\tau_m(a) \neq a$ existiert ein $\tau_k \in \{\tau_1, \ldots, \tau_{m-1}\}$ mit $\tau_k(a) \neq a$.

III. Unter all diesen Elementen τ_k sei τ_k so gewählt, dass k maximal ist.

Es ist

$$\tau_k = (ab) \quad \text{für ein} \quad b \neq a.$$

1. Fall $k = m - 1$: Dann ist

$$\mathrm{id} = \tau_1 \circ \cdots \circ \tau_{m-2} \circ (ab) \circ (ax).$$

Ist $b = x$, so ist $(ab) \circ (ax) = \mathrm{id}$, also

$$\mathrm{id} = \tau_1 \circ \cdots \circ \tau_{m-2},$$

im Widerspruch zu I.
Ist $b \neq x$, so ist $(ab) \circ (ax) = (xb) \circ (ab)$ (nachrechnen!), also

$$\mathrm{id} = \tau_1 \circ \cdots \circ \tau_{m-2} \circ (xb) \circ (ab),$$

im Widerspruch zu II.
2. Fall $k < m - 1$: Sei $\tau_{k+1} = (cd)$, also $c \neq a \neq b$. Dann ist

$$\mathrm{id} = \tau_1 \circ \cdots \circ \tau_{k-1} \circ (ab) \circ (cd) \circ \cdots \circ \tau_m.$$

Nun gilt (man rechne nach!):

$$\begin{array}{lcl} b \notin \{c, d\} & \Rightarrow & (ab) \circ (cd) = (cd) \circ (ab) \\ b = c & \Rightarrow & (ab) \circ (cd) = (ab) \circ (bd) = (ad) \circ (ab) \\ b = d & \Rightarrow & (ab) \circ (cd) = (ab) \circ (cb) = (ac) \circ (ab) \end{array}$$

Alle drei Möglichkeiten widersprechen III. □

3 Die Determinante

Wir beginnen mit ein paar Vorbemerkungen.

3.1 *Sei $A \in K^{m\times n}$ und $B \in K^{l\times m}$. Dann ist $(BA)^T = A^T B^T$.*

Beweis: Sei $C = (c_{ij}) = BA$ und $D = (d_{ij}) = A^T B^T$. Dann ist

$$\begin{aligned} c_{ij} &= \sum_{k=1}^{m} b_{ik} a_{kj} \\ d_{ij} &= \sum_{k=1}^{m} a_{ki} b_{jk}. \end{aligned}$$

□

Notation: Die Menge aller invertierbaren Matrizen in $K^{n\times n}$ bezeichnen wir mit $GL(n, K)$.

3.2 *a)* $A \in GL(n, K) \Rightarrow A^{-1} \in GL(n, K)$

b) $A, B \in GL(n, K) \Rightarrow AB \in GL(n, K)$ *und* $(AB)^{-1} = B^{-1}A^{-1}$

c) $A \in GL(n, K) \Rightarrow A^T \in GL(n, K)$ *und* $(A^T)^{-1} = (A^{-1})^T$

Beweis:

a) $A^{-1}A = A^{-1}A = \mathbf{1}_n$

b) $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = \mathbf{1}_n$ und $(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}B = \mathbf{1}_n$

c) Aus $A^{-1}A = \mathbf{1}_n$ folgt mit 3.1 $\mathbf{1}_n = \mathbf{1}_n^T = (A^{-1}A)^T \stackrel{!}{=} A^T(A^{-1})^T$. Also ist A^T invertierbar (beachte 2.12) und $(A^T)^{-1} = (A^{-1})^T$. □

Notation: Sei $A = (a_{ij}) \in K^{n\times n}$.

a) A ist eine *untere Dreieckmatrix*, falls $a_{ij} = 0$, wenn immer $i < j$. A ist *obere Dreiecksmatrix*, wenn A^T eine untere Dreiecksmatrix ist.

b) A ist *Diagonalmatrix*, falls $a_{ij} = 0$, wenn immer $i \neq j$.

c) Ist A Dreiecksmatrix, so sei

$$d(A) := \prod_{i=1}^{n} a_{ii}.$$

d) $[A]_{Z_1}$ bzw. $[A]_{S_1}$ sei die Menge aller Matrizen $A' \in K^{n\times n}$, die aus A durch Umformungen vom Typ Z_1 bzw. vom Typ S_1 hervorgehen.

e) $[A]_D := [A]_{Z_1} \cup [A]_{S_1}$

3.3 *Sei $A \in K^{n\times n}$ Dreiecksmatrix. Dann gilt*

$$A \in GL(n,K) \Leftrightarrow d(A) \neq 0.$$

Beweis: Sei o. E. A eine obere Dreiecksmatrix (andernfalls argumentiere man mit A^T) und sei A' eine Stufenmatrix, die aus A durch Z_1, Z_2, Z_3 hervorgeht. Wir haben gelernt (2.11 und 1.7): Genau dann liegt A in $GL(n,K)$, wenn $A' = \mathbf{1}_n$.

Sei $d(A) = 0$: Dann existiert ein $i \in \{1,\ldots,n\}$ mit $a_{ii} = 0$. Man überlege sich, dass dann A in eine Stufenmatrix umgewandelt werden kann, die eine Nullzeile enthält. Also ist A nicht invertierbar. Sei umgekehrt $d(A) \neq 0$. Dann ist $a_{ii} \neq 0$ für alle $i \in \{1,\ldots,n\}$. Also kann A durch Operatoren Z_1, Z_2, Z_3 in die Einheitsmatrix $\mathbf{1}_n$ übergeführt werden. Es folgt $A \in GL(n,K)$. □

Anhand von Beispielen mache man sich klar:

3.4 *Sei $A \in K^{n\times n}$. In $[A]_{Z_1}$ bzw. $[A]_{S_1}$ liegt eine obere bzw. untere Dreiecksmatrix.* □

Wir geben zwei Beispiele:

1) Sei

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = K^{2\times 2}.$$

Im Falle $(a,c) = (0,0)$ hat

$$A' := A = \begin{pmatrix} 0 & b \\ 0 & d \end{pmatrix}$$

Dreiecksgestalt und es ist $d(A') = 0$.

Fall $a \neq 0$:

$$A \overset{Z_1}{\to} \begin{pmatrix} a & b \\ c - \frac{c}{a}a & d - \frac{c}{a}b \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & d - \frac{c}{b}b \end{pmatrix} =: A'$$

A' ist eine obere Dreiecksmatrix und es ist

$$d(A') = a(d - \frac{c}{a}b) = ad - bc.$$

Im Falle $ad - bc \neq 0$ ist auch $u := d - \frac{c}{a}b \neq 0$, und

$$A' \overset{Z_1}{\to} \begin{pmatrix} a & b - \frac{b}{u}u \\ 0 & u \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & u \end{pmatrix}.$$

Fall $a = 0, c \neq 0$:

$$A \overset{Z_1}{\to} \begin{pmatrix} c & b + d \\ c & d \end{pmatrix} \overset{Z_1}{\to} \to \begin{pmatrix} c & b + d \\ 0 & -b \end{pmatrix} =: A'$$

A' ist eine obere Dreiecksmatrix und es ist $d(A') = -cb = ad - bc$.

2) Sei $A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ mit $a, b \neq 0$.

$$A \overset{Z_1}{\to} \begin{pmatrix} a & 0 \\ 1 & b \end{pmatrix} \overset{Z_1}{\to} \begin{pmatrix} 1 & (1-a)b \\ 1 & b \end{pmatrix}$$

$$\overset{Z_1}{\to} \begin{pmatrix} 1 & (1-a)b \\ 0 & b - (1-a)b \end{pmatrix} \overset{Z_1}{\to} \begin{pmatrix} 1 & 0 \\ 0 & ab \end{pmatrix} =: A'.$$

Es ist $d(A) = d(A')$.

Diese zwei Beispiele zeigen auch:

3.5 *Sei $A \in K^{n \times n}$. Ist A invertierbar, so liegt in $[A]_{Z_1}$ und in $[A]_{S_1}$ eine Diagonalmatrix.* □

Da Z_1–Umformungen und S_1–Umformungen rückgängig gemacht werden können, gilt (vergleiche die Ausführungen über Äquivalenzrelation am Ende von 1):

3.6 $[A']_D = [A]_D$ *für* $A' \in [A]_D$. □

3.7 *Seien $A, B \in K^{n\times n}$ und $A' \in [A]_{Z_1}$, $B' \in [A]_{S_1}$. Dann gilt*

a) $[A'B] \in [AB]_{Z_1}$ *und b)* $AB' \in [AB]_{S_1}$.

Beweis: Sei etwa $A' = (a'_{ij})$ von der Form $a'_{1*} = a_{1*} + \lambda a_{2*},\ a'_{i*} = a_{i*},\ i = 2, \ldots, n$ und sei $C = AB,\ C' = A'B$. Dann ist

$$c_{1j} = a_{1*}\ b^T_{*j} \quad , \quad c_{2j} = a_{2*}\ b^T_{*j}$$

und

$$c'_{1j} = (a_{1*} + \lambda a_{2*})\ b^T_{*j} = a_{1*}\ b^T_{*j} + \lambda a_{2*}\ b^T_{*j},$$

also $c'_{1*} = c_{1*} + \lambda c_{2*}$. Durch sukzessive Wiederholung dieser Überlegung erhält man die Aussage a). Die Aussage b) folgt analog. □

3.8 Satz: *Sei $A \in K^{n\times n}$ und seien A', A'' zwei Dreiecksmatrizen aus $[A]_D$. Dann ist $d(A') = d(A'')$.*

Den nichttrivialen Beweis dieses Satzes stellen wir zurück und ziehen zunächst Folgerungen aus ihm.

Definition: Sei $A \in K^{n\times n}$ und $A' \in [A]_D$ eine Dreiecksmatrix. Die Zahl $d(A')$ heißt die *Determinante* von A. Wir schreiben $\det A := d(A')$.

Nach 3.8 ist $\det A$ wohldefiniert. Zum Beispiel ist $\det A = d(A)$, wenn A eine Dreiecksmatrix ist; insbesondere ist

$$\det(\mathbf{1}_n) = 1.$$

Wir kennen auch schon die Determinante einer Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in K^{2\times 2}$:

$$\det A = ad - bc$$

Als erstes geben wir eine **geometrische** Interpretation der Determinante:

Seien $a_{1*}, \ldots, a_{n*}$ Vektoren in K^n (Zeilen der Matrix A). Die Menge

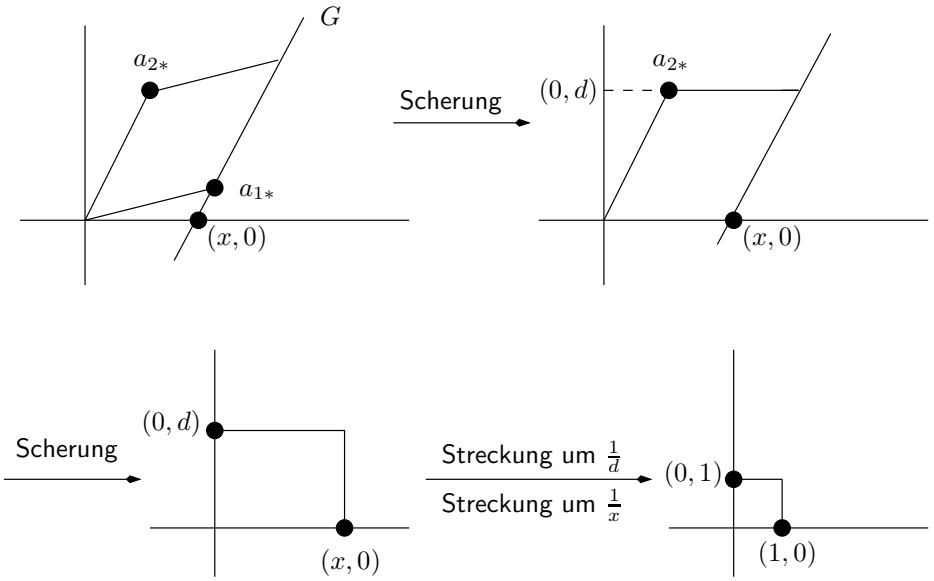
$Q := Q_A := \{\sum_{i=1}^n \lambda_i a_{i*} \mid 0 \le \lambda_i \le 1\}$ [8]

heißt im Fall $n = 2$ *Parallelogramm*, im Fall $n = 3$ *Parallelotop*; für beliebiges n

[8] Achtung: Eine Ungleichung ist nur in $K = \mathbb{R}$ erlaubt!

nennen wir Q einen *Quader*. Wir interpretieren $\det A$ als *Volumen* V von Q und führen dies beispielhaft im Fall $n = 2$ vor.

Sei $n = 2,\ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, also $a_{*1} = (a, b),\ a_{*2} = (c, d)$.



Es folgt $V = dx$. Es ist

$$G = Ka_{2*} + a_{1*} = \{(\lambda c + a, \lambda d + b) \mid \lambda \in K\}.$$

Aus $\lambda d + b = 0$ folgt $\lambda = -\frac{b}{d}$. Also ist $x = -\frac{b}{d}c + a$ und $V = ad - bc = \det A$.

Vertauscht man hier die Vektoren a_{1*}, a_{2*}, so folgt mit derselben Rechnung $\det A = -(ad - bc)$. Man merke sich:

- *Die Determinante misst das Volumen bis auf ein Vorzeichen.*

Sind die Vektoren $a_{1*}, a_{2*} \neq (0, 0)$, und liegen sie auf einer Geraden, welche durch den Nullpunkt geht, so ist $a_{1*} = \lambda a_{2*}$ mit $\lambda \neq 0$, d.h. $a = \lambda c,\ b = \lambda d$. Es folgt

$$\det A = ad - bc = (\lambda c)\frac{b}{\lambda} - bc = 0,$$

in Übereinstimmung mit der Anschauung.

Folgende zwei Eigenschaften von $\det$ sind fundamental:

3.9 *Seien* $A, B \in K^{n\times n}$.

a) $A \in GL(n, k) \ \Leftrightarrow\ \det A \neq 0$

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

Beweis: a) ergibt sich aus 3.3.

b) Sei $\det A = 0$. Nach a) ist A nicht invertierbar und nach 2.12 ist die Abbildung $\alpha : K^n \to K^n$ mit $\underline{x} \mapsto A\underline{x}$ nicht surjektiv. Sei $\beta : K^n \to K^n$ mit $\underline{x} \mapsto B\underline{x}$. Dann ist auch $\alpha \circ \beta : K^n \to K^n$ nicht surjektiv, also AB nicht invertierbar (2.12 und 2.6). Mit a) folgt $\det(AB) = 0$. Also ist

$$0 = \det(AB) = 0(\det B) = (\det A)(\det B).$$

Ist $\det B = 0$, so folgt analog die Behauptung.

Sei $\det A \neq 0 \neq \det B$. Nach a) sind dann A und B invertierbar. Nach 3.5 existiert eine Diagonalmatrix A' in $[A]_{Z_1}$ und eine Diagonalmatrix $B' \in [B]_{S_1}$. Durch wiederholte Anwendung von 3.7 folgt

$$[A'B'] \in [A]_D.$$

Ist $(a'_{11}, \ldots, a'_{nn})$ die Diagonale von A' und $(b'_{11}, \ldots, b'_{nn})$ die von B', so ist offenbar $(a'_{11}b'_{11}, \ldots, a'_{nn}b'_{nn})$ die von $A'B'$. Es folgt

$$\det AB = \det A'B' = \prod_{i=1}^{n} a'_{ii}b'_{ii} = d(A')d(B') \stackrel{!}{=} (\det A)(\det B).$$

□

Schreibweise: Für $A \in K^{n\times n}$

$$A = \begin{pmatrix} a_{1*} \\ \vdots \\ a_{n*} \end{pmatrix} = (a_{*1}, \ldots, a_{*n}).$$

3.10 *Sei $A \in K^{n\times n}$. Es gelten folgende Determinantenregeln:*

$\mathbf{D}_1$ $\det A = \det A'$ *für* $A' \in [A]_D$.

$\mathbf{D}_2$ $\det A = \det A^T$.

$\mathbf{D}_3$ *Ist eine Spalte oder Zeile von A gleich 0, so ist* $\det A = 0$.

$\mathbf{D}_4$ *Die Matrix A' entstehe aus A, indem man eine Zeile oder Spalte von A mit $\lambda \in K$ multipliziert. Dann ist* $\lambda \det A = \det A'$.

$\mathbf{D}_5$ *Entsteht A aus A' durch Vertauschung zweier Zeilen oder zweier Spalten, so ist* $\det A = -\det A'$.

$\mathbf{D}_6$ *Sind zwei Zeilen (Spalten) von A gleich, so ist* $\det A = 0$.

$\mathbf{D}_7$ $\det(\ldots, \underset{i}{\underline{b} + \underline{c}}, \ldots) = \det(\ldots, \underset{i}{b}, \ldots) + \det(\ldots, \underset{i}{c}, \ldots)$,

und

$$\det \begin{pmatrix} \vdots \\ \underline{b} + \underline{c} \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ \underline{b} \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ \underline{c} \\ \vdots \end{pmatrix}.$$

$\mathbf{D}_8$ *Hat A die Form*

$$A = \left(\begin{array}{c|c} B & * \\ \hline 0 & C \end{array}\right) \quad \textit{mit} \quad B \in K^{m\times m} \quad \textit{und} \quad C \in K^{(n-m)\times(n-m)},$$

so ist $\det A = (\det B)(\det C)$.

Beweis:

$\mathbf{D}_1$) folgt aus 3.6.

$\mathbf{D}_2$) Dies ergibt sich aus der Definition von $\det$.

$\mathbf{D}_3$) Sei etwa $a_{*k} = 0$. Dann berühren Z_1-Umformungen von A die k-te Spalte nicht. Wandelt man also A durch Z_1-Umformungen in eine obere Dreiecksmatrix $A' = (a'_{ij})$, so gilt insbesondere $a'_{kk} = 0$, also $\det A = d(A') = 0$.

$\mathbf{D}_4$) B entstehe aus A, indem man etwa die j-te Spalte a_{*j} durch λa_{*j} ersetzt. Sei zunächst $\lambda = 0$, also $\lambda a_{*j} = \mathbf{0}$. Mit $\mathbf{D}_3$ folgt $\det B = 0 = 0 \cdot \det A$. Sei $\lambda \neq 0$. Entsteht B' aus B durch gewisse Z_1-Umformungen, so ist die j-te Spalte von B' das λ-fache der j-ten Spalte derjenigen Matrix A', die aus A durch *dieselben* Z_1-Umformungen entsteht. Es folgt: Ist $A' = (a'_{ij})$ in Dreiecksform, so ist auch $B' = (b'_{ij})$ in Dreiecksform und $b'_{jj} = \lambda a'_{jj}$. Also ist $\det A = d(A') = \lambda d(B') = \lambda \det B$.

$\mathbf{D}_5$) $\det(\dots, a_{*i}, \dots, a_{*j} \dots) =$

$$\overset{D_1}{=} \det(\dots, a_{*i}, \dots, a_{*j} - a_{*i}, \dots) \overset{D_1}{=} \det(\dots, a_{*j}, \dots, a_{*j} - a_{*i}, \dots)$$
$$\overset{D_1}{=} \det(\dots, a_{*j}, \dots, -a_{*i}, \dots) \overset{D_4}{=} (-1)\det(\dots, a_{*j}, \dots, a_{*i}, \dots).$$

$\mathbf{D}_6$) Dies folgt aus $\mathbf{D}_5$.

$\mathbf{D}_7$) Wegen $\mathbf{D}_5$ können wir $i = n$ annehmen. Es sei also

$$a_{*n} = \underline{b} + \underline{c}.$$

Sei $B = (a_{*1}, \dots, a_{*n-1}, \underline{b})$ und $C = (a_{*1}, \dots, a_{*n-1}, \underline{c}) \in K^{n \times n}$.
Durch gewisse Z_1-Umformungen, die nicht die letzte Spalte von A „berücksichtigen“, kann A in eine obere Dreiecksmatrix A' umgewandelt werden. Durch *dieselben* Z_1-Umformungen geht dann B in eine Dreiecksmatrix B' mit der letzten Spalte $\underline{b}'$ und C in eine Dreiecksmatrix C' mit der letzten Spalte $\underline{c}'$ über, sodass

$$a'_{*n} = \underline{b}' + \underline{c}'.$$

Es folgt

$$\begin{aligned}\det A &= d(A') = a'_{11}\dots a'_{n-1,n-1}\ a'_{n,n}\\ &= a'_{11}\dots a'_{n-1,n-1}\ (b'_{n,n}+c'_{n,n})\\ &= a'_{11}\dots a'_{n-1,n-1}\ b'_{n,n} + a'_{11}\dots a'_{n-1,n-1}\ c'_{n,n}\\ &= d(B')+d(C') = \det B + \det C.\end{aligned}$$

$\mathbf{D}_8$) Dies folgt unmittelbar aus der Definition der Determinante. □

Wir fassen $\mathbf{D}_4$ und $\mathbf{D}_7$ zusammen:

3.11 *Die Abbildung* $\det : K^{n\times n} \to K$ *ist* linear *in den Zeilen (Spalten), d.h.*

$$\det\begin{pmatrix}\vdots\\ \lambda\underline{b}+\mu\underline{c}\\ \vdots\end{pmatrix} = \lambda\det\begin{pmatrix}\vdots\\ \underline{b}\\ \vdots\end{pmatrix} + \mu\det\begin{pmatrix}\vdots\\ \underline{c}\\ \vdots\end{pmatrix},\quad \lambda,\mu\in K.$$

3.12 Cramersche Regel: *Sei* $A\in GL(n,K)$. *Dann hat das lineare Gleichungssystem*

$$A\underline{x} = \underline{b}\quad,\quad \underline{x},\underline{b}\in K^n$$

die Lösung $\underline{x} = (x_1,\dots,x_n)$ *mit* [9]

$$x_j = \frac{\det(a_{*1},\dots,a_{*j-1},\underline{b},a_{*j},\dots,a_{*n})}{\det A}\quad,\quad ij = 1,\dots,n.$$

Beweis: Für $j=1,\dots,n$ gilt nach $\mathbf{D}_4$

$$x_j\cdot(\det A) = \det(\dots,x_ja_{*j},\dots) = \det(\dots,-\sum_{i\neq j}x_ia_{*i}+\underline{b},\dots)$$

$$\overset{3.11}{=}\left(\sum_{i\neq j}-x_i\det(\ \dots,\underset{j}{a_{*i}},\dots\ )\right)+\det(\ \dots,\underset{j}{\underline{b}},\dots\ )\ \overset{D_6}{=}\ \det(\ \dots,\underset{j}{\underline{b}},\dots\ ).$$

□

Notation: A_{ij} sei die Matrix aus $K^{(n-1)\times(n-1)}$, welche aus A durch Streichen der i-ten Zeile und j-ten Spalte entsteht.

[9] Beachte 3.9 a).

3.13 $\det(a_{*1}, \ldots, a_{*j-1}, e_i, a_{*j+1}, \ldots, a_{*n}) = (-1)^{i+j} \det A_{ij} \quad , \quad i \in \{1, \ldots, n\}.$

Beweis: Sei $a_{*j} = e_i$. Nach der Spaltenpermutation

$$(21) \circ (32) \circ \cdots \circ (j-1, j-2) \circ (j, j-1)$$

hat A die Form

$$A' = (e_i, a_{*2}, \ldots, a_{*j-1}, a_{*j+1}, \ldots, a_{*n})$$

und nach der Zeilenpermutation

$$(21) \circ (32) \circ \cdots \circ (i-1, i-2) \circ (i, i-1)$$

hat A' die Form

$$A'' = \begin{pmatrix} 1 & a_{i2} \ldots a_{in} \\ 0 & \\ \vdots & A_{ij} \\ 0 & \end{pmatrix}.$$

Aus $\mathbf{D}_8$ folgt

$$\det A'' = 1 \cdot \det A_{ij}.$$

Da wir $(j-1) + (i-1)$ Vertauschungen vorgenommen haben, ergibt sich mit $\mathbf{D}_5$

$$\det A = (-1)^{j-1+i-1} \det A'' = (-1)^{i+j} \det A_{ij}.$$

□

3.14 *Sei A invertierbar und $\tilde{A} = (\tilde{a}_{ij}) \in K^{n \times n}$ mit*

$$\tilde{a}_{ij} := (-1)^{i+j} \det A_{ji}.$$

Dann ist

$$A^{-1} = \frac{1}{\det A} \tilde{A}.$$

Beweis: Betrachte das Gleichungssystem

$$A\underline{x} = \underline{b} \quad , \quad \text{d.h.} \quad \underline{x} = A^{-1}\underline{b}.$$

Wir wissen: Die i-te Spalte von A^{-1} ist $A^{-1}e_i$. Setzt man also $\underline{b} = e_i$, so ist die zugehörige Lösung $\underline{x} = (x_1, \ldots, x_n)$ die i-te Spalte von A^{-1}. Aus 3.12 folgt:

$$x_j = \frac{1}{\det A} \cdot \det(\ \ldots, e_i, \ldots\ ) \overset{3.13}{=} \frac{1}{\det A} \cdot (-1)^{i+j} \det A_{ij}$$
$$\qquad\qquad\qquad

$$= \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \det \begin{pmatrix} e_{j_1} \\ e_{j_2} \\ a_{3*} \\ \vdots \\ a_{n*} \end{pmatrix} = \dots$$

$$= \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \cdots \sum_{j_n=1}^{n} a_{1,j_1} a_{2,j_2} \dots a_{n,j_n} \det \begin{pmatrix} e_{j_1} \\ \vdots \\ e_{j_n} \end{pmatrix}.$$

Sei

$$\pi : \{1, \dots, n\} \to \{1, \dots, n\} \quad \text{mit} \quad j \mapsto j_i.$$

Ist π keine Permutation, so existieren j_k, j_l mit $j_k = j_l$ und $k \neq l$. Mit $\mathbf{D}_3$ folgt

$$\det \begin{pmatrix} e_{j_1} \\ \vdots \\ e_{j_n} \end{pmatrix} = 0.$$

Ist π eine Permutation, so existieren nach 2.15 Transpositionen $\tau_1, \dots, \tau_m$ mit

$$\pi = \tau_1 \circ \cdots \circ \tau_m.$$

Die Matrix $\begin{pmatrix} e_{j_1} \\ \vdots \\ e_{j_n} \end{pmatrix}$ entsteht daher aus der Einheitsmatrix $\mathbf{1}_n$, indem man nacheinander die Transpositionen $\tau_1, \tau_2, \dots, \tau_m$ auf die Zeilen von $\mathbf{1}_n$ anwendet.
Mit $\mathbf{D}_5$ folgt

$$\det \begin{pmatrix} e_{j_1} \\ \vdots \\ e_{j_n} \end{pmatrix} = (-1)^m \det \mathbf{1}_n = (-1)^m \overset{2.19}{=} \operatorname{sgn} \pi. \qquad \square$$

Zum Schluss dieses Abschnitts beweisen wir nun endlich Satz 3.8, welcher der Ausgangspunkt unserer Determinantenregeln war.

Beweis von 3.8: Wir *definieren* für die Matrix A die Zahl

$$\delta(A) := \sum_{\pi \in S_n} (\operatorname{sgn} \pi) a_{1,\pi(1)} \cdots a_{n,\pi(n)}$$

und zeigen I) und II):

I) $A' \in [A]_D \;\Rightarrow\; \delta(A) = \delta(A')$

II) Ist A eine Dreiecksmatrix, so gilt $\delta(A) = d(A)$.

Seien nun A', A'' Dreiecksmatrizen aus $[A]_D$. Aus I) folgt $\delta(A') = \delta(A'')$ und mit II)

$$d(A') \;=\; \delta(A') \;=\; \delta(A'') \;=\; d(A''),$$

also die Behauptung.

Dem Beweis von I) und II) seien zwei Bemerkungen vorausgeschickt.

(1) $\quad \delta(A) = \delta(A^\top)$.

Dies ergibt sich aus

$$\begin{aligned}
\delta(A) &= \sum_{\pi\in S_n} (\operatorname{sgn}\pi) a_{1,\pi(1)}, \ldots, a_{n,\pi(n)} \\
&= \sum_{\pi\in S_n} (\operatorname{sgn}\pi) a_{\pi^{-1}(1),1}, \ldots, a_{\pi^{-1}(n),n} \\
&\overset{!}{=} \sum_{\pi^{-1}\in S_n} (\operatorname{sgn}\pi^{-1}) a_{\pi^{-1}(1),1}, \cdots, a_{\pi^{-1}(n),n} \\
&= \delta(A^\top).
\end{aligned}$$

(2) Sind zwei Spalten (Zeilen) von A gleich, so ist $\delta(A) = 0$.

Sei dazu etwa $a_{*i} = a_{*j},\ i \neq j$ und $\tau = (ij) \in S_n$. Dann gilt

$$\sum_{\pi\in S_n} (\operatorname{sgn}\pi) a_{1,\pi(1)}, \ldots, a_{n,\pi(n)} = \sum_{\pi\in S_n} (\operatorname{sgn}\pi) a_{1,(\pi\circ\tau)(1)}, \ldots, a_{n,(\pi\circ\tau)(n)}$$

$$\overset{2.20}{=} \sum_{\pi\in S_n} (-\operatorname{sgn}(\pi\cdot\tau)) a_{1,\pi\circ\tau(1)}, \ldots, a_{n,\pi\circ\tau(n)}.$$

Da mit $\pi \in S_n$ auch $\pi \circ \tau$ ganz S_n durchläuft (2.21), folgt $\delta(A) = -\delta(A)$, also (2).

Beweis von I) : Wegen (1) können wir uns auf Z_1–Umformungen beschränken. Es ist

$$a_{1,\pi(1)} \ldots a_{i,\pi(i)} \ldots (a_{j,\pi(j)} + \lambda a_{i,\pi(i)}) \ldots a_{n,\pi(n)}$$

$$= \left(\prod_{l=1}^{n} a_{l,\pi(l)}\right) + \lambda\big(a_{1,\pi(1)} \dots a_{i,\pi(i)} \dots a_{i,\pi(i)} \dots a_{n,\pi(n)}\big),$$

also

$$\delta \begin{pmatrix} \vdots \\ a_{i*} \\ \vdots \\ a_{j*} + \lambda a_{i*} \\ \vdots \end{pmatrix} = \delta(A) + \lambda\delta \begin{pmatrix} \vdots \\ a_{i*} \\ \vdots \\ a_{i*} \\ \vdots \end{pmatrix} \overset{(2)}{=} \delta(A).$$

Beweis von II): Wegen (1) können wir annehmen, dass A eine obere Dreiecksmatrix ist. Sei $\pi \in S_N$, $\pi \neq \mathrm{id}$. Da π eine Permutation ist, existiert ein $i \in \{1, \dots, n\}$ mit $\pi(i) < i$. Dann ist $a_{i,\pi(i)} = 0$, also

$$a_{1,\pi(1)} \dots a_{i,\pi(i)} \dots a_{n,\pi(n)} = 0.$$

Es folgt $\delta(A) = (\mathrm{sgn}\,\mathrm{id}) a_{11} \dots a_{nn} = d(A)$. □

4 Der Vektorraum

Definition:

a) Sei V eine nichtleere Menge, in der eine Verknüpfung "$+$" erklärt ist, d.h. je zwei Elementen v, w ist ein Element $x \in V$ zugeordnet – wir schreiben $x = v + w$. Dann heißt V eine *abelsche Gruppe*, wenn Folgendes gilt:

$u + v = v + u.$ (Kommutativgesetz)

$u + (v + w) = (u + v) + w.$ (Assoziativgesetz)

Es existiert ein Element $0 \in V$, sodass

$v + 0 = v.$ (neutrales Element)

Es existiert ein Element $-v \in V$, sodass

$v + (-v) = 0.$ (inverses Element)

Wir schreiben $v - u := v + (-u)$. Ist V abelsche Gruppe mit der Verknüpfung "$+$", so schreiben wir auch gerne $V = V(+)$.

b) Sei $V = V(+)$ eine abelsche Gruppe. Auf V sei eine *skalare Multiplikation* mit Zahlen $\lambda \in K$ erklärt, d.h. jedem $v \in V$ und jedem $\lambda \in K$ ist ein Element $x \in V$ zugeordnet – wir schreiben $x = \lambda v$. Dann heißt V Vektorraum über K, wenn folgende Rechenregeln gelten ($v, w \in V,\ \lambda, \mu \in K$).

$$1v = v$$

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

$$\lambda(v + u) = \lambda v + \lambda u$$

$$(\lambda + \mu)v = \lambda v + \mu v.$$

Hauptbeispiel: $V := K^n$. Wir haben in **1** eine Addition und skalare Multiplikation definiert

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n)$$

$$\lambda(x_1, \ldots, x_n) := (\lambda x_1, \ldots, \lambda x_n) \qquad \lambda \in K$$

Dann ist $V(+)$ eine abelsche Gruppe und die skalare Multiplikation genügt den obigen Regeln. Also ist K^n ein Vektorraum.

Im Folgenden sei V ein Vektorraum.

4.1 Rechenregeln:

a) Die Gleichung $x + v = w$ hat in V genau eine Lösung, nämlich $x = w - v$.

b) $0v = 0$ [10]

c) $\lambda 0 = 0$ [11]

d) $(-\lambda v) = -(\lambda v) = \lambda(-v)$. *Insbesondere ist* $(-1)v = -v$

e) Aus $\lambda v = 0$ folgt $\lambda = 0$ oder $v = 0$.

Beweis: a) $(w - v) + v = (w + (-v)) + v = w + ((-v) + v) = w + 0 = w$.

[10] Die linke 0 ist in K und die rechte in V

[11] Beide Nullen liegen in V

b) $0 + 0v = 0v = (0 + 0)v = 0v + 0v$. Nach a) ist $0v = 0$.

c) $\lambda 0 = \lambda(0 + 0) \overset{\mathcal{V}_3}{=} \lambda 0 + \lambda 0$. Man schließe wie in b).

d) $0 = 0v = (\lambda + (-\lambda))v = \lambda v + (-\lambda v)$.

e) Sei $\lambda v = 0$ und $\lambda \neq 0$. Dann existiert $\lambda^{-1} \in K$. Es folgt

$$v \overset{!}{=} 1v = (\lambda^{-1}\lambda)v \overset{!}{=} \lambda^{-1}(\lambda v) = \lambda^{-1}0 \overset{c)}{=} 0.$$

Definition: Eine nichtleere Teilmenge U von V heißt *Unterraum* von V, falls U bez. der in V erklärten Addition und bez. der in V erklärten skalaren Multiplikation wieder ein Vektorraum ist.

4.2 *Ist U eine nichtleere Teilmenge von V, so ist U schon dann Unterraum, wenn gilt*

$$\begin{aligned} u, v \in U &\Rightarrow u + v \in U \\ u \in U,\ \lambda \in K &\Rightarrow \lambda u \in U \end{aligned}$$

Beweis: Es ist U bez. der in V erklärten Verknüpfung "$+$" eine abelsche Gruppe, denn mit $u \in U$ liegt auch $-u = (-1)u$ und $0 = u - u$ in U. Die Regeln der skalaren Multiplikation gelten auch in U, denn sie gelten in V. □

Offenbar sind $U = 0$ [12] und $U = V$ Unterräume von V.

Sei $A \in K^{m\times n}$ und U die Lösungsmenge des homogenen Gleichungssystems $A\underline{x} = \underline{0},\ \underline{x} \in K^n$. Nach 1.5 bzw. 1.1 ist U Unterraum des Vektorraums K^n.

Definition: Seien V und W Vektorräume. Eine Abbildung $\alpha : V \to W$ heißt *linear*, wenn gilt $(u, v \in V,\ \lambda \in K)$

$$\begin{aligned} \alpha(u + v) &= \alpha(u) + \alpha(v) \\ \alpha(\lambda v) &= \lambda\alpha(v) \end{aligned}$$

Ist α bijektiv, so heißt α *Isomorphismus*. In diesem Fall sind V, W *isomorphe* Vektorräume; wir schreiben $V \cong W$.

[12] Eigentlich $U = \{0\}$.

4.3 *Sei $\alpha : V \to W$ linear. Dann ist*

$$\operatorname{Bild}\alpha := \{\alpha(v) \mid v \in V\}$$

Unterraum von W und

$$\operatorname{Kern}\alpha := \{v \in V \mid \alpha(v) = 0\}$$

Unterraum von V.

Beweis: Wir wenden 4.2 an. Seien $w_1, w_2 \in \operatorname{Bild}\alpha$ und seien $v_1, v_2 \in V$ mit $\alpha(v_i) = w_i,\ i = 1, 2$. Dann ist

$$w_1 + w_2 = \alpha(v_1) + \alpha(v_2) \stackrel{!}{=} \alpha(v_1 + v_2) \in \operatorname{Bild}\alpha$$

und

$$\lambda w_1 = \lambda\alpha(v_1) \stackrel{!}{=} \alpha(\lambda v_1) \in \operatorname{Bild}\alpha.$$

Seien $v_1, v_2 \in \operatorname{Kern}\alpha$. Dann ist

$$\alpha(v_1 + v_2) \stackrel{!}{=} \alpha(v_1) + \alpha(v_2) = 0 + 0 = 0,$$

also $v_1 + v_2 \in \operatorname{Kern}\alpha$, und

$$\alpha(\lambda v_1) = \lambda\alpha(v_1) = \lambda 0 = 0,$$

also $\lambda v_1 \in \operatorname{Kern}\alpha$. □

4.4 *Sei $\alpha : V \to W$ linear. Dann gilt*

$$\operatorname{Kern}\alpha = \{0\} \ \Leftrightarrow\ \alpha \textit{ ist injektiv.}$$

Beweis: Ist α injektiv, so folgt die Behauptung aus 2.2. Die umgekehrte Behauptung folgt aus

$$\alpha(v) = \alpha(u) \ \Leftrightarrow\ \alpha(v) - \alpha(u) = 0 \ \Leftrightarrow\ \alpha(v - u) = 0 \ \Leftrightarrow\ v - u \in \operatorname{Kern}\alpha.$$

□

Definition: Seien $v_1, \ldots, v_m$ aus V. Die Menge

$$\langle v_1, \ldots, v_m\rangle := \left\{\sum_{i=1}^{m} \lambda_i v_i \mid \lambda_i \in K\right\}$$

ist die *lineare Hülle (Aufspann, Spann, Erzeugnis)* von $v_1, \ldots, v_n$. Eine Summe $\sum_{i=1}^{m} \lambda_i v_i$ heißt *Linearkombination* von $v_1, \ldots, v_m$.

4.5 a) $\langle v_1, \ldots, v_m \rangle$ *ist Unterraum von* V.

b) $\langle v_1, \ldots, v_m \rangle$ *ist der kleinste Unterraum von* V*, der die Vektoren* $v_1, \ldots, v_n$ *enthält.*

Beweis:

a) Seien $v = \sum_i \lambda_i v_i$ und $u = \sum_i \mu_i v_i$ aus $\langle v_1, \ldots, v_n \rangle$. Dann ist

$$u + v = \left(\sum_i \lambda_i v_i \right) + \left(\sum_i \mu_i v_i \right) \stackrel{!}{=} \sum_i (\lambda_i v_i + \mu_i v_i)$$
$$\stackrel{!}{=} \sum_i (\lambda_i + \mu_i) v_i \in \langle v_1, \ldots, v_m \rangle$$

und

$$\lambda v = \lambda \left(\sum_i \lambda_i v_i \right) \stackrel{!}{=} \sum_i \lambda(\lambda_i v_i) \stackrel{!}{=} \sum_i (\lambda \lambda_i) v_i \in \langle v_1, \ldots, v_m \rangle .$$

b) Sei U Unterraum mit $v_1, \ldots, v_m \in U$. Dann liegen auch die Linearkombinationen $\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n$ in U. □

Definition: V heißt *endlich erzeugt*, wenn es in V endlich viele Vektoren $v_1, \ldots, v_m$ gibt, sodass $V = \langle v_1, \ldots, v_m \rangle$.

Beispiel: 1) $V = K^n$ ist endlich erzeugt, denn $V = \langle e_1, \ldots, e_n \rangle$.

2) Sei X eine nichtleere Menge und $V = K^X$ die Menge aller Abbildungen $f : X \to K$. Definiert man auf V eine (*punktweise*) Addition und skalare Multiplikation durch $(f, g \in V,\ \lambda \in K)$

$$f + g : X \to K \qquad \text{mit} \qquad x \mapsto f(x) + g(x)$$
$$\lambda f : X \to K \qquad \text{mit} \qquad x \mapsto \lambda f(x),$$

so wird V zu einem Vektorraum. Zum Beispiel ist die Nullabbildung (also die Abbildung f mit $f(x) = 0$ für alle $x \in X$) die 0 in V.

Ist X eine endliche Menge, etwa $X = [n],\ n \in \mathbb{N}$, so ist der Vektorraum K^X isomorph zum Vektorraum K^n (siehe **2**).

Ist X eine *unendliche Menge*, d.h. keine endliche Menge, so behaupten wir,

dass der Vektorraum V nicht endlich erzeugt ist. Stellvertretend für eine solche unendliche Menge X argumentieren wir im Fall $X = \mathbb{N}$.

Jedes $f \in K^{\mathbb{N}}$ ist definiert durch eine Folge

$$(x_1, x_2, \ldots, x_i, x_{i+1}, \ldots) = (x_i)_{i\in\mathbb{N}},$$

wobei

$$x_i = f(x_i) \quad , \quad i \in \mathbb{N}.$$

Somit ist $K^{\mathbb{N}}$ der Vektorraum aller Folgen $\underline{x} = (x_i)_{i\in\mathbb{N}}$ mit

$$\begin{aligned} \underline{x} + \underline{y} &= (x_i + y_i)_{i\in\mathbb{N}} \\ \lambda\underline{x} &= (\lambda x_i)_{i\in\mathbb{N}}. \end{aligned}$$

Wir nehmen nun an, dass V endlich erzeugt ist und führen diese Annahme zu einem Widerspruch.

Seien $v_1, \ldots, v_n \in V$ mit $V = \langle v_1, \ldots, v_n\rangle$, und

$$v_j = (a_{ij})_{i\in\mathbb{N}} \quad , \quad j = 1, \ldots, n.$$

Zu jedem $\underline{b} = (b_i)_{i\in\mathbb{N}}$ exisiteren Zahlen $x_1, \ldots, x_n \in K$, sodass

$$\sum_{j=1}^{n} x_j v_j = \underline{b},$$

d.h.

$$\sum_{j=1}^{n} x_j a_{ij} = b_i \quad , \quad i = 1, 2, 3, \ldots$$

Sei $m > n$ und $A = (a_{ij}) \in K^{m\times n}$, wobei $i = 1, \ldots, m$ und $j = 1, \ldots, n$. Dann ist das Gleichungssystem $A\underline{x} = \underline{b}$ für jedes $\underline{b} \in K^m$ lösbar. Sei A' eine Stufenmatrix, die aus A durch Z_1, Z_2, Z_3 hervorgeht. Wegen $m > n$ besitzt A' eine Nullzeile. Es existiert also ein $b' \in K^m$, sodass das Gleichungssystem $A'\underline{x} = b'$ nicht lösbar ist. Da A aus A' durch Z_1, Z_2, Z_3–Umformungen hervorgeht, existiert nach 1.7 ein $\underline{b} \in K^m$, sodass das Gleichungssystem $A\underline{x} = b$ keine Lösung hat – im Widerspruch zur Annahme. Also ist $V = K^{\mathbb{N}}$ nicht endlich

erzeugt.

Im Folgenden sei V Vektorraum.

Definition: a) Vektoren $v_1, \ldots, v_n \in V$ heißen *linear abhängig*, wenn es Skalare $\lambda_1, \ldots, \lambda_n \in K$ gibt, nicht alle $= 0$, sodass

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

b) Vektoren $v_1, \ldots, v_n \in V$ heißen *linear unabhängig*, wenn sie nicht linear abhängig sind, d.h. es gilt die Implikation

$$\sum_{i=1}^{n} \lambda_i v_i = 0 \qquad \Rightarrow \qquad \lambda_1 = \lambda_2 = \ldots = \lambda_n = 0.$$

Z. B. ist $v_1 = 0$ linear abhängig, denn $1 \cdot v_1 = 0$. Ein Vektor $v_1 \neq 0$ ist linear unabhängig.

Seien v_1, v_2 linear abhängig, etwa $\lambda v_1 + \mu v_2 = 0$ mit $\mu \neq 0$. Dann ist

$$v_2 = -\frac{\lambda}{\mu} v_1 \in \langle v_1 \rangle .$$

Seien v_1, v_2, v_3 linear abhängig, etwa $\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3$ mit $\lambda_2 \neq 0$. Dann ist

$$v_2 = -\frac{1}{\lambda_2}(\lambda_1 v_1 + \lambda_3 v_3) = \left(-\frac{\lambda_1}{\lambda_2}\right) v_1 + \left(-\frac{\lambda_3}{\lambda_2}\right) v_3 \in \langle v_1, v_3 \rangle .$$

Auf dieselbe Weise folgt

4.6 *Seien $v_1, \ldots, v_n \in V$, $n > 1$, linear abhängig. Dann existiert ein $i \in \{1, \ldots, n\}$ mit $v_i \in \langle v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n \rangle$.*
Insbesondere ist $\langle v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n \rangle = \langle v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n \rangle$. □

Definition: Vektoren $v_1, \ldots, v_n$ von V bilden eine *Basis* von V, wenn gilt

$v_1, \ldots, v_n$ sind linear unabhängig
$V = \langle v_1, \ldots, v_n \rangle$

4.7 Satz: *Ein endlich erzeugter Vektorraum $V \neq 0$ besitzt eine Basis.*

Beweis: Es existieren $w_1, \ldots, w_m \in V$ mit $V = \langle w_1, \ldots, w_m \rangle$. Sind $w_1, \ldots, w_m$ linear unabhängig, so ist $w_1, \ldots, w_m$ Basis. Andernfalls existiert nach 4.6 ein $w_i \in \{w_1, \ldots, w_m\}$ mit

$$V = \langle w_1, \ldots, w_{i-1}, w_{i+1}, \ldots, w_m \rangle .$$

Sind die Vektoren $w_1, \ldots, w_{i-1}, w_{i+1}, \ldots, w_m$ linear unabhängig, so bilden sie eine Basis. Andernfalls existiert $w_j \in \{w_1, \ldots, w_{i-1}, w_{i+1}, \ldots, w_n\}$, sodass die Vektoren $w_1, \ldots, w_{j-1}, w_{j+1}, \ldots, w_{i-1}, w_{i+1}, \ldots, w_n$ linear abhängig sind. So fortfahrend erhält man eine Basis von V. □

Der Beweis zeigt auch:

4.8 *Seien $v_1, \ldots, v_m \neq 0$ aus V. Dann enthält $v_1, \ldots, v_m$ eine Basis von $\langle v_1, \ldots, v_m \rangle$.* □

Beispiel: Sei $V = K^n$. Jedes $\underline{x} = (x_1, \ldots, x_n) \in K^n$ ist eine Linearkombination

$$\underline{x} = \sum_{i=1}^{n} \lambda_i e_i \quad , \quad \lambda_i := x_i.$$

Also ist $V = \langle e_1, \ldots, e_n \rangle$. Aus $\underline{x} = 0$ folgt $\lambda_i = 0,\ i = 1, \ldots, n$. Also sind die Vektoren $e_1, \ldots, e_n$ linear unabhängig, sie bilden also eine Basis von $V = K^n$.

Als zweites **Beispiel** betrachten wir eine Stufenmatrix.

4.9 *Sei $A \in K^{m \times n}$ Stufenmatrix und $r \in [m]$ die Anzahl der Zeilen $\neq \underline{0}$.*

a) Die Zeilen $a_{1*}, \ldots, a_{r*}$ sind linear unabhängig in K^n.

b) Seien $a_{*j_1}, \ldots, a_{*j_r}$ die Spalten von A, in denen ein Pivot vorkommt. Dann ist $a_{*j_1}, \ldots, a_{*j_r}$ eine Basis des Vektorraums $\langle a_{*1}, \ldots, a_{*n} \rangle$.

Beweis: Man mache sich Beispiele. □

Als drittes **Beispiel** nehmen wir eine Stufenmatrix

$$A = \begin{pmatrix} 1 & 2 & 0 & 3 & 4 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix} \in K^{2 \times 5}.$$

Wir betrachten das Gleichungssystem

$(*) \qquad A\underline{x} = \underline{0} \, , \qquad \underline{x} \in K^n.$

Lösungen sind

$$u_5 = (-4, 0, -1, 0, 1)$$

$$u_4 = (-3, 0, -1, 1, 0)$$

$$u_2 = (-2, 1, 0, 0, 0).$$

Dann ist

$$U =< u_2, u_4, u_5 >$$

der Lösungsraum von $(*)$. Ist z. B. $\underline{x} = (x_1, x_2, x_3, x_4, x_5) \in U$ mit

$$x_2 = 4, \quad x_4 = 2, \quad x_5 = 5,$$

so ist

$$\underline{x} = 4u_2 + 2u_4 + 5u_5 = (-8 - 6 - 20, 4, -5 - 2, 2, 5).$$

Sei nun $A \in K^{m \times n}$ eine beliebige Stufenmatrix und r die Anzahl der von $\underline{0}$ verschiedenen Zeilen, $d := n - r$. Seien $a_{*j_1}, \ldots, a_{*j_d}$ die Spalten von A, die keine Pivotspalten sind. Für $j_i \in \{j_1, \ldots, j_d\}$ bekommen wir eine Lösung u_{j_i} von $(*)$, indem man die freien Variablen $x_{j_1}, \ldots, x_{j_d}$ zu

$$x_{j_l} = 0, \ l \neq i \quad \text{und} \quad x_{j_i} = 1$$

setzt; man nennt diese Lösungen *Fundamentallösungen*.

Man mache sich anhand des Beispiels allgemein klar:

4.10 *Ist A eine Stufenmatrix in $(*)$ und sind $u_{j_1}, \ldots, u_{j_d}$ die Fundamentallösungen von $(*)$, so ist $u_{j_1}, \ldots, u_{j_d}$ eine Basis von U.* □

4.11 Satz: *Sei $v_1, \ldots, v_n$ eine Basis des Vektorraums V.*

a) *Zu jedem $v \in V$ existiert genau ein Tupel $\underline{\lambda} = (\lambda_1, \ldots, \lambda_n) \in K^n$ mit*

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n.$$

b) *Die Abbildung*

$$\kappa := \kappa_{v_1,\ldots,v_n} : V \to K^n \quad \textit{mit} \quad v \mapsto \underline{\lambda}$$

ist ein Isomorphismus von V *auf den Vektorraum* K^n. *Es ist* $\kappa(v_i) = e_i,\ i = 1, \ldots, n$.

Beweis: a) Sei

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

Dann ist

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

Da $v_1, \ldots, v_n$ linear unabhängig sind, folgt $\lambda_i - \mu_i = 0$, d.h. $\lambda_i = \mu_i$ für $i = 1, \ldots, n$.

b) Nach a) ist κ bijektiv. Seien $v, w \in V$ und $\underline{\lambda} = \kappa(v)$, $\underline{\mu} = \kappa(w)$. Dann ist

$$v + w = \left(\sum_i \lambda_i v_i\right) + \left(\sum_i \mu_i v_i\right) \stackrel{!}{=} \sum_i (\lambda_i + \mu_i) v_i,$$

d.h. $\kappa(v + w) = \kappa(v) + \kappa(w)$. Sei $\rho \in K$. Dann ist

$$\rho v = \rho \left(\sum_i \lambda_i v_i\right) = \sum_i \rho(\lambda_i v_i) = \sum_i (\rho \lambda_i) v_i,$$

d.h. $\kappa(\rho v) = \rho \kappa(v)$. □

Sprechweise: Das Tupel $\underline{\lambda} = \kappa_{v_1,\ldots,v_n}(v)$ ist das *Koordinatentupel* von v bez. der Basis $v_1, \ldots, v_n$ und die Zahlen $\lambda_1, \ldots, \lambda_n$ sind die *Koordinaten* von v bez. $v_1, \ldots, v_n$.

Notation: Für $A \in K^{m \times n}$ sei $[A]_Z$ bzw. $[A]_S$ die Menge aller Matrizen $A' \in K^{m \times n}$, die aus A durch Umformungen Z_1, Z_2, Z_3 bzw. S_1, S_2, S_3 hervorgehen.

4.12 Proposition I) *Seien* $a_{1*}, \ldots, a_{m*} \in K^n$. *Wir bilden die Matrix* $A \in K^{m \times n}$ *mit den Zeilen* $a_{1*}, \ldots, a_{m*}$. *Sei* $A' \in [A]_Z$ *und seien* $a'_{1*}, \ldots, a'_{m*}$ *die Zeilen von* A'.

a) $\langle a_{1*}, \ldots, a_{m*} \rangle = \langle a'_{1*}, \ldots, a'_{m*} \rangle$

b) *Die Vektoren $a_{1*}, \ldots, a_{m*}$ sind linear unabhängig, genau dann, wenn es die Vektoren $a'_{1*}, \ldots, a'_{m*}$ sind.*

Beweis: $\mathbf{Z}_1$) Wir addieren das λ–fache der ersten Zeile zur zweiten; also ist

$$a'_{2*} = a_{2*} + \lambda a_{1*} \in \langle a_{1*}, \ldots, a_{m*} \rangle$$

und $\langle a'_{1*}, \ldots, a^*_{m*} \rangle \subseteq \langle a_{1*}, \ldots, a_{m*} \rangle$. Wir addieren das $(-\lambda)$–fache von a'_{1*} $(= a_{1*})$ zu a'_{2*}, also ist

$$a'_{2*} - \lambda a'_{1*} = a_{2*} \in \langle a'_{1*}, \ldots, a'_{m*} \rangle$$

und $\langle a_{1*}, \ldots, a'_{m*} \rangle \subseteq \langle a'_{1*}, \ldots, a'_{m*} \rangle$. Es folgt a). Sei

$$\sum_{i=1}^{m} \lambda_i \; a'_{*i} = 0$$

angenommen, also

$$0 = \lambda_1 a_{*1} + \lambda_2 (a_{2*} + \lambda a_{1*}) + \lambda_3 a_{3*} + \ldots = (\lambda_1 + \lambda_2 \lambda_1) a_{1*} + \lambda_2 a_{2*} + \lambda_3 a_{3*} + \ldots$$

Sind $a_{*1}, \ldots, a_{m*}$ linear unabhängig, so folgt $\lambda_1 + \lambda\lambda_2 = 0 = \lambda_2 = \ldots = \lambda_m$, slso auch $\lambda_1 = 0$. Dies beweist b).

$\mathbf{Z}_2$) Sei $a'_{1*} = \lambda a_{1*},\ \lambda \neq 0$, und $a'_{i*} = a_{i*}$ für $i \neq 1$. Dann ist $a_{1*} = \lambda^{-1} a'_{1*}$, es folgt a). Sei

$$\sum_{i=1}^{m} \lambda_i \; a'_{i*} = 0$$

angenommen, also

$$0 = \lambda_1 (\lambda a_{1*}) + \lambda_2 a^*_{2*} + \cdots$$

Sind $a_{1*}, \ldots, a_{m*}$ linear unabhängig, so folgt $\lambda_1 \lambda = 0 = \lambda_2 = \ldots = \lambda_m$. Wegen $\lambda \neq 0$ ist auch $\lambda_1 = 0$. Dies beweist b).

$\mathbf{Z}_3$) Da in $V(+)$ das Kommutativgesetz gilt, ist die Behauptung in a) und b) trivial. □

Daraus folgt

4.13 *Seien $a_{1*}, \ldots, a_{m*}$ linear unabhängige Vektoren in K^n. Dann ist $m \leq n$.*

Beweis: Wir bilden die Matrix $A \in K^{m \times n}$ mit den Zeilen $a_{*1}, \ldots, a_{*m}$. In $[A]_Z$ existiert eine Stufenmatrix A' (1.8). Da die Zeilen $a'_{1*}, \ldots, a'_{m*}$ nach 4.12 linear unabhängig sind, kann A' keine Nullzeile enthalten. Es folgt $m \leq n$. □

4.14 Proposition II): *Seien $a_{*1}, \ldots, a_{*n}$ Vektoren in K^m. Wir bilden die Matrix $A = (a_{ij}) \in K^{m \times n}$ mit den Spalten $a_{*1}, \ldots, a_{*n}$. Sei $A' \in [A]_Z$ mit den Spalten $a'_{*1}, \ldots, a'_{*n}$. Genau dann sind die Vektoren $a_{*1}, \ldots, a_{*n}$ linear unabhängig, wenn es die Vektoren $a'_{*1}, \ldots, a'_{*n}$ sind.*

Beweis: Seien $a_{*1}, \ldots, a_{*n}$ linear unabhängig.

$\mathbf{Z}_1$) Sei $a'_{2*} = a_{2*} + \lambda a_{1*}$ und $a'_{i*} = a_{i*}$ für $i \neq 2$. Sei

$$\sum_{j=1}^{n} \lambda_j a'_{*j} = \underline{0}$$

angenommen, d.h.

$$\begin{aligned} 0 &= \sum_{j=1}^{n} \lambda_j a_{1j} \\ 0 &= \sum_{j=1}^{n} \lambda_j (a_{2j} + \lambda a_{1j}) \\ 0 &= \sum_{j=1}^{n} \lambda_j a_{ij} \quad , \quad i > 2. \end{aligned}$$

Es folgt

$$0 = \sum_{j=1}^{n} \lambda_j (a_{2j} + \lambda a_{1j}) = \sum_{j=1}^{n} \lambda_j a_{2j} + \lambda \sum_{j=1}^{n} \lambda_j a_{1j} = \sum_{j=1}^{n} \lambda_j a_{2j},$$

also insgesamt $\sum_j \lambda_j a_{*j} = \underline{0}$. Da $a_{*1}, \ldots, a_{*n}$ linear unabhängig sind, erzwingt dies $\lambda_1 = \ldots = \lambda_n = 0$.

Sind $a'_{1*}, \ldots, a'_{*n}$ linear unabhängig, so folgt wegen $A \in [A']_{Z_1}$ genauso, dass $a_{1*}, \ldots, a_{*n}$ linear unabhängig sind.

$\mathbf{Z}_2$) Sei $a'_{1*} = \lambda a_{1*},\ \lambda \neq 0$, und $a'_{i*} = a_{i*}$ für $i \geq 2$. Sei

$$\sum_{i=1}^{n} \lambda_i a'_{i*} = 0$$

angenommen, d.h.

$$\begin{aligned} 0 &= \sum_{j=1}^{n} \lambda_j (\lambda a_{1j}) = \lambda \left(\sum_{j=1}^{n} \lambda_j a_{1j} \right) \\ 0 &= \sum_{j=1}^{n} \lambda_j a_{ij} \quad , \quad i \geq 2. \end{aligned}$$

Wegen $\lambda \neq 0$ folgt $\sum_{j=1}^{n} \lambda_j a_{1j} = 0$, also insgesamt $\sum_{j=1}^{n} \lambda_j a_{*j} = 0$, d.h. $\lambda_1 = \ldots = \lambda_n = 0$.

$\mathbf{Z}_3$) Sei $a'_{1*} = a_{2*},\ a'_{2*} = a_{1*}$ und $a_{i*} = a_{i*}$ für $i \geq 3$. Sei

$$\sum_{j=1}^{n} \lambda_j a'_{*j} = 0$$

angenommen, d.h.

$$\begin{aligned} \sum_{j=1}^{n} \lambda_j a_{2j} &= 0 \\ \sum_{j=1}^{n} \lambda_j a_{1j} &= 0 \\ \sum_{j=1}^{n} \lambda_i a_{ij} &= 0 \quad , \quad i \geq 3. \end{aligned}$$

Also ist $\sum_{j=1}^{n} \lambda_j a_{*j} = \underline{0}$ und $\lambda_1 = \ldots = \lambda_n = 0$. □

Im Folgenden sei V Vektorraum.

4.15 *Seien $u_1, \ldots, u_m \in V$ linear unabhängig und sei u_{m+1} aus V, aber nicht in $\langle u_1, \ldots, u_m \rangle$. Dann sind auch die Vektoren $u_1, \ldots, u_m, u_{m+1}$ linear unabhängig.*

Beweis: Sei

$$\left(\sum_{i=1}^{m} \lambda_i u_i\right) + \lambda_{m+1} u_{m+1} = 0$$

angenommen. Im Fall $\lambda_{m+1} = 0$, $\sum_{i=1}^{n} \lambda_i u_i = 0$, also $\lambda_1 = \ldots = \lambda_n = 0$, da $u_1, \ldots, u_n$ linear unabhängig sind. Also ist $\lambda_{m+1} \neq 0$ und

$$u_{m+1} = -\frac{1}{\lambda_{m+1}} \sum_{i=1}^{n} \lambda_i u_i \in \langle u_1, \ldots, u_n \rangle$$

im Widerspruch zur Voraussetzung. □

4.16 Satz: *Sei V endlich erzeugt und U Unterraum von V.*

a) U ist endlich erzeugt.

b) Sei $v_1, \ldots, v_n$ eine Basis von V und $u_1, \ldots, u_m$ eine Basis von U (4.7). Dann ist $m \leq n$.

Beweis: Wegen 4.10 und 4.11 b) können wir $V = K^n$ annehmen. Sei $0 \neq u_1 \in U$. Dann ist u_1 linear unabhängig. Seien $u_1, \ldots, u_m \in U$ linear unabhängig. Aus 4.13 folgt $m \leq n$. Seien $u_1, \ldots, u_m$ so gewählt, dass m maximal ist. Aus 4.15 folgt $U = \langle u_1, \ldots, u_m \rangle$. Also ist $u_1, \ldots, u_m$ Basis. □

Wir können $U = V$ setzen. Also folgt das

4.17 Korollar I) *Seien $v_1, \ldots, v_n$ und $u_1, \ldots, u_m$ zwei Basen von V. Dann ist $n = m$.* □

Definition: Besitzt V eine Basis $v_1, \ldots, v_n$, so heißt n die *Dimension* von V; wir schreiben $n = \dim V$. Im Fall $V = 0$ sei $\dim V = 0$ gesetzt. Nach 4.17 ist $\dim V$ wohldefiniert.

4.18 Korollar II) *Seien V und W endlich erzeugte Vektorräume. Dann gilt*

$$V \cong W \;\Leftrightarrow\; \dim V = \dim W.$$

Beweis: Nach 4.11 b) können wir $V = K^n$ und $W = K^m$ annehmen. Dann ist $\dim V = n$ und $\dim W = m$. Im Fall $n = m$ ist $V = W$ und im Fall $K^n \cong K^m$ ist $n = \dim K^n = \dim K^m = m$. □

Im Folgenden sei $A = (a_{ij}) \in K^{m\times n}$ mit den Zeilenvektoren $a_{1*}, \ldots, a_{m*} \in K^n$ und Spaltenvektoren $a_{*1}, \ldots, a_{*n} \in K^m$. Wir setzen

$$\begin{aligned} \mathrm{rg_Z}\, A &:= \dim \langle a_{1*}, \ldots, a_{m*}\rangle \quad \text{(Zeilenrang)} \\ \mathrm{rg_S}\, A &:= \dim \langle a_{*1}, \ldots, a_{*n}\rangle \quad \text{(Spaltenrang)} \end{aligned}$$

4.19 *a) Sei $A' \in [A]_Z$ bzw. $A' \in [A]_S$. Dann ist $\mathrm{rg_Z}\, A = \mathrm{rg_Z}\, A'$ bzw. $\mathrm{rg_S}\, A = \mathrm{rg_S}\, A'$.*

b) Sei $A' \in [A]_Z$. Dann ist $\mathrm{rg_S}\, A = \mathrm{rg_S}\, A'$.

Beweis: a) folgt aus 4.12 a).

b) Sei $r := \mathrm{rg_S}\, A$. Nach 4.8 enthält $a_{*1}, \ldots, a_{*n}$ eine Basis von $\langle a_{*1}, \ldots, a_{*n}\rangle$, sei $a_{*j_1}, \ldots, a_{*j_r}$ eine solche. Nach 4.13 sind dann die Spalten $a'_{*j_1}, \ldots, a'_{*j_r}$ von A' linear unabhängig. Bilden sie keine Basis von

$$\langle a'_{*1}, \ldots, a'_{*n}\rangle \overset{4.12a)}{=} \langle a_{*1}, \ldots, a_{*n}\rangle,$$

so existiert nach 4.15 ein

$$a'_{*j_{r+1}} \in \{a'_{*1}, \ldots, a'_{*n}\} \backslash \{a'_{*j_1}, \ldots, a'_{*j_r}\},$$

sodass $a'_{*j_1}, \ldots, a'_{*j_{r+1}}$ linear unabhängig sind. Wegen $A \in [A']_Z$ sind dann auch die Spalten $a_{*j_1}, \ldots, a_{*j_{r+1}}$ linear unabhängig. Dies widerspricht $r = \dim \langle a_{*1}, \ldots, a_{*n}\rangle$. Also ist $a'_{*j_1}, \ldots, a_{*j_r}$ Basis von $\langle a_{*1}, \ldots, a_{*n}\rangle$. Es folgt die Behauptung.

In $[A]_Z$ liegt eine Stufenmatrix A'. Sei r die Anzahl der Zeilen $\neq \underline{0}$ von A'. Nach 4.9 ist

$$\mathrm{rg_Z}\, A' = r = \mathrm{rg_S}\, A'.$$

Also folgt mit 4.19:

4.20 Satz: $\mathrm{rg_Z}\, A = \mathrm{rg_S}\, A$

Definition: Die Zahl $\mathrm{rg_Z}\, A = \mathrm{rg_S}\, A$ heißt der *Rang* $(= \mathrm{rg}\, A)$ von A.

Aus 4.10 folgt

4.21 Satz: *Sei $A \in K^{m\times n}$ und U der Lösungsraum des linearen Gleichungssystems $A\underline{x} = \underline{0}$. Dann ist $\dim U = n - \mathrm{Rang}\, A$.* □

4.22 Korollar: *Seien V, W endlich erzeugte Vektorräume und $\alpha : V \to W$ eine lineare Abbildung. Dann gilt*

$$\dim V = \dim(\operatorname{Kern}\alpha) + \dim(\operatorname{Bild}\alpha) \qquad \textit{(Dimensionsformel)}$$

Beweis: Ist $n = \dim V$ und $m := \dim W$, so können wir nach 4.11

$$V = K^n \quad , \quad W = K^m$$

annehmen, wir haben also eine Situation wie in **2**. Nach 2.4 existiert $A \in K^{m\times n}$ mit

$$\alpha(\underline{x}) = A\underline{x} \quad , \quad \underline{x} \in K^n.$$

Es ist

$$\operatorname{Kern}\alpha = \{\underline{x} \in K^n \mid A\underline{x} = 0\}$$

und

$$\operatorname{Bild}\alpha = \left\{ \sum_{j=1}^{n} x_j a_{*j} \mid x_j \in K \right\}.$$

Also ist

$$\dim(\operatorname{Bild}\alpha) = \dim\langle a_{*1}, \ldots, a_{*n}\rangle = \operatorname{rg}_S A$$

und nach 4.21

$$\dim(\operatorname{Kern}\alpha) = n - \operatorname{rg}_Z A.$$

Mit 4.20 folgt die Behauptung. □

4.23 *Sei V endlich erzeugt und U Unterraum von V. Ist $\dim U = \dim V$, so ist $U = V$.*

Beweis: Wir können wieder $V = K^n$ annehmen. Sei $a_{*1}, \ldots, a_{*n}$ eine Basis von U und sei $A \in K^{n\times n}$ die Matrix mit den Spalten $a_{*1}, \ldots, a_{*n}$. Dann ist

$$\alpha : K^n \to K^n \quad \text{mit} \quad \underline{x} \mapsto A\underline{x}$$

eine lineare Abbildung mit $\operatorname{Bild}\alpha = U$. Nach 4.22 ist $\operatorname{Kern}\alpha = 0$, also α injektiv (4.4). Nach 2.11 ist dann α auch surjektiv, d.h. $U = \operatorname{Bild}\alpha = K^n$. □

Wir sind in 1.1 vom Lösungsraum eines homogen linearen Gleichungssystems ausgegangen. Abschließend können wir feststellen:

4.24 Lemma: *Sei U ein Unterraum von K^n. Dann existiert eine Matrix $A \in K^{m\times n}$, sodass U Lösungsraum des Gleichungssystems $A\underline{x} = \underline{0}$ ist.*

Beweis: Im Fall $U = 0$ ist $A = \mathbf{1}_n$ und im Fall $U = V$ ist $A = \mathbf{O}$ die Nullmatrix. Sei $0 \neq U \neq V$ und $u_1, \ldots, u_k$ Basis von V, also $1 \leq k < n$ (4.23). Sei $B \in K^{k\times n}$ mit den Zeilen $u_1, \ldots, u_k \in K^n$ und $U^\perp$ der Lösungsraum von

$$B\underline{x} = \underline{0}.$$

Da $u_1, \ldots, u_k$ linear unabhängig sind, ist $\operatorname{rg} B = k$. Aus 4.21 folgt $\dim U^\perp = n - k$. Sei $v_1, \ldots, v_{n-k} \in K^n$ eine Basis von $U^\perp$; man nehme etwa die Fundamentallösungen. Sei $m := n - k$ und $A \in K^{m\times n}$ die Matrix mit den Zeilen $v_1, \ldots, v_m$. Dann ist $\operatorname{rg} A = m$. Sei U' der Lösungsraum von $A\underline{x} = \underline{0}$. Nach 4.21 ist $\dim U' = n - m = k$ und nach Konstruktion ist $U \subseteq U'$. Wegen $\dim U = \dim U' = k$ folgt $U = U'$ aus 4.23. □

Beispiel: Sei $n = 4,\ u_1 = (1, 1, 1, 1),\ u_2 = (1, 1, 2, 2)$ und $U = \langle u_1, u_2\rangle$. Wir bilden

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

und berechnen $U^\perp$:

$$B \overset{z_1}{\to} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \overset{z_1}{\to} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Also sind $u_1 = (0, 0, -1, 1),\ u_2 = (-1, 1, 0, 0)$ die Fundamentallösungen. Es ist $m = n - k = 2$ und

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

Lösungen des Gleichungssystems $A\underline{x} = \underline{0}$ sind u_1 und u_2. Also ist U Lösungsraum von $A\underline{x} = \underline{0}$.

5 Gruppen

Sei G eine nichtleere Menge und R eine Vorschrift, die jedem Paar (a, b) von Elementen $a, b \in G$ in eindeutiger Weise ein Element $R(a, b) \in G$ zuordnet.

Man nennt R eine *Verknüpfung*, oder *(zweistellige) Relation* auf G. Oft wird die Verknüpfung R *multiplikativ* geschrieben; man definiert also für $a, b \in G$

$$ab := R(a, b). \qquad \text{(Produkt von } a, b\text{)}$$

Wird R *additiv* geschrieben, so ist

$$a + b := R(a, b). \qquad \text{(Summe von } a, b\text{)}$$

Beispiele:

1.) $G \in \{\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$ und $R(a, b) := a + b$

2.) $G \in \{\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$ und $R(a, b) := ab$

Notation: Gibt es in der Menge X ein Element 0, so sei immer $X^* = \{x \in X \mid x \neq 0\}$.

3.) $G \in \{\mathbb{Z}^*, \mathbb{Q}^*, \mathbb{R}^*\}$ und $R(a, b) := ab$

Hier definiert $R(a, b) := a + b$ keine Verknüpfung *auf* G, denn es ist

$$R(a, -a) = 0 \notin G.$$

4.) $G = K^n$ und $R(a, b) = a + b$

5.) $G = K^{m \times n}$ und $R(A, B) = A + B$ (Matrixaddition)

6.) $G = K^{n \times n}$ und $R(A, B) = AB$ (Matrixprodukt)

7.) $G = GL(n, K)$ und $R(A, B) = AB$

8.) $G = S_n$, und $R(\alpha, \beta) = \alpha \circ \beta$.

9.) Sei X eine Menge und S_X die Menge aller Bijektionen von X nach X.[13] Sei $G := S_X$ und

$$R(\alpha, \beta) = \alpha \circ \beta.$$

[13] $S_n = S_X$ für $X = \{1, \ldots, n\}$

Eine Verknüpfung R auf G heißt *assoziativ*, wenn für alle $a, b, c \in G$ gilt

$$R(R(a,b),c) \;=\; R(a,R(b,c)) \qquad \text{(Assoziativgesetz)}$$

Multiplikativ geschrieben lautet das Assoziativgesetz

$$(ab)c \quad = \quad a(bc)$$

Man setzt fest

$$abc \;:=\; (ab)c \;=\; a(bc)$$

Es können also mehrere Elemente von G verknüpft werden. Seien $a_1, a_2, \ldots, a_{n-1}, a_n \in G$ und sei das „Produkt“ $a_1 a_2 \ldots a_{n-1}$ bereits definiert. Dann definiert man

$$a_1 a_2 \ldots a_{n-1} a_n := (a_1 a_2 \ldots a_{n-1}) a_n.$$

Aus dem Assoziativgesetz folgt nun, dass jede sinnvolle Klammerung des Ausdrucks $a_1 a_2 \ldots a_{n-1} a_n$ dasselbe Element ergibt.

Beispiel:

$$a((bc)d) = (a(bc))d = ((ab)c)d = (abc)d = abcd$$

Insbesondere lassen sich die Produkte eines Elements $a \in G$ mit sich bilden. Man definiert

$$a^1 \;:=\; a, \;\; a^2 := aa, \;\; a^3 := a^2 a = aaa, \ldots.$$

$$a^n \;:=\; (a^{n-1})a = \underbrace{aa\ldots a}_{n} \qquad (n \in \mathbb{N}) \qquad \text{(Potenz von } a\text{)}$$

Es folgen die *Potenzgesetze* $(n, m \in \mathbb{N})$

$$a^{n+m} \;=\; a^n a^m$$

$$(a^n)^m \;=\; a^{nm}$$

Schreibt man die Verknüpfung R additiv, so definiert man anstatt a^n

$$na := \underbrace{a + \cdots + a}_{n}. \qquad \text{(}\textit{Vielfaches}\text{ von } a\text{)}$$

Die Potenzgesetze lauten dann:

$$\begin{aligned}(n+m) &= na+ma \\ m(na) &= (mn)a\end{aligned}$$

Definition: Sei G eine nichtleere Menge mit einer assoziativen Verknüpfung R.

a) Ein Element $e \in G$, für das gilt

$$R(a,e) = R(e,a) = a \quad \text{für alle} \quad a \in G,$$

heißt ein *neutrales Element* von G.

b) Besitzt G ein neutrales Element, so heißt G *Monoid*.

Bemerkung: In einem Monoid G existiert genau ein neutrales Element. Hat nämlich $e' \in G$ dieselbe Eigenschaft wie e, so gilt

$$e' = ee' = e.$$

In einem multiplikativ geschriebenen Monoid wird e oft mit 1 und in einem additiv geschriebenen Monoid wird e oft mit 0 bezeichnet.

Definiert man in einem multiplikativen Monoid G

$$a^0 := e \quad , \quad 0 \in \mathbb{N},$$

und in einem additiven Monoid G

$$0a := e \quad , \quad 0 \in \mathbb{N},$$

so gelten die Potenzgesetze nun für alle Zahlen $n, m \in N_0$.

5.1 *Sei G ein (multiplikativ geschriebenes) Monoid. Folgende zwei Aussagen sind äquivalent.*

(i) Seien $a, b \in G$. Dann existiert genau ein $x \in G$ und genau ein $y \in G$, sodass

$$ax = b \quad \textit{und} \quad ya = b.$$

(ii) Sei $a \in G$. Es existiert ein $x \in G$ mit $ax = e$ und $xa = e$.

Insbesondere existiert genau ein $x \in G$ mit $ax = xa = e$. Dieses x wird mit a^{-1} (additiv mit $-a$) bezeichnet, und das zu a inverse *Element genannt.*

Beweis: (ii) $\Rightarrow$ (i). Sei auch $x' \in G$ mit

$$ax' = x'a = e.$$

Dann folgt

$$x' = ex' = (xa)x' = x(ax') = xe = x.$$

Also besitzt a genau ein inverses Element $= a^{-1}$. Nun sei $x \in G$ mit $ax = b$. Es folgt

$$\begin{aligned} a^{-1}(ax) &= a^{-1}b \\ (a^{-1}a)x &= a^{-1}b \\ ex &= a^{-1}b \\ x &= a^{-1}b \end{aligned}$$

Umgekehrt ist $x := a^{-1}b$ Lösung von $ax = b$.
Ähnlich ist $y := ba^{-1}$ die eindeutig bestimmte Lösung von $ya = b$.

(i) $\Rightarrow$ (ii) Setzt man $b = e$ so existieren $x, y \in G$ mit

$$ax = e \quad \text{und} \quad ya = e.$$

Es folgt

$$yax = (ya)x = ex = x,$$

und darum

$$ay = aye = (ay)(ax) = a(yax) = ax,$$

also $ax = ay = e$, und deswegen $x = y$. □

Definition: Ein Monoid G heißt *Gruppe*, falls eine der äquivalenten Bedingungen in 5.1 gilt.

Eine Gruppe G heißt *abelsch*, falls zusätzlich noch das *Kommutativgesetz* gilt:

$$ab = ba \quad \text{für alle} \quad a, b \in G.$$

Abelsche Gruppen schreibt man gerne additiv; hier ist also $a + b = b + a$ und man definiert

$$a - b := a + (-b).$$

Sei G eine Gruppe (multiplikativ geschrieben)

1) In G gilt die *Kürzregel*

$$ab = ac \Rightarrow b = c \qquad (a, b, c \in G)$$

$$ba = ca \Rightarrow b = c$$

2) Ist $G = \{a_1, \ldots, a_n\}$ eine *endliche* Gruppe, so lässt sich die Verknüpfung in G durch eine *Verknüpfungstafel* beschreiben. Als Beispiel stellen wir die Verknüpfungstafel für $G = S_3$ auf. Wir haben $S_3 = \{\mathrm{id}, a_1, a_2, a_3, a_4, a_5\}$ mit $a_1 = (123),\ a_2 = (132),\ a_3 = (12),\ a_4 = (13)$ und $a_5 = (2, 3)$. Sei $1 = \mathrm{id}$. An der (i, j)–ten Stelle der folgenden Tabelle steht das Element $a_i a_j$.

	1	a_1	a_2	a_3	a_4	a_5
1	1	a_1	a_2	a_3	a_4	a_5
a_1	a_1	a_2	1	a_4	a_5	a_3
a_2	a_2	1	a_1	a_5	a_3	a_4
a_3	a_3	a_5	a_4	1	a_2	a_1
a_4	a_4	a_3	a_5	a_1	1	a_2
a_5	a_5	a_4	a_3	a_2	a_1	1

3) Für $a \in G$ und $n \in \mathbb{N}$ definiert man

$$a^{-n} := (a^n)^{-1} = \underbrace{a^{-1} \ldots a^{-1}}_{n}.$$

Dann gelten die Potenzgesetze

$$\begin{aligned} a^{z_1+z_2} &= a^{z_1} a^{z_2} \\ (a^{z_1})^{z_2} &= a^{z_1 z_2} \end{aligned}$$

für alle $z_1, z_2 \in \mathbb{Z}$.

Definition: Sei U eine nichtleere Teilmenge von G. Dann heißt U *Untergruppe* von G, falls U bez. der in G erklärten Verknüpfung eine Gruppe ist.

Dafür genügt es zu fordern:

$\mathcal{U}_1$ $\quad a, b \in U \quad \Rightarrow \quad ab \in U$ (Abgeschlossenheit)

$\mathcal{U}_2$ $\quad a \in U \quad \Rightarrow \quad a^{-1} \in U.$

Beispiele: $\quad U = \{e\}$ und $U = G$ sind Untergruppen von G.

Definition: Sei X eine nichtleere Teilmenge von G. Die Menge aller Produkte

$$x_1^{z_1}\, x_2^{z_2} \dots x_{z_{n-1}}^{z_{n-1}}\, x_{z_n}, \;\; n \in \mathbb{N}, \; x_1, \dots, x_n \in X, \; z_1, \dots, z_n \in \mathbb{Z},$$

ist das *Erzeugnis* von X und wird mit $< X >$ bezeichnet.

Für $X = \{a\}$ ist

$$< a >:=< X >= \{a^z | z \in \mathbb{Z}\}$$

die Menge aller *Potenzen* von a. Für $X = \{a_1, \dots, a_n\}$ sei

$$< a_1, \dots, a_n >:=< X > .$$

5.2 $\quad < X >$ *ist eine Untergruppe von* G.

Beweis: Seien $x_1^{z_1} \dots x_n^{z_n}$ und $y_1^{z_1'} \dots y_m^{z_m'} \in < X > \quad (y_i \in X, z_1', \dots, z_m' \in \mathbb{Z})$. Dann ist

$$x_1^{z_1} \dots x_n^{z_n} y_1^{z_1'} \dots y_m^{z_m'} \in < X >,$$

und

$$(x_1^{z_1} \dots x_n^{z_n})^{-1} = x_n^{-z_n} \dots x_1^{-z_1} \in < X > .$$

□

Definition: Seien G_1, G_2 zwei Gruppen (mit den Verknüpfungen R_1, R_2). Eine Abbildung $\varphi : G_1 \to G_2$ heißt *Homomorphismus*, falls für alle $a, b \in G_1$ gilt

$$R_2(\varphi(a), \varphi(b)) = \varphi(R_1(a, b)),$$

d.h., falls R_1 und R_2 multiplikativ

$$\varphi(a)\varphi(b) = \varphi(ab),$$

falls R_1, R_2 additiv

$$\varphi(a) + \varphi(b) = \varphi(a + b)$$

falls R_1 multiplikativ und R_2 additiv

$$\varphi(a) + \varphi(b) = \varphi(ab),$$

falls R_2 multiplikativ, R_1 additiv

$$\varphi(a)\varphi(b) = \varphi(a + b).$$

Ist φ bijektiv, so heißen G_1, G_2 *isomorph*, wir schreiben $G_1 \cong G_2$.

Seien nun G_1, G_2 multiplikativ geschriebene Gruppen und

$$\varphi : G_1 \to G_2$$

ein Homomorphismus. Es sei e_1 das neutrale Element von G_1 und e_2 das neutrale Element von G_2. Wir setzen

$$\begin{array}{lcll} \text{Bild}\,\varphi & := & \{\varphi(a) | a \in G_1\} & (\subseteq G_2) \\ \text{Kern}\,\varphi & := & \{a \in G_1 | \varphi(a) = e_2\} & (\subseteq G_1) \end{array}$$

5.3 *Sei $a \in G$. Für alle Potenzen a^z von a $(z \in \mathbb{Z})$ gilt*

$$\varphi(a^z) = \varphi(a)^z.$$

Insbesondere erhält man für $z = 0$

$$\varphi(e_1) = e_2,$$

und für $z = -1$

$$\varphi(a^{-1}) = \varphi(a)^{-1}.$$

Beweis: Sei zunächst $z = 0$, es ist also $\varphi(e_1) = e_2$ zu zeigen. Doch dies folgt aus der Gleichung

$$\varphi(e_1)\varphi(e_1) \stackrel{!}{=} \varphi(e_1 e_1) = \varphi(e_1)$$

mit Hilfe der Kürzregel in G_2.
Für $z = n \in \mathbb{N}$ ist[14]

$$\varphi(a^n) = \varphi(a^{n-1}a) \stackrel{!}{=} \varphi(a^{n-1})\varphi(a) \stackrel{(')}{=} \varphi(a)^{n-1}\varphi(a) = \varphi(a)^n.$$

Für $z = -n < 0$ ist

$$e_2 = \varphi(e_1) = \varphi(a^{-n}a^n) \stackrel{!}{=} \varphi(a^{-n})\varphi(a^n) = \varphi(a^z)\varphi(a)^n,$$

d.h.

$$\varphi(a^z) = \varphi(a)^{-n} = \varphi(a)^z.$$

□

Zu folgender Aussage vergleiche 4.3:

5.4 *a) Kern φ ist eine Untergruppe von G_1.*
b) Bild φ ist eine Untergruppe von G_2.

Beweis: a) Seien $a, b \in$ Kern φ. Dann ist

$$\varphi(ab) = \varphi(a)\varphi(b) = e_2e_2 = e_2,$$

also $ab \in$ Kern φ. Wegen

$$\varphi(a^{-1}) \stackrel{5.3}{=} \varphi(a)^{-1} = e_2^{-1} = e_2,$$

ist auch $a^{-1} \in$ Kern φ.
b) Seien $\varphi(a), \varphi(b) \in$ Bild $\varphi\,(a, b \in G_1)$. Dann ist

$$\varphi(a)\varphi(b) = \varphi(ab) \in \text{ Bild}\,\varphi,$$

und

$$\varphi(a)^{-1} \stackrel{5.3}{=} \varphi(a^{-1}) \in \text{ Bild}\varphi.$$

Zu folgender Aussage vergleiche 4.4:

[14] Dabei haben wir im Schritt $(')$ eine triviale Induktion nach n benützt.

5.5 *Der Homomorphismus φ ist injektiv genau dann, wenn* $\mathrm{Kern}\ \varphi = \{e_1\}$ *gilt.*

Beweis: Für $a, b \in G_1$ gilt

$$\begin{aligned} \varphi(a) = \varphi(b) &\Leftrightarrow \varphi(a)^{-1}\varphi(b) = e_2 \\ &\overset{5.3}{\Leftrightarrow} \varphi(a^{-1})\varphi(b) = e_2 \\ &\Leftrightarrow \varphi(a^{-1}b) = e_2 \\ &\Leftrightarrow a^{-1}b \in \mathsf{Kern}\varphi. \end{aligned}$$

Im Falle Kern $\varphi = \{e_1\}$ hat man also

$$\varphi(a) = \varphi(b) \Leftrightarrow a^{-1}b = e_1 \Leftrightarrow a = b,$$

d.h. φ ist injektiv. Ist umgekehrt φ injektiv, so ist wegen $\varphi(e_1) = e_2$

$$\varphi(a) = e_2 \Leftrightarrow a = e_1.$$

Dies bedeutet Kern $\varphi = \{e_1\}$. □

Beispiele: 1) Sei G eine Gruppe. Die Abbildungen

$$\begin{aligned} \mathsf{id} &: G \to G \quad \mathsf{mit} \quad a \mapsto a \\ o &: G \to G \quad \mathsf{mit} \quad a \mapsto e \end{aligned}$$

sind (triviale) Homomorphismen.[15] Hierbei gilt

$$\mathsf{Kern}\ \ \mathsf{id} = \{e\} \quad \mathsf{und} \quad \mathrm{Kern}\ o = G.$$

2) Die Menge $\mathbb{Z}$ ist bez. der Addition eine Gruppe $\mathbb{Z}(+)$. Sei G eine weitere (multiplikativ geschriebene) Gruppe und $a \in G$. Sei

$$\varphi : \mathbb{Z}(+) \to G \quad \mathsf{mit} \quad z \mapsto a^z.$$

Die Potenzgesetze besagen

$$\varphi(z_1 + z_2) = a^{z_1+z_2} = a^{z_1}a^{z_2} = \varphi(z_1)\varphi(z_2).$$

Also ist φ Homomorphismus.

[15] o ist die *Nullabbildung*.

3) K^n und K^m sind bez. der Addition Gruppen. Jede lineare Abbildung $\alpha : K^n \to K^m$ ist ein Homomorphismus dieser Gruppen.

4) Für $A \in GL(n, K)$ ist $\det A \neq 0$, d.h. $\det A \in K^*$. Die Abbildung

$$\det : GL_n(K) \to K^* \quad \text{mit} \quad A \mapsto \det A$$

ist ein Homomorphismus, denn 3.9 b) besagt

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

5) Die Menge S_n aller Permutationen von $[n]$ ist bez. der Komposition eine Gruppe (siehe **2**). Die Menge $C := \{1, -1\} \subseteq K^*$ ist eine Untergruppe von K^*. Die Abbildung

$$\text{sgn} : S_n \to C \quad \text{mit} \quad \pi \mapsto \text{sgn}\,\pi$$

ist ein Homomorphismus, denn 2.17 besagt

$$\text{sgn}\,(\pi \circ \rho) = (\text{sgn}\,\pi)(\text{sign}\ \rho).$$

Es ist Kern $(\text{sgn}\,) = A_n$ die Menge aller geraden Permutationen von S_n.

Im Folgenden sind die betrachteten Gruppen *abelsch* (d.h. die Verknüpfung ist kommutativ), wir schreiben sie daher *additiv*. Also ist 0 das neutrale Element, und $-a$ das zu a inverse Element. Man setzt $0a := 0$ und für $n \in \mathbb{N}$

$$na := \underbrace{a + \ldots + a}_{n}$$

$$(-n)a := -(na) \stackrel{!}{=} \underbrace{-a - a \ldots - a}_{n}$$

Dann lauten die Potenzgesetze $(z_1, z_2 \in \mathbb{Z})$

$$\begin{aligned} (z_1 + z_2)a &= z_1 a + z_2 a \\ (z_1 z_2)a &= z_1(z_2 a) \end{aligned}$$

Definition: Eine Gruppe G heißt *zyklisch*, wenn ein $a \in G$ existiert, sodass jedes Element x von G ein Vielfaches $za, z \in \mathbb{Z}$, von a ist (multiplikativ $x = a^z$ ist Potenz von a), d.h. es ist

$$G =< a > .$$

Bemerkung: Wegen $(z_1, z_2 \in \mathbb{Z})$

$$z_1 a + z_2 a = (z_1 + z_2)a = (z_2 + z_1)a = z_2 a + z_1 a$$

ist eine zyklische Gruppe abelsch.

Das Element a heißt *erzeugendes* Element von G.
Die additive Gruppe $\mathbb{Z}(+)$ der ganzen Zahlen ist eine zyklische Gruppe mit dem erzeugenden Element $a = 1$; denn jedes $z = z1$ ist ein Vielfaches von a.

Notation: Für $n \in \mathbb{N}_0$ sei

$$n\mathbb{Z} := \{zn | z \in \mathbb{Z}\}$$

die Menge aller durch n teilbaren Zahlen. Offenbar gilt

- $n\mathbb{Z} = \{0\}$ für $n = 0$
- $n\mathbb{Z} = \mathbb{Z}$ für $n = 1$
- $n\mathbb{Z}$ ist eine Untergruppe von $\mathbb{Z}(+)$
- $n\mathbb{Z}$ ist eine zyklische Gruppe mit dem erzeugenden Element n.

Notation: Für $n \in \mathbb{N}$ sei[16]

$$\mathbb{Z}_n := \{0, 1, 2, \ldots, n-1\} \qquad (\subseteq \mathbb{Z}).$$

Im Folgenden benötigen wir zwei fundamentale Eigenschaften der ganzen Zahlen, die wir – ohne sie zu problematisieren – benutzen.

$\mathcal{N}_{\min}$ *Ist X eine nichtleere Teilmenge von $\mathbb{N}$, so existiert ein $n \in X$ mit $n \leq x$ für alle $x \in X$; wir schreiben*

$$n = \min X.$$

$\mathcal{N}_{\text{rest}}$ *Sei $n \in \mathbb{N}$. Zu jeder Zahl $a \in \mathbb{Z}$ existiert genau ein $\rho(a) \in \mathbb{Z}_{n-1}$ mit folgender Eigenschaft: Es existiert ein $t \in \mathbb{Z}$, sodass gilt:*[17]

$(*)$ $\qquad a = tn + \rho(a).$

[16] also $Z_1 = \{0\}$

[17] Man "teile" a durch n und erhält einen Rest $\rho(a) \in \{0, 1, \ldots, n-1\}$. Im Übrigen ist – wie man sich leicht überlegt – auch die Zahl t durch die Gleichung $(*)$ eindeutig bestimmt.

Offenbar gilt

$$\rho(a) = a \;\Leftrightarrow\; a \in \mathbb{Z}_n.$$

Beispiel: Sei $n = 7$. Dann ist

$$\rho(6) = \rho(-1) = \rho(13) = \rho(-8) = 6,$$

und

$$\rho(-3) = \rho(4) = \rho(11) = 4.$$

5.6 *Sei U eine Untergruppe von $\mathbb{Z}(+)$. Dann existiert ein $n \in \mathbb{N}_0$ mit $U = n\mathbb{Z}$.*

Beweis: Da U Untergruppe von $\mathbb{Z}(+)$ ist, gilt $0 \in U$. Im Falle $U = \{0\}$ ist $U = n\mathbb{Z}$ für $n = 0$. Sei $U \neq \{0\}$. Da mit z auch $-z$ in U liegt, ist

$$X := U \cap \mathbb{N} \neq \emptyset.$$

Wegen $\mathcal{N}_{\min}$ existiert

$$n := \min X.$$

Es folgt

$(')$ $\quad U \cap \mathbb{Z}_n = \{0\}.$

Wegen $n \in U$ ist auch jedes Vielfache zn von n in U, d.h.

$$n\mathbb{Z} \subseteq U.$$

Sei $a \in U$. Ist $\rho(a)$ und t wie in $(*)$ unter $\mathcal{N}_{\text{rest}}$, so gilt wegen $tn \in U$ auch

$$\rho(a) = a - tn \in U,$$

also $\rho(a) \in U \cap Z_n$. Aus $(')$ folgt $\rho(a) = 0$, und daher $a = tn \in n\mathbb{Z}$. Es ist also auch

$$U \subseteq n\mathbb{Z},$$

und deshalb $U = n\mathbb{Z}$. □

Die Abbildung

$$\rho : \mathbb{Z} \to \mathbb{Z}_n \quad \text{mit} \quad a \mapsto \rho(a)$$

nennen wir *Restabbildung modulo* n. Es ist

$$\rho(z) = z \;\Leftrightarrow\; z \in \mathbb{Z}_{n-1} \quad \text{und} \quad \rho(z) = 0 \;\Leftrightarrow\; z \in n\mathbb{Z}.$$

Zwei Zahlen $z_1, z_2 \in \mathbb{Z}$ heißen *kongruent modulo* n ($z_1 \equiv z_2 \pmod n$), wenn $\rho(z_1) = \rho(z_2)$, d.h.

$$z_1 - z_2 \in n\mathbb{Z}.$$

Die Menge $\mathbb{Z}_n$ ist bez. der Addition und Multiplikation in $\mathbb{Z}$ nicht abgeschlossen. Deswegen **definiert** man auf $\mathbb{Z}_n$ eine Addition und Multiplikation modulo n. Wir erklären also zwei Verknüpfungen R_1, R_2 auf $\mathbb{Z}_n$

$$\begin{aligned} R_1(a,b) &= \rho(a+b) \\ R_2(a,b) &= \rho(ab) \end{aligned}$$

Zum Beispiel ist für $n = 3$ die Verknüpfungstafel zu R_1,

	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

und zu R_2

	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Wir schreiben R_1 additiv

$$a \tilde{+} b := R_1(a,b) \qquad (Addition\ modulo\ n)$$

und R_2 multiplikativ

$$a \circ b := R_2(a,b) \qquad (Multiplikation\ modulo\ n)$$

5.7 *Für $a, b \in \mathbb{Z}$ gilt*

a) $\rho(a+b) = \rho(a) \tilde{+} \rho(b)$

b) $\rho(ab) = \rho(a) \circ \rho(b)$

Beweis: Sei $a = tn + \rho(a)$ und $b = sn + \rho(b),\ t, s \in \mathbb{Z}$. Dann ist

$$a + b = (t + s)n + \rho(a) + \rho(b).$$

Dies bedeutet

$$\rho(\rho(a) + \rho(b)) = \rho(a + b).$$

Also ist $\rho(a) \tilde{+} \rho(b) = \rho(a + b)$; dies ist a). Es ist

$$ab = (tn + \rho(a))(sn + \rho(b) = (ts + t\rho(b) + \rho(a)s)n + \rho(a)\rho(b).$$

Daraus folgt

$$\rho(ab) = \rho(\rho(a)\rho(b)).$$

Also ist $\rho(a) \circ \rho(b) = \rho(ab)$; dies ist b). □

5.8 *Bezüglich "$\tilde{+}$" ist $\mathbb{Z}_n$ eine zyklische Gruppe.*

Beweis: Wir wissen, dass $\mathbb{Z}(+)$ eine zyklische Gruppe ist. Via der Restabbildung ρ übertragen wir dies auf $\mathbb{Z}_n$. Dazu benutzen wir 5.7 a).

Assoziativgesetz: Für $a, b, c \in \mathbb{Z}_n$ gilt

$$\begin{aligned} a \tilde{+} (b \tilde{+} c) &= \rho(a) \tilde{+} (\rho(b) \tilde{+} \rho(c)) \stackrel{!}{=} \rho(a) \tilde{+} \rho(b + c) \\ &= \rho(a + (b + c)) \stackrel{!}{=} \rho((a + b) + c) \\ &= \rho(a + b) \tilde{+} \rho(c) = (\rho(a) \tilde{+} \rho(b)) \tilde{+} \rho(c) \\ &= (a \tilde{+} b) \tilde{+} c \end{aligned}$$

Die 0 in $\mathbb{Z}_n$ ist das neutrale Element, und für $0 \neq a \in \mathbb{Z}_n$ ist $n - a$ das zu a inverse Element. Schließlich ist $\mathbb{Z}_n$ zyklisch, denn $1 \in \mathbb{Z}_n$ ist erzeugendes Element.

□

Bemerkung: Die Beziehung $\rho(a + b) = \rho(a) \tilde{+} \rho(b)$ bedeutet nun, dass die Abbildung $\rho : \mathbb{Z} \to \mathbb{Z}_n$ ein Homomorphismus der Gruppe $\mathbb{Z}(+)$ auf die Gruppe $\mathbb{Z}_n$ ist. Es ist $\operatorname{Kern} \rho = n\mathbb{Z}$.

5.9 *Die Menge $\mathbb{Z}_n$ ist bez. "R_2" ein kommutatives Monoid mit neutralem Element* 1.

Beweis: Das Assoziativgesetz folgt mit 5.7 b) aus dem Assoziativgesetz für die Multiplikation in $\mathbb{Z}$:

$$\begin{aligned}(\rho(a) \circ \rho(b)) \circ \rho(c) &= \rho(ab) \circ \rho(c) \\ &= \rho((ab)c) \stackrel{!}{=} \rho(a(bc)) = \rho(a) \circ \rho(bc) \\ &= \rho \circ (\rho(a) \circ \rho(c)).\end{aligned}$$

Die Multiplikation ist kommutativ, denn sie ist in $\mathbb{Z}$ kommutativ:

$$\rho(a) \circ \rho(b) = \rho(ab) = \rho(ba) = \rho(b) \circ \rho(a).$$

Das neutrale Element ist $1 \in \mathbb{Z}_n$. □

5.10 *In $\mathbb{Z}_n$ gelten die Distributivgesetze: Für alle $a, b, c \in Z_n$ gilt*

$$\begin{aligned}a \circ (b \tilde{+} c) &= a \circ b \tilde{+} a \circ c \\ (b \tilde{+} c) \circ a &= b \circ c \tilde{+} c \circ a.\end{aligned}$$

Beweis: Das zweite Distributivgesetz folgt aus dem ersten, da die Multiplikation in $\mathbb{Z}_n$ kommutativ ist. Aus 5.7 folgt

$$\begin{aligned}a \circ (b \tilde{+} c) &= \rho(a) \circ (\rho(b) \tilde{+} \rho(c)) = \\ &= \rho(a) \circ (\rho(b + c)) = \rho(a(b + c)) \\ &\stackrel{!}{=} \rho(ab + ac) = \rho(ab) \tilde{+} \rho(ac) \\ &= \rho(a) \circ \rho(b) \tilde{+} \rho(a) \circ \rho(c) \\ &= a \circ b \tilde{+} a \circ c.\end{aligned}$$

□

6 Ringe und Körper

Definition: Sei R eine nichtleere Menge, auf der zwei Verknüpfungen, nämlich eine Addition und eine Multiplikation, erklärt sind. Dann heißt R ein *Ring*, falls gilt

$\mathcal{R}_1$ Bezüglich der Addition ist R eine abelsche Gruppe (mit neutralem Element 0).

$\mathcal{R}_2$ Bezüglich der Multiplikation ist R ein Monoid (mit neutralem Element 1). Es ist $0 \neq 1$.

$\mathcal{R}_3$ Es gelten in R die *Distributivgesetze*

$$\begin{aligned} a(b+c) &= ab+ac \\ (b+c)a &= ba+ca. \end{aligned}$$

Ist die Multiplikation in R kommutativ, gilt also $ab = ba$ für alle $a, b \in R$, so heißt R *kommutativer* Ring.

6.1 *Sei R ein Ring, dann gilt für alle $a, b \in R$*

a) $0a = a0 = 0$

b) $(-1)a = -a = a(-1)$

c) $(-a)b = -(ab) = a(-b)$

d) $(-a)(-b) = ab$

Beweis: a) Es ist

$$0a = (0+0)a \overset{\mathcal{R}_3}{=} 0a + 0a, \text{ und } a0 = a(0+0) = a0 + a0.$$

Es folgt $0a = a0 = 0$.
b) Es ist

$$(-1)a + a \overset{\mathcal{R}_2}{=} (-1)a + 1a = (-1+1)a = 0a \overset{a)}{=} 0,$$

also $(-1)a = -a$. Ein symmetrisches Argument ergibt $a(-1) = a$.

c) Es ist

$$(-a)b \overset{b)}{=} ((-1)a)b \overset{\mathcal{R}_2}{=} (-1)(ab) \overset{b)}{=} -ab, \text{ und}$$
$$a(-b) \overset{b)}{=} a(b(-1)) \overset{\mathcal{R}_2}{=} (ab)(-1) \overset{b}{=} -ab.$$

d) In der additiven Gruppe $R(+)$ ist

$$(-1) + 1 = 0, \quad \text{also} \quad -(-1) = 1.$$

Es folgt

$$(-a)(-b) \overset{b)}{=} a(-1)(-1)b \overset{b)}{=} a(-(-1))b = a1b = ab.$$

□

Beispiele: **1)** Die Zahlbereiche $\mathbb{Z}, \mathbb{Q}$ und $\mathbb{R}$ sind Ringe.[18]

2) Die Menge $K^{n\times n}$ ist bez. der Matrixaddition und Matrixmultiplikation ein Ring. Das Nullelement von $K^{n\times n}$ ist die Matrix mit lauter Nullen und das Einselement ist die Einheitsmatrix $\mathbf{1}_n$.[19]

Dieser Ring ist für $n \geq 2$ nicht kommutativ.

3) Für $n \in \mathbb{N},\ n \geq 2$ ist $Z_n = \{0, 1, \ldots, n-1\}$ ein kommutativer Ring bezüglich der Addition und Multiplikation modulo n (siehe 5.8, 5.9, 5.10).

Definition: Sei R ein Ring.

a) Ein Element $a \in R$ heißt *Nullteiler* von R, falls $a \neq 0$ und es ein $b \neq 0$ in R gibt mit

$$ab = ba = 0.$$

b) Besitzt R keine Nullteiler, so heißt R *nullteilerfrei,* d.h. für alle $a, b \in R$ gilt

$$ab = 0 \ \Rightarrow \ a = 0 \quad \text{oder} \quad b = 0.$$

Bemerkung: Ist R nullteilerfrei, so ist

$$R^* := \{a \in R \,|\, a \neq 0\}$$

[18] bezüglich der in der Schule geübten Addition und Multiplikation

[19] Der Nachweis der Distributivgesetze ist hier mit einiger Schreibarbeit verbunden.

bez. der in R erklärten Multiplikation ein Monoid.

Definition. Eine Zahl $p \in \mathbb{Z}$ heißt *Primzahl*, wenn $p \neq 1, -1, 0$ und für alle $x, y \in \mathbb{Z}$ gilt

$$xy \equiv 0 \pmod{p} \quad \Rightarrow \quad x \equiv 0 \pmod{p} \quad \text{oder} \quad y \equiv 0 \pmod{p},$$

d.h.

$$\rho(xy) = 0 \quad \Rightarrow \quad \rho(x) = 0 \quad \text{oder} \quad \rho(y) = 0.$$

Vertrauter ist folgende Charakterisierung von Primzahlen.

6.2 *Sei p eine Zahl aus $\mathbb{N}$, verschieden von 1. Folgende zwei Aussagen sind äquivalent:*

(i) *p ist eine Primzahl.*

(ii) *1 und p sind die einzigen Teiler von p, die in $\mathbb{N}$ liegen.*

Beweis: (i) Sei $p = xy$ mit $x, y \in \{2, \ldots, p-1\}$ angenommen. Dann ist $\rho(xy) = 0$ und $\rho(x) \neq 0 \neq \rho(y)$ entgegen der Voraussetzung.

(ii) $\Rightarrow$ (i) Seien $x, y \in \mathbb{Z}$ mit $xy \equiv 0 \pmod{p}$. Dann ist p ein Teiler von xy. Es existiert also ein Teiler x_1 von x und ein Teiler y_1 von y mit $p = x_1 y_1$[20]. Wegen $p > 0$ können wir dabei $x_1, y_1 > 0$ annehmen. Also ist $x_1 = p$ oder $y_1 = p$, und daher $x_1 \equiv 0 \pmod{p}$ oder $y_2 \equiv 0 \pmod{p}$. □

6.3 *Sei $n \in \mathbb{N},\ n > 1$. Genau dann ist der Ring Z_n nullteilerfrei, wenn n eine Primzahl ist.*

Beweis: Seien $a, b \in \mathbb{Z}_n$ mit

$$a \circ b = 0, \text{ d.h.} \quad \rho(ab) = 0.$$

Sei n Primzahl. Dann folgt mit 6.2

$$a = \rho(a) = 0 \quad \text{oder} \quad b = \rho(b) = 0,$$

d.h. Z_n ist nullteilerfrei.

Sei n keine Primzahl. Dann existieren $a, b \in \mathbb{Z}_n$ mit $n = ab$. Dann ist $\rho(ab) = 0$, also $a \circ b = 0$. Also ist a Nullteiler. □

[20] Hier benützen wir die Tatsache, dass sich jede Zahl in eindeutiger Weise als Produkt von Primzahlen schreiben lässt.

6.4 *Sei R ein kommutativer Ring mit endlich vielen Elementen. Ist R nullteilerfrei, so ist das Monoid R^* eine Gruppe.*

Beweis: Seien $a, b \in R^*$. Nach 5.1 genügt es zu zeigen, dass genau ein $x \in R^*$ existiert mit

$$ax = b.$$

Dazu betrachten wir die Abbildung

$$\alpha : R^* \to R^* \quad \text{mit} \quad x \mapsto ax.$$

1) α ist *definiert,* denn R ist nullstellenfrei. Mit a, x liegt also auch ax in R^*.

2) α ist *injektiv:* Seien $x, y \in R^*$ mit $\rho(x) = \rho(y)$, also $ax = ay$. Es folgt

$$0 = ax - ay = a(x - y),$$

und daraus $x - y = 0$, da R nullteilerfrei ist. Es ist also $x = y$.[21]

3) α ist *surjektiv:* Nach Voraussetzung ist $R^* = \{x_1, \ldots, x_n\}$ eine endliche Menge. Wegen 2) sind $\alpha(x_1), \ldots, \alpha(x_n)$ genau n Elemente von R^*. Es folgt $\alpha(R^*) = R^*$.

α ist also *bijektiv*. Danach existiert genau ein $x \in R$ mit $\alpha(x) = b$, also $ax = b$. □

Definition: Sei K ein kommutativer Ring. K heißt *Körper*, wenn K^* eine Gruppe ist.

Zum Beispiel sind $K = \mathbb{Q}$ und $K = \mathbb{R}$ Körper. Aus 6.4 und 6.3 folgt:

6.5 *Der Ring Z_n ist genau dann ein Körper, wenn n eine Primzahl ist.*

Wir werden später noch den Körper $K = \mathbb{C}$ der *komplexen Zahlen* kennen lernen.

In den vorigen Kapiteln haben wir für $K = \mathbb{R}$ nur die Grundrechnungsarten benutzt, die in jedem Körper K gelten. Wir können also feststellen:

> **Alles, was vorher für $K = \mathbb{R}$ formuliert ist, gilt für jeden Körper K.**

[21] In einem nullteilerfreien Ring gilt daher die Kürzregel.

7 Polynome, mehr Körper und Codes

Ein Polynom

$$A = A(X) = a_0 + a_1X + a_2X^2 + \ldots + a_nX^n, \quad a_i \in K$$

ist eine *Rechenvorschrift*, in der man für die *Variable* X Zahlen $\xi \in K$ (u.a.) *einsetzen* kann:

$$A(\xi) := a_0 + a_1\xi + \ldots + a_n\xi^n \ \in K$$

Wir schreiben

$$A = \sum_{i\in\mathbb{N}_0} a_iX^i = \sum_i a_iX^i = \sum a_iX^i,$$

mit der Verabredung, dass nur endlich viele Zahlen a_i, $i \in \mathbb{N}_0$, von 0 verschieden sind.

Achtung: Zu $A \in K[X]$ gehört eine Abbildung (Funktion)

$$\overline{A} : K \to K \quad \text{mit} \quad \xi \mapsto A(\xi).$$

Zum Beispiel ist

$\overline{A}$ *konstant*, wenn $A = a_0$,

$\overline{A}$ *linear*, wenn $A = a_0 + a_1X,\ a_1 \neq 0$,

$\overline{A}$ *quadratisch*, wenn $A = a_0 + a_1X + a_2X^2,\ a_2 \neq 0$.

Ist K ein *endlicher* Körper, so existieren Polynome A, B mit $\overline{A} = \overline{B}$ aber $A \neq B$. Zum Beispiel gilt dies im Fall $K = \mathbb{Z}_2$ für die Polynome

$$A = 1 + X + X^2 \quad \text{und} \quad B = 1,$$

denn es ist $A(0) = 1 = B(0)$ und $A(1) = 1 = B(1)$.

Sei $\lambda \in K$ und $B = \sum b_iX^i$ ein weiteres Polynom. Aus den Rechenregeln, die in K gelten, ergeben sich folgende Beziehungen:

$$\lambda A(\xi) = \lambda a_0 + (\lambda a_1)\xi + (\lambda a_2)\xi^2 + \ldots$$

$$A(\xi) + B(\xi) = (a_0 + b_0) + (a_1 + b_1)\xi + (a_2 + b_2)\xi^2 + \ldots$$

$$A(\xi)B(\xi) = a_0b_0 + (a_0b_1 + a_1b_0)\xi + (a_0b_2 + a_1b_1 + a_2b_0)\xi^2 + \ldots$$

Demzufolge *definiert* man

$$\lambda A := \quad \sum_i (\lambda a_i) X^i$$

$$A + B := \quad \sum_i (a_i + b_i) X^i$$

$$AB := \quad \sum_i c_i X^i, \quad \text{mit} \quad c_i := a_0b_i + a_1b_{i-1} + \ldots + a_{i-1}b_1 + a_ib_0$$

Wir haben also folgende *Einsetzungsregeln*, die wir ständig benutzen werden.

$$(\lambda A)(\xi) = \lambda(A(\xi)), \quad (A + B)(\xi) = A(\xi) + B(\xi), \quad (AB)(\xi) = A(\xi)B(\xi)$$

Sei $K[X]$ die Menge aller Polynome (über K). Bezüglich der definierten Addition und skalaren Multiplikation ist $K[X]$ ein K-Vektorraum. Dabei ist $A = 0$ das *Nullpolynom*, wenn $a_i = 0$ für alle Koeffizienten a_i. Setzt man $X^0 := 1$, so bilden die *Monome* X^i, $i \in \mathbb{N}_0$, eine *Basis* von $K[X]$, d.h. jedes Polynom ist in eindeutiger Weise eine Linearkombination der Monome $X^i,\ i \in \mathbb{N}$. Die definierte Multiplikation von Polynomen macht $K[X]$ außerdem zu einem Ring, der *kommutativ* ist, denn es gilt $AB = BA$. Man hat also auch das Distributivgesetz

$$A(B + C) = AB + AC,$$

und

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

Dabei ist $1 = X^0$ das neutrale Element bez. der Multiplikation.

Wir definieren den *Grad* eines Polynoms $A = \sum_{i=0}^{n} a_i X^i$. Im Falle $a_n \neq 0$ ist $\operatorname{grad} A := n$. Sind alle $a_i = 0$, also $A = 0$, so sei $\operatorname{grad} A := -1$. Man verifiziert sofort folgende Beziehungen:

$\operatorname{grad} \lambda A = \operatorname{grad} A,$ falls $\lambda \neq 0$

$\operatorname{grad}(A + B) \leq$ max $\{\operatorname{grad} A, \operatorname{grad} B\}$, mit Gleichheit, wenn $\operatorname{grad} A \neq \operatorname{grad} B$

$\operatorname{grad} AB = \operatorname{grad} A + \operatorname{grad} B,$ wenn $A \neq 0 \neq B$

Wir nennen diese Beziehungen *Gradformeln* und verwenden sie im Folgenden

meistens ohne besonderen Hinweis.

Ist $A \neq 0$ mit $\operatorname{grad} A = n$, so ist a_n der *Leitkoeffizient* von A. Ist $a_n = 1$, so heißt A *normiert*.

Sei $n \in \mathbb{N}$. Mit $K_n[X]$ bezeichnen wir die Menge aller Polynome vom Grad $\leq n-1$. Offenbar ist $K_n[X]$ ein Unterraum des Vektorraumes $K[X]$ und die Abbildungsvorschrift

$$(a_0 \ldots a_{n-1}) \mapsto \sum_{i=0}^{n-1} a_i X^i$$

erklärt einen Isomorphismus der Vektorräume K^n und $K_n[X]$. Der Basis des K^n, welche die Einheitsvektoren des K^n bilden, entspricht hier die Basis $1, X \ldots X^{n-1}$ von $K_n[X]$. Außer im Falle $n = 1$ ist $K_n[X]$ nicht abgeschlossen bez. der Multiplikation. Sei $n = 1$. Identifiziert man $1 = X^0$ mit dem Einselement $1 \in K$, so ist

$$K_1[X] = \{\lambda X^0 \mid \lambda \in K\} = \{\lambda \in K\} = K$$

Damit ist $K_1[X]$ mit dem Körper K identifiziert. Insbesondere ist $K_1[X]$ abgeschlossen bezüglich der Multiplikation in $\mathbb{F}[X]$. Der Körper K ist also *eingebettet* in $K[X]$. Er besteht aus den Polynomen vom Grad ≤ 0, also aus den *konstanten* Polynomen $(\overline{A} = \text{const})$. Wir machen ein paar Beispiele.

1) Es sei $K := \mathbb{Z}_2 = \{0, 1\}$ (siehe 6.5) und $R := K_2[X]$, d.h.

$$R = \{0, 1, X, X+1\}.$$

Wir setzen $a := X,\ b := X + 1 = a + 1$, also

$$R = \{0, 1, a, b\}.$$

Die Addition in R ist die Addition von Polynomen. Also ist

$$\begin{aligned} 1+1 &= 0 \quad , \quad 1+a = b \quad , \quad 1+b = a \\ a+a &= (1+1)a = 0a = 0 \\ a+b &= a+(a+1) = (a+a)+1 = 1 \\ b+b &= (1+1)b = 0 \end{aligned}$$

R ist Vektorraum über K; die Addition in R haben wir gerade geklärt. Die Skalaren von R sind die Elemente $\lambda = 0,\ \lambda = 1$, also ist $0v = 0$ und $1v = v$ für alle $v \in R$. Es ist $1, a$ eine Basis von R und daher R isomorph zu dem Vektorraum K^2 (4.11). Die Koordinaten $\kappa(v)$ von $v \in R$ bez. der Basis $1, a$ sind

$$\kappa(0) = (0,0) \quad , \quad \kappa(1) = (1,0) \quad , \quad \kappa(a) = (0,1) \quad , \quad \kappa(b) = (1,1).$$

Sei $N := X^2 + 1 \in K[X]$. Wir erklären eine kommutative Multiplikation "$\circ$" *modulo* N auf R: Wir setzen $0 \circ v := 0$ und $1 \circ v := v$ für alle $v \in R$. Es ist

$$\begin{array}{lll} aa = X^2 = N + 1 & \text{, also} & a \circ a := 1 \\ ab = X^2 + X = N + b & \text{, also} & a \circ b := a \\ bb \stackrel{!}{=} X^2 + 1 = N + 0 & \text{, also} & b \circ b := 0 \end{array}$$

Bezüglich der Addition und dieser Multiplikation ist R ein Ring und b ist Nullteiler von R.

Sei $N := X^2 + X + 1$. Wir erklären auf R wieder eine kommutative Multiplikation modulo N: Wir setzen wieder $0 \circ v := 0$ und $1 \circ v := v$ für alle $v \in R$. Es ist

$$\begin{array}{lll} aa = X^2 = N + (X + 1) & \text{, also} & a \circ a := b \\ ab = X^2 + X = N + 1 & \text{, also} & a \circ b := 1 \\ bb = X^2 + 1 = N + X & \text{, also} & b \circ b := a \end{array}$$

Der dadurch definierte Ring R ist also nullteilerfrei und somit Körper (6.3, 6.4). Die multiplikative Gruppe R^* ist zyklisch mit erzeugendem Element a. Man schreibt $R = \mathbb{F}_4$ oder $R = GF(4)$ und nennt R *Galoisfeld*.

Bemerkung: Sei $N = X^2 + 1 \in K[X]$. Dann ist $N = (X + 1)^2$, das Polynom N ist also *zerlegbar*.

Sei $N = X^2 + X + 1$. Dann ist $N(0) = 1 = N(1)$. Das Polynom N hat keine *Nullstelle* in K, und es gibt keine *Faktorisierung* von N.

2) Sei wieder $K = \mathbb{Z}_2$ und nun $R = \mathbb{F}_3[X]$. Dann ist

$$R = \{0, 1, X, X^2, X + 1, X^2 + X, X^2 + X + 1, X^2 + X\}.$$

R ist Vektorraum über K und $1, X, X^2$ ist Basis von R. Bezüglich dieser Basis haben die Elemente von R die Form

$$R = \{(0,0,0), (1,0,0), (0,1,0), (0,0,1), (1,1,0), (0,1,1), (1,1,1), (0,1,1)\}.$$

Sei $N = X^3 + X + 1$ (Es ist $N(0) = N(1) = 1$ und N nicht zerlegbar!).

Wir erklären auf R eine Multiplikation "$\circ$" *modulo* N: Sei $a := X$

$$\begin{aligned} a^2 &= X^2 = 0N + X^2 \quad \text{, also} \quad a \circ a = X^2 \\ a^3 &= X^3 = N + (X+1) \quad \text{, also} \quad a \circ a \circ a = X + 1 \\ a^4 &= X^4 = XN + (X^2 + X) \quad \text{, also} \quad a \circ a \circ a \circ a = X^2 + X \\ a^5 &= X^5 = X^2N + (X^3 + X^2) = X^2N + N + (X^2 + X + 1) \\ &= (X^2+1)N + (X^2 + X + 1), \end{aligned}$$

also $a \circ a \circ a \circ a \circ a = X^2 + X + 1$

$$a^6 = X^3N + (X^4 + X^3) = X^3N + (XN + X^2 + X) = (X^3 + X)N + (X^2 + X),$$

also $a \circ a \circ a \circ a \circ a \circ a = X^2 + X$

$$\begin{aligned} a^7 &= X^4N + (X^5 + X^4) = X^4N + (X^2N + (X^3 + X^2) + X^4) \\ &= X^4N + X^2N + (XN + (X^2 + X) + N + (X+1) + X^2) \\ &= (X^4 + X^2 + X + 1)N + (X^2 + X + X + 1 + X^2) \\ &= (X^4 + X^2 + X + 1)N + 1, \end{aligned}$$

also $a \circ a \circ a \circ a \circ a \circ a \circ a = 1$.

Also ist R^* eine zyklische Gruppe mit a als erzeugendes Element, d.h. R ist Körper; man schreibt $R = \mathbb{F}_8$ oder $R = GF(8) = GF(2^3)$.

3) Sei $K := \mathbb{R}$ und $R := K_2[X]$. Dann ist

$$R = \{a + bX \mid a, b \in \mathbb{R}\}$$

ein Vektorraum mit Basis $1, X$ und R isomorph zum Vektorraum $\mathbb{R}^2$

$$\kappa(a + bX) = (a, b) \in \mathbb{R}^2.$$

Sei $N := X^2 + 1 \in K[X]$. Wir erklären eine Multiplikation "$\circ$" modulo N auf R. Es ist

$$X^2 = N + (-1) \quad \text{, also} \quad X \circ X = -1.$$

Seien $z_1 := a_1 + b_1 X$ und $z_2 := a_2 + b_2 X \in R$. Es ist

$$\begin{aligned} z_1 z_2 &= (a_1 + b_1 X)(a_2 + b_2 X) = a_1 a_2 + (a_1 b_2 + a_2 b_1) X + b_1 b_2 X^2 \\ &= a_1 a_2 + (a_1 b_2 + b_1 a_2) X + b_1 b_2 (N - 1) \\ &= (a_1 a_2 - b_1 b_2) + (a_1 b_2 + b_1 a_2) X + b_1 b_2 N \end{aligned}$$

Setzen wir $i := X$, so folgt

- $z_1 \circ z_2 = (a_1 a_2 - b_1 b_2) + i(a_1 b_2 + b_1 a_2)$

Die Addition in R lautet

- $(a_1 + i b_1) + (a_2 + i b_2) = (a_1 + a_2) + i(b_1 + b_2)$ (*komponentenweise*)

Die Elemente $z = a + ib$ von R heißen *komplexe Zahlen*; i heißt *imaginäre Einheit*, es ist $i^2 = -1$. Die reellen Zahlen

$$a = \mathrm{Re}\, z \quad \text{bzw.} \quad b = \mathrm{Im}\, z$$

heißen *Realteil* bzw. Imaginarteil von z.

Wir setzen $\mathbb{C} := R$ und nennen $\mathbb{C}$ den Körper der *komplexen Zahlen*, denn es gilt

7.1 $\mathbb{C}$ *ist Körper.*

Beweis: Sei $z_1 = a_1 + i b_1 \neq 0$ aus $\mathbb{C}$. Wir zeigen die Existenz eines (zu z_1 inversen) Elements $z_2 = a_2 + i b_2 \in \mathbb{C}$ mit $z_1 \circ z_2 = 1$, d.h.

$$(*) \qquad \begin{aligned} a_1 a_2 - b_1 b_2 &= 1 \\ b_1 a_2 + a\, b_2 &= 0 \end{aligned}$$

Wir betrachten $(*)$ als lineares Gleichungssystem mit Koeffizientenmatrix $\begin{pmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{pmatrix}$ und Unbekannten a_2, b_2. Es ist $\det \begin{pmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{pmatrix} = a_1^2 + b_1^2 > 0$, denn $z_1 \neq 0$. Also existiert genau eine Lösung z_2 von $(*)$. Nach der Cramerschen Regel 3.12

ist

$$a_2 = \frac{\det \begin{pmatrix} 1 & -b_1 \\ 0 & a_1 \end{pmatrix}}{a_1^2 + b_1^2} = \frac{a_1}{a_1^2 + b_1^2}$$

$$b_1 = \frac{\det \begin{pmatrix} a_1 & 1 \\ b_1 & 0 \end{pmatrix}}{a_1^2 + b_1^2} = \frac{-b_1}{a_1^2 + b_1^2}$$

□

Wir setzen diese drei Beispiele in einen allgemeinen Kontext.

Sei $N \in K[X]$ ein Polynom vom Grad $n \geq 1$.

7.2 Division mit Rest: *Sei $A \in K[X]$. Es existieren eindeutig bestimmte Polynome F und R, sodass gilt*

$$A = FN + R \quad , \quad \text{grad } R < n.$$

Beweis: Anstatt eines allgemeinen Beweises erklären wir dies anhand von zwei Beispielen.

1) Sei $K = \mathbb{Z}_2$ und $N = X^3 + X + 1$ sowie $A = X^7$ wie im vorigen Beispiel 2.

$$\begin{array}{l} X^7 : (X^3 + X + 1) = (X^4 + X^2 + X + 1) + 1 \\ \underline{X^7 + X^5 + X^4} \\ \quad\; X^5 + X^4 \\ \quad\; \underline{X^5 + X^3 + X^2} \\ \qquad\; X^4 + X^3 + X^2 \\ \qquad\; \underline{X^4 + X^2 + X} \\ \qquad\quad\; X^3 + X \\ \qquad\quad\; \underline{X^3 + X + 1} \\ \qquad\qquad\quad\; 1 \end{array}$$

Also ist $X^4 = FN + R$ mit $F = X^4 + X^2 + X + 1$ und $R = 1$.

2) Sei $K = \mathbb{Q},\ A = 2X^3 + 4X + 1$ und $N = 3X + 1$.

$$
\begin{array}{l}
(2X^3+4X+1):(3X+1) = \left(\frac{2}{3}X^2+\frac{2}{0}X+\frac{34}{27}\right)+\frac{25}{9} \\
\underline{-2X^2-\frac{2}{3}X^2} \\
\qquad -\frac{2}{3}X^2+4X+1 \\
\qquad \underline{-\frac{2}{3}X^2-\frac{2}{9}X} \\
\qquad\qquad -\frac{34}{9}X+1 \\
\qquad\qquad \underline{-\frac{34}{9}X-\frac{34}{9}} \\
\qquad\qquad\qquad \frac{25}{9}
\end{array}
$$

Also ist $A = FN + R$ mit $F = \left(\frac{2}{3}X^2+\frac{2}{9}X+\frac{34}{25}\right)+\frac{25}{9}$ und $R = \frac{25}{9}$. □

Wie in **5** haben wir also eine Restabbildung modulo N

$$\rho : K[X] \to K_n[X] \quad \text{mit} \quad \rho A = R.$$

7.3 *Die Abbildung ρ ist linear.*

Beweis: Seien $A, A' \in \mathbb{F}[X]$ und sei

$$A = FN + \rho(A) \quad , \quad A' = F'N + \rho(A').$$

Dann ist

$$A + B = (F + F')N + (\rho(A) + \rho(A'))$$

und

$$\operatorname{grad}(\rho(A) + \rho(A')) \leq \max\{\operatorname{grad} A, \operatorname{grad} A\} < n.$$

Also $\rho(A+B) = \rho(A) + \rho(B)$. Es ist $\lambda A = (\lambda F)N + \lambda\rho(A)$ und $\operatorname{grad}\lambda\rho(A) \leq \operatorname{grad}\rho(A)$; es folgt $\rho(\lambda A) = \lambda\rho(A)$.

Analog wie $\mathbb{Z}_n$ machen wir nun $K_n[X]$ zu einem Ring. Die Addition sei die Polynomaddition in $K[X]$; beachte (anders als in $\mathbb{Z}_n$)

$$A, B \in K_n[X] \Rightarrow A + B \in K_n[X].$$

Die Multiplikation "$\circ$" in $K_n[X]$ ist die Multiplikation in $K[X]$ modulo N:

$$A \circ B := \rho(AB)$$

Ganz analog wie $\mathbb{Z}$ in **5** wird dadurch $K_n[X]$ zu einem kommutativen Ring, den wir jetzt mit $K_N[X]$ bezeichnen.

Definition: a) Ein Polynom $N,\ n = \operatorname{grad} N \geq 1$, heißt *zerlegbar*, wenn $A, B \in K[X]$ existieren, sodass

$$N = AB \quad , \quad \operatorname{grad} A,\ \operatorname{grad} B < n.$$

b) Ist N nicht zerlegbar, so heißt N *unzerlegbar* oder *irreduzibel*.

c) N heißt *Primelement*, wenn für $A, B \in K[X]$ gilt

$$\rho(AB) = 0 \qquad \Rightarrow \qquad \rho(A) = 0 \quad \text{oder} \quad \rho(B) = 0.$$

Ist N Primelement, so ist N offenbar unzerlegbar. Die umgekehrte Aussage ist auch richtig, aber etwas umständlich zu beweisen; wir übergehen den Beweis; vergleiche Beweis von 6.2.

Analog zu 6.3 folgt nun

7.4 *Der Ring $K_N[X]$ ist genau dann nullteilerfrei, wenn N ein irreduzibles Polynom ist.* □

Daraus folgt

7.5 Satz: *Der Ring $K_N[X]$ ist genau dann ein Körper, wenn N ein irreduzibles Polynom ist.*

Beweis (vergleiche Beweis von 6.4): Sei $R := K_N[X]$. Ist N nicht irreduzibel, also zerlegbar, so besitzt R Nullteiler und R ist kein Körper. Sei N irreduzibel, und seien $A, B \neq 0$ aus $K_N[X] =: R$. Wir zeigen, dass genau ein Polynom P existiert mit

$$A \circ P = B.$$

Dazu betrachten wir die Abbildung

$$\alpha : R^* \to R^* \quad \text{mit} \quad P \mapsto A \circ P.$$

1) α ist definiert, denn R ist nullteilerfrei (7.4).

2) α ist injektiv: Seien $A, B \in R^*$ mit $\alpha(A) = \alpha(B)$, d.h. $A \circ P = B \circ P$. Es folgt

$$0 = A \circ P - B \circ P = (A - B) \circ P$$

und daraus $A - B = 0$, d.h. $A = B$, da R nullteilerfrei ist.

3) α ist linear:

$$\begin{aligned} \alpha(P+Q) &= A \circ (P+Q) = A \circ P + A \circ Q = \alpha(P) + \alpha(Q) \\ \alpha(\lambda P) &= A \circ (\lambda P) \stackrel{!}{=} \lambda(A \circ P) = \lambda\alpha(P) \end{aligned}$$

Bei ! haben wir 7.3 benutzt:

$$A \circ \lambda P = \rho(A(\lambda P)) = \rho(\lambda AP) = \lambda\rho(AP) = \lambda(A \circ B)$$

4) α ist injektiv: $\alpha(P) = \alpha(Q)$ bedeutet $A \circ P = A \circ Q$, d.h. $0 = A \circ P - A \circ Q = A \circ (P - Q)$. Es folgt $P - Q = 0$, d.h. $P = Q$, da R nullteilerfrei ist.

Wir haben in 4.22 und schon in 2.11 gelernt, dass eine injektive lineare Abbildung α eines Vektorraums (hier $\mathbb{F}_n[X]$) in sich auch surjektiv ist.

Insgesamt ist $\alpha : R^* \to R^*$ bijektiv. Es existiert also genau ein $P \in K_N[X]$ mit $A \circ P = 1$. Damit ist R^* Gruppe und der Ring R Körper. □

7.6 Satz: *Sei $p \in \mathbb{N}$ Primzahl und $K := \mathbb{Z}_p$. Sei $N \in K[X]$ ein irreduzibles Polynom vom Grad n. Dann ist $K_N[X]$ ein endlicher Körper, er hat p^n Elemente.*

Beweis: Angesichts 7.5 ist nur noch $|K_N[X]| = p^n$ zu zeigen. In unseren Beispielen haben wir gesehen, dass der Vektorraum $K_n[X]$ mit dem Vektorraum K^n identifiziert werden kann. Es ist $|K^n| = p^n$, da $|K| = p$. □

Ein vergleichsweise tief liegender Satz besagt:

• Sei $K = \mathbb{Z}_p$. Dann existiert zu jedem $n \in \mathbb{N}$ ein irreduzibles Polynom $N \in K[X]$ vom Grad n.

Zum Beispiel ist im Fall $K = \mathbb{Z}_2 = \{0, 1\}$ das Polynom

$$N = 1 + X + X^3 + X^4 + X^8$$

irreduzibel und definiert den endlichen Körper $\mathbb{F}_{2^8}$, welcher in der Informationstechnologie eine bedeutende Rolle spielt.

Wir kommen nun zu einer weiteren wichtigen Eigenschaft von Polynomen.

7.7 Satz: *Sei $0 \neq A \in K[X]$ und $\lambda \in K[X]$ eine Nullstelle von A, d.h. $A(\lambda) = 0$. Dann existiert $B \in K[X]$ mit $A = (X - \lambda)B$.*

Beweis: Wir teilen A durch $X - \lambda$ (Division mit Rest) und erhalten $A = F(X - \lambda) + R$, mit $\operatorname{grad} R < 1$. Es ist

$$0 = A(\lambda) = F(\lambda)(\lambda - \lambda) + R(\lambda),$$

also $R(\lambda) = 0$. Es folgt $R = 0$. □

7.8 Korollar: *Seien $\lambda_1, \ldots, \lambda_m$ verschiedene Nullstellen von A. Dann ist A von der Form*

$$A = (X - \lambda_1)(X - \lambda_2) \ldots (X - \lambda_m)C \quad , \quad C \in \mathbb{F}[X].$$

Beweis: Es ist $A = (X - \lambda_1)B_1,\ B_1 \in \mathbb{F}[X]$ und $0 = A(\lambda_2) = (\lambda_2 - \lambda_1)B_1(\lambda_2)$. Wegen $\lambda_2 \neq \lambda_1$ folgt $B_1(\lambda_2) = 0$. Also ist $B_1 = (X - \lambda_2)B_2$. Es ist $B_2(\lambda_3) = 0$, also $B_2 = (X - \lambda_3)B_3$. So fortfahrend erhält man die Behauptung, denn

$$\operatorname{grad} A > \operatorname{grad} B_1 > \operatorname{grad} B_2 > \ldots$$

□

Daraus folgt:

7.9 *Sei $n := \operatorname{grad} A$. Dann hat A höchstens n Nullstellen.* □

Bezeichnung: Hat A die Form $A = (a_1X + a_0)B,\ a_1 \neq 0$, so nennt man $a_1X + a_0$ einen *Linearfaktor* von A.

Bemerkung: $a_1X + a_0 = a_1(X - \lambda)$ mit $\lambda := -\frac{a_0}{a_1}$. Also ist λ Nullstelle von $a_1X + a_0$ und A. Bis auf eine Normierung haben die Linearfaktoren von A die Form $X - \lambda$.

Bezeichnung: Das Polynom A *zerfällt* in Linearfaktoren, wenn A ein Produkt von Linearfaktoren ist. In diesem Fall ist

$$A = a_n \prod_{i=1}^{m} (X - \lambda_i)^{e_i} \quad , \quad e_i \in \mathbb{N},$$

wobei $\lambda_1, \ldots, \lambda_m$ die verschiedenen Nullstellen von A sind. Die Zahl e_i heißt *Vielfachheit* der Nullstelle $\lambda_i,\ i = 1, \ldots, m$.

Ein verhältnismäßig tief liegender Satz ist der

7.10 Fundamentalsatz der Algebra: *Sei* $A \in \mathbb{C}[X],\ \operatorname{grad} A \leq 1$. *Dann existiert ein* $\lambda \in \mathbb{C}$ *mit* $A(\lambda) = 0$.

7.11 Korollar: *In* $\mathbb{C}[X]$ *zerfällt jedes Polynom* A *vom Grad* ≥ 1 *in Linearfaktoren.* □

Zum Beispiel gilt

$$X^2 + 1 = (X - i)(X + i)$$

in $\mathbb{C}[X]$. In $\mathbb{R}[X]$ ist das Polynom $X^2 + 1$ nicht zerlegbar, also irreduzibel.

Der Fundamentalsatz der Algebra besagt, dass die irreduziblen Polynome in $\mathbb{C}[X]$ genau die Polynome vom Grad 1 sind.

Daraus folgt, dass die irreduziblen Polynome in $\mathbb{R}[X]$ Grad ≤ 2 haben. Dies werden wir später mal begründen.

Ist $A \in K[X]$ und $\lambda \in K$ Nullstelle von A, so ist $A = (X - \lambda)B$, also A nicht irreduzibel.

7.12 *Sei* $\operatorname{grad} A = 2$ *oder* $\operatorname{grad} A = 3$. *Besitzt* A *keine Nullstelle in* K*, so ist* A *irreduzibel.*

Beweis: Wir nehmen an, dass A nicht irreduzibel ist. Dann existiert eine Faktorisierung $A = BC$ mit $\operatorname{grad} B = 2$ und $\operatorname{grad} C = 1$. Als lineares Polynom besitzt C eine Nullstelle λ. Dann ist $A(\lambda) = B(\lambda)C(\lambda) = 0$ im Widerspruch zur Annahme. □

Wir kommen nun zu einem weiteren Aspekt, nämlich der *Interpolation*.

Im Folgenden seien $\xi_1, \xi_2, \ldots, \xi_n$ verschiedene Zahlen aus K; wir nennen sie *Stützstellen*. Sei

$$y = (y_1, y_2, \ldots, y_n) \in K^n.$$

Die Interpolation löst die Aufgabe, ein Polynom $A \in \mathbb{F}_n[X]$ zu finden, sodass

$$A(\xi_i) = y_i \quad , \quad i = 1, \ldots, n.$$

Schreiben wir $A = \sum_{i=0}^{n-1} a_i X^i$, so bedeutet dies

$$
(*) \qquad \begin{array}{ccccccccccc}
a_0 & + & a_1\xi_1 & + & a_2\xi_1^2 & + & \dots & + & a_{n-1}\xi_1^{n-1} & = & y_1 \\
\vdots & & \vdots & & \vdots & & & & \vdots & & \vdots \\
a_0 & + & a_1\xi_n & + & a_2\xi_n^2 & + & \dots & + & a_{n-1}\xi_n^{n-1} & = & y_n
\end{array}
$$

Dies ist ein lineares Gleichungssystem in den n Unbekannten $a_0, \dots, a_{n-1}$ mit der Koeffizientenmatrix

$$
V := \begin{pmatrix}
1 & \xi_1 & x_1^2 & \dots & \xi_1^{n-1} \\
1 & \xi_2 & x_2^2 & \dots & \xi_2^n \\
\vdots & \vdots & & & \\
1 & \xi_n & x_n^2 & \dots & \xi_{n-1}^{n-1}
\end{pmatrix} \qquad (Vandermondsche\ Matrix)
$$

In Matrixschreibweise lautet $(*)$

$$
Va^T = y^T \quad , \quad a = (a_0, a_1, \dots, a_{n-1}).
$$

Wir zeigen, dass V invertierbar ist, also ist

$$
a^T = V^{-1}y^T
$$

und dies ist die Lösung unseres Interpolationsproblems.

In der Praxis geht man anders vor. Wir erklären die *Newtoninterpolation* zunächst an einem Beispiel:

Sei $K = \mathbb{Q}$ und $n = 3$ sowie $\xi_1 = 1,\ \xi_2 = 2,\ \xi_3 = 3$. Wir interpolieren das Wertetupel $y = (0, 1, 2)$ und machen den Ansatz

$$
A = \lambda_0 + \lambda_1(X - 1) + \lambda_2(X - 1)(X - 2).
$$

Es ist

$$
\begin{array}{rcl}
0 & = & A(1) = \lambda_0 \\
1 & = & A(2) = \lambda_0 + \lambda_1(2 - 1) \\
2 & = & A(3) = \lambda_0 + \lambda_1(3 - 1) + \lambda_2(3 - 1)(3 - 2)
\end{array}
$$

Sukzessive folgt: $\lambda_0 = 0$,

$$
\begin{array}{rcl}
1 & = & 0 + \lambda_1 \ \Rightarrow\ \lambda_1 = -1 \\
2 & = & 0 + (-1) \cdot 2 + \lambda_2 \cdot 2 \ \Rightarrow\ \lambda_2 = 2
\end{array}
$$

Also ist

$$A = -(X-1) + 2(X-1)(X-2) = 2X^2 - 5X + 3.$$

Im allgemeinen Fall definiert man Polynome N_i vom Grad i, $i = 1, \ldots, n-1$ durch

$$N_i = (X - \xi_1)(X - \xi_2) \ldots (X - \xi_{i-1}).$$

Man macht folgenden Ansatz

$$A = \alpha_0 + \alpha_1 N_1 + \alpha_2 N_2 + \ldots + a_{n-1} N_n$$

und berechnet sukzessive die Koeffizienten $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$ aus folgendem Schema:

$$\begin{array}{rcll} y_1 & = & A(\xi_1) & = \alpha_0 \\ y_2 & = & A(\xi_2) & = \alpha_0 + \alpha_1 N_1(\xi_2) \\ y_3 & = & A(\xi_3) & = \alpha_0 + \alpha_1 N_1(\xi_3) + \alpha_2 N_2(\xi_3) \\ \vdots & & \vdots & \\ y_n & = & A(\xi_n) & = \alpha_0 + \alpha_1 N_1(\xi_n) + \alpha_2 N_2(\xi_n) + \alpha_{n-1} N_{n-1}(\xi_n) \end{array}$$

Dabei beachte man $N_j(\xi_{j+1}) \neq 0$, denn $\xi_i \neq \xi_j$ für $i \neq j$. Insbesondere folgt:

7.13 Interpolationssatz: *Zu $y = (y_1, \ldots, y_n) \in K^n$ existiert genau ein Polynom $A \in K_n[X]$ mit $A(\xi_i) = y_i,\ i = 1, \ldots, n.$* □

Bemerkung 1: Das Gleichungssystem $(*)$ $Va^T = y^T$ hat also genau eine Lösung. Die Vandermondsche Matrix V ist daher invertierbar (2.11).

Bemerkung 2: Sei $A \in K_n[X]$ und $y_i = A(\xi_i),\ y = (y_1, \ldots, y_n)$. Dann existiert genau ein Tupel $(\alpha_0, \ldots, \alpha_{n-1}) \in K^n$ mit

$$A = \alpha_0 + \alpha_1 N_1 + \ldots + \alpha_{n-1} N_{n-1}.$$

Dies bedeutet, dass die Polynome $1, N_1 + \ldots, N_{n-1}$ eine Basis des Vektorraums $K_n[X]$ ist.

Sei K ein endlicher Körper. Ein verhältnismäßig tief liegender Satz besagt

- *Die multiplikative Gruppe K^* ist eine zyklische Gruppe.*

Wir machen folgende Voraussetzung:

Z *Sei $\beta \in K^*$ und das Erzeugnis $\langle\beta\rangle$ eine zyklische Gruppe mit n Elementen ($\cong \mathbb{Z}_n$); insbesondere ist $\langle\beta\rangle = \{1, \beta, \beta^2, \ldots, \beta^{n-1}\}$.*

Für unsere Interpolation nehmen wir als Stützstellen die Zahlen $1, \beta, \ldots, \beta^{n-1}$. Sei $y = (y_0, \ldots, y_{n-1}) \in K^n$. Wir bestimmen $A \in \mathbb{F}_n[X]$ mit $A(\beta^i) = y_i$, d.h. für $i = 0, \ldots, n-1$ gilt

(i) $$y_i = a_0 + a_1\beta^i + a_2\beta^{2i} + \ldots + a_{n-1}\beta^{(n-1)i}$$

Die Zahl $\tilde{n} \in K$ sei definiert durch

$$\tilde{n} = \underbrace{1 + 1 + \ldots + 1}_{n} \quad , \quad 1 \in K$$

Im Fall $K = \mathbb{Z}_2$ und $n = 2$ ist z.B. $\tilde{n} = 0$; im Fall $K = \mathbb{Z}_5$ und $n = 4$ ist $\tilde{n} \neq 0$.

Wir bilden das Polynom

$$B := \frac{1}{\tilde{n}}(y_0 + y_1X + \ldots + y_{n-1}X^{n-1}).$$

Unter diesen Voraussetzungen gilt

7.14 $a_j = B(\beta^{-j}),\ j = 0, \ldots, n-1$

Beweis: Sei $\lambda \in \langle\beta\rangle$ und

$$s := 1 + \lambda + \lambda^2 + \ldots + \lambda^{n-1}.$$

Dann ist

$$\lambda s = \lambda + \lambda^2 + \ldots + \lambda^{n-1} + \lambda^n \stackrel{!}{=} s,$$

da $\lambda^n = 1$ in der zyklischen Gruppe $\langle\beta\rangle$. Also ist

(ii) $$s = \begin{cases} 0 & \lambda \neq 1 \\ \text{falls} & \\ \tilde{n} & \lambda = 1 \end{cases}$$

Eine Umsummerierung ergibt

$$\begin{aligned}\tilde{n}B(\beta^{-j}) &= \sum_{i=0}^{n-1} y_i\beta^{-ji} \overset{(i)}{=} \sum_{i=0}^{n-1}\left(\sum_{k=0}^{n-1} a_k\beta^{ik}\right)\beta^{-ji} \\ &\overset{!}{=} \sum_{k=0}^{n-1} a_k\left(\sum_{i=0}^{n-1}\beta^{ik}\beta^{-ji}\right) = \sum_{k=0}^{n-1} a_k\left(\sum_{i=0}^{n-1}\beta^{i(k-j)}\right) \\ &= \sum_{k=0}^{n-1} a_k\left(\sum_{i=0}^{n-1}\lambda^{-i}\right),\end{aligned}$$

dabei ist $\lambda := \beta^{(k-i)}$ gesetzt. Aus (ii) folgt $\tilde{n}B(\beta^{-j}) = \tilde{n}a_j$, und daraus die Behauptung. □

Diese Beziehung ist – insbesondere in der Elektrotechnik – unter dem Stichwort *diskrete Fourieranalyse* bekannt. Sie zeigt, dass – unter der speziellen Voraussetzung **Z** – die i. Allg. sehr rechenaufwändige *Interpolation* eine (viel *schnellere*) *Auswertung* ist. In der Praxis der digitalen Nachrichtenübertragung ist dies von entscheidender Bedeutung.

Wir kommen nun zu dem Thema CODES, welches wir am Beispiel der *Reed–Solomon* Codes erläutern.[22]

Seien $t, k, n \in \mathbb{N}$ mit

$$n = k + 2t.$$

Wir wollen k–Tupel $a = (a_0, \ldots, a_{n-1}) \in K^k$ (*Informationen*) in einem fehlerhaften Kanal so übertragen, dass danach die ursprüngliche Information a wieder rekonstruiert werden kann.

Dazu wird a zunächst *codiert*: Seien n *Stützstellen* $\xi_1, \ldots, \xi_n \in \mathbb{F}$ fest vorgegeben. Man bildet das Polynom

$$A = a_0 + a_1X + \ldots + a_{n-1}X^{n-1} \in K_n[X]$$

und wertet es aus:

$$y_i := A(\xi_i) \qquad i = 1, \ldots, n$$

[22] Reed–Solomon Codes arbeiten in jedem CD–Spieler, Computer, Handy, ...

Man erhält ein *Wertetupel* $y = (y_1, \ldots, y_n)$. Der Reed–Solomon Code RS besteht aus allen solchen Tupeln

$$\mathrm{RS} = \{(y_1, \ldots, y_n) \in K^n \mid y_i = A(\xi_i),\ A \in K[X]\}$$

Der entscheidende Punkt ist:

- *In je k Komponenten y_i von y ist die Information a codiert.*

Um dies zu begründen, seien $y_{i_1}, \ldots, y_{i_k}$ verschiedene Komponenten von y. Nach dem Interpolationssatz 7.13 existiert genau ein Polynom $B \in K_k[X]$ mit

$$B(\xi_{i_j}) = y_{i_j} \quad , \quad j = 1, \ldots, k.$$

Nach Konstruktion ist auch $A(\xi_{i_j}) = y_{i_j}$ und $A \in K_k[X]$. Es folgt $A = B$.

Anstatt $a \in K^k$ überträgt man nun die *codierte Information* $y \in K^n$ in einem fehlerhaften Kanal. Der Empfänger erhält ein Tupel $\tilde{y} = (\tilde{y}_1, \ldots, \tilde{y}_n) \in K^n$, das i. Allg. verschieden von y ist. Sei

$$\mathbf{F} := \{i \in \{1, \ldots, n\} \mid y_i \neq \tilde{y}_i\}$$

die *Fehlerstellenmenge*. Wir machen nun eine Voraussetzung an den Übertragungskanal, nämlich

$$|\mathbf{F}| \leq t$$

und begründen, dass unter dieser Voraussetzung die codierte Information y, und damit a, aus $\tilde{y}$ zurückgewonnen werden kann (*Decodierung*).

Für $x = (x_1, \ldots, x_n),\ y = (y_1, \ldots, y_n) \in K^n$ definieren wir den *Hammingabstand* $d(x, y)$, und zwar ist $d(x, y)$ die Anzahl der $i \in \{1, \ldots, n\}$ mit $x_i \neq y_i$. Die Voraussetzung $|\mathbf{F}| \leq t$ bedeutet also $d(y, \tilde{y}) \leq t$.

7.15 Lemma: *Für alle $x, y \in \mathrm{RS},\ x \neq y$, ist $d(x, y) \geq 2t + 1$.*[23]

Beweis: Sei y die Auswertung von $A \in K_k[X]$ und x die Auswertung von $B \in K_k[X]$. Sei $d(x, y) \leq 2t$ angenommen. Wegen $n = k + 2t$ existieren dann mindestens k Stellen $i \in \{1, \ldots, n\}$ mit $x_i = y_i$. In k Stellen ist aber die Information codiert. Es folgt $A = B$, also $x = y$ entgegen der Annahme $x \neq y$. □

Das **Decodierungsprinzip** ist nun folgendes:

[23] Man kann zeigen, dass diese Ungleichung *scharf* ist.

7.16 *Das Element $y \in \mathrm{RS}$ ist das einzige Element $x \in \mathrm{RS}$ mit $d(\tilde{y}, x) \leq t$.*

Beweis: Für $z \in \mathrm{RS}$ gelte ebenfalls $d(\tilde{y}, z) \leq t$. Die Vektoren y und z unterscheiden sich also von $\tilde{y}$ an jeweils höchstens t Stellen $i \in \{1, \ldots, n\}$. Dann unterscheiden sich aber y und z an höchstens $2t$ Stellen. Also ist $d(y, z) \leq 2t$. Aus 7.15 folgt $z = y$. □

In der Praxis ist die in 7.16 geschilderte *Sucherei* viel zu zeitaufwendig. Es existieren daher sehr raffinierte Verfahren, um $\tilde{y}$ zu decodieren. Dabei benutzt man wesentlich die schnelle Interpolation 7.14; letzten Endes läuft die Decodierung auf das Lösen von linearen Gleichungssystemen hinaus (Gausselimination).

Typische Parameter eines RS–Codes sind in der Praxis

$$n = 255 = 2^8 - 1 \quad , \quad t = 12 \quad , \quad k = 231.$$

Dabei nimmt man den Körper $K = GF(2^8)$. Die Stützstellen ξ_i sind die Elemente $\neq 0$ von k; sie bilden also eine zyklische Gruppe!

8 Die komplexen Zahlen

Eine komplexe Zahl hat die Form

$$z = a + ib \quad , \quad a = \mathrm{Re}\, z \quad , \quad b = \mathrm{Im}\, z \in \mathbb{R}.$$

Sie bilden den komplexen Zahlkörper $\mathbb{C}$, siehe 7.1.

Seien $z_1 = a_1 + ib_1,\ z_2 = a_2 + ib_2 \in \mathbb{C}$. Dann ist

$$z_1 + z_2 = (a_1 + a_2) + i(b_1 + b_2)$$

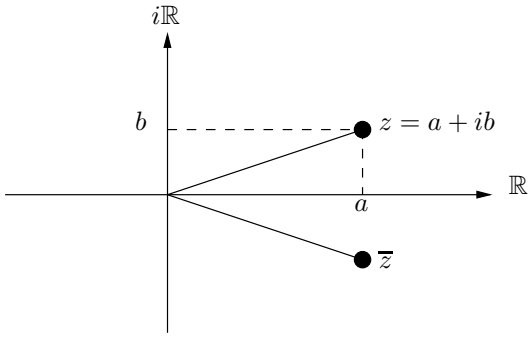
und $(i^2 = -1)$

$$\begin{aligned} z_1 z_2 &= (a_1 + ib_1)(a_2 + ib_2) = a_1 a_2 + i a_1 b_2 + i b_1 a_2 + i^2 b_1 b_2 \\ &= (a_1 a_2 - b_1 b_2) + i(a_1 b_2 + b_1 a_2). \end{aligned}$$

Der Körper $\mathbb{R}$ liegt in $\mathbb{C}$, denn

$$\mathbb{R} = \{z \in \mathbb{C} \mid \mathrm{Im}\, z = 0\}.$$

Der Körper $\mathbb{C}$ ist insbesondere ein Vektorraum über dem Körper $K = \mathbb{R}$; die Zahlen $1, i$ sind eine Basis dieses Vektorraums. Identifiziert man $\mathbb{C}$ bez. dieser



Basis mit dem $\mathbb{R}^2$, so spricht man von der *komplexen Zahlebene*.

Die reelle Zahl

$$|z| := \sqrt[+]{a^2 + b^2} \geq 0$$

heißt *Betrag* von z. Sie misst den Abstand des Punktes z zum Nullpunkt. Im Fall $z \in \mathbb{R}$ ist $z = a$, also $|z|$ der Betrag der reellen Zahl r.

Die Zahl

$$\overline{z} := a - ib$$

heißt die zu z *konjugiert komplexe* Zahl.

Wir stellen nun eine Reihe von *Rechenregeln* zusammen.

1) $z = 0 \Leftrightarrow |z| = 0$

2) $z \in \mathbb{R} \Leftrightarrow \operatorname{Im} z = 0 \Leftrightarrow z = \overline{z}$

3) $z\overline{z} = |z|^2$, denn: $(a + ib)(a - ib) = a^2 + b^2 + i(ab - ba)$.

4) $\overline{z_1 + z_2} = \overline{z}_1 + \overline{z}_2,\ \overline{z_1 z_2} = \overline{z}_1 \overline{z}_2$,
denn: $\overline{z_1 z_2} = (a_1 a_2 - b_1 b_2) - i(a_1 b_2 + b_1 a_2) = (a_1 - ib_1)(a_2 - ib_2)$.

5) $|z_1 z_2| = |z_1||z_2|$,
denn: $|z_1 z_2|^2 \overset{3)}{=} (z_1 z_2)(\overline{z_1 z_2}) \overset{4)}{=} z_1 z_2 \overline{z}_1 \overline{z}_2 = z_1 \overline{z}_1 z_2 \overline{z}_2 = |z_1|^2 |z_2|^2$.

Bemerkung: Der Betrag ist also *multiplikativ*, nicht aber *additiv*, sondern es gilt

schwächer die *Dreiecksungleichung*

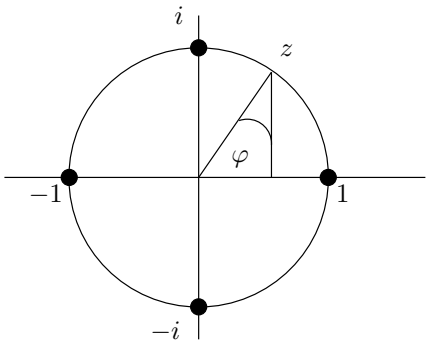
$$|z_1 + z_2| \leq |z_1| + |z_2|.$$

Wir kommen darauf zurück.

6) $z^{-1} = \frac{\overline{z}}{|z|^2}$ für $z \neq 0$, denn: $z\frac{\overline{z}}{|z|^2} = \frac{z\overline{z}}{|z|^2} \overset{3)}{=} 1$.

7) $\left|\frac{z}{|z|}\right| = 1$ für $z \neq 0$, denn: $\left|z\frac{1}{|z|}\right| \overset{5)}{=} |z| \left|\frac{1}{|z|}\right| = |z| \cdot \frac{1}{|z|}$.

8) *Ist* $|z| = 1$, *d.h. liegt* z *auf dem Einheitskreis, so existiert genau ein Winkel* φ, $0 \leq \varphi < 2\pi$ *mit* $z = \cos\varphi + i\sin\varphi$.



Ist $z \in \mathbb{C}$ beliebig, $z \neq 0$, so ist $\left|\frac{z}{|z|}\right| = 1$. Also folgt mit 8)

9) $z = |z|(\cos\varphi + i\sin\varphi),\ 0 \leq \varphi < 2\pi$

Diese Darstellung der komplexen Zahl nennt man *Polardarstellung* von z. Dabei ist $|z|$ der *Radius* und φ das Argument

$$\varphi = \arg z.$$

Zum Beispiel gilt $\arg\overline{z} = 2\pi - \varphi$.

10) $\arg(z_1 z_2) = \arg z_1 + \arg z_2$ (modulo 2π)
Denn:

$$\begin{aligned} z_1 z_2 &= |z_1||z_2|(\cos\varphi_1 + i\sin\varphi_1)(\cos\varphi_2 + i\sin\varphi_2) \\ &= |z_1||z_2|(\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2 + i(\cos\varphi_1\sin\varphi_2 + \sin\varphi_1\cos\varphi_2) \end{aligned}$$

Nun beachte die trigonometrischen Formeln[24]

$$\begin{aligned}\cos(\varphi_1+\varphi_2) &= \cos\varphi_1\cos\varphi_2-\sin\varphi_1\sin\varphi_2\\ \sin(\varphi_1+\varphi_2) &= \sin\varphi_1\cos\varphi_2+\cos\varphi_1\sin\varphi_2\end{aligned}$$

Aus 10) folgt

11) *Sei* $z=|z|(\cos\varphi+i\sin\varphi)$ *und* $w:=\sqrt[+]{|z|}\left(\cos\frac{\varphi}{2}+i\sin\frac{\varphi}{2}\right)$. *Dann ist* $w^2=z$. *Jede komplexe Zahl besitzt also eine Wurzel.*

Daraus folgt, dass jede quadratische Gleichung

$$a_0+a_1X+a_2X^2=0 \quad , \quad a_i\in\mathbb{C} \quad , \quad a_2\neq 0$$

eine Lösung $z\in\mathbb{C}$ hat (wie in der Schule). Nach dem Fundamentalsatz der Algebra 7.10 ist sogar jede polynomiale Gleichung

$$a_0+a_1X+a_2X^2+\ldots+a_nX^n=0 \quad , \quad a_i\in\mathbb{C} \quad , \quad a_n\neq 0$$

in $\mathbb{C}$ lösbar (sofern $n\geq 1$).

Zum Schluss lösen wir ein Problem, welches wir schon in **7** gestellt haben.

8.1 Satz: *Sei* $A\in\mathbb{R}[X]$ *ein irreduzibles Polynom. Dann ist* $\operatorname{grad}A\leq 2$.

Beweis: Als irreduzibles Polynom besitzt A nach 7.7 keine Nullstelle in $\mathbb{R}$. Wegen $\mathbb{R}\subseteq\mathbb{C}$ liegt A auch in $\mathbb{C}[X]$. Nach 7.10 existiert also ein $z\in\mathbb{C}$ mit $A(z)=0$. Dabei ist $z\notin\mathbb{R}$, also $z\neq\overline{z}$. Es ist

$$0=A(z)=a_0+a_1z+a_1z^2+\ldots+a_nz^n \quad , \quad a_i\in\mathbb{R}.$$

Unter wiederholter Anwendung der Regeln 4) und 5) folgt

$$\begin{aligned}0 &= \overline{0}=\overline{A(z)}=\overline{a_0+a_1z+a_2z^2+\ldots+a_nz^n}\\ &= \overline{a}_0+\overline{a_1z}+\overline{a_2z^2}+\ldots+\overline{a_nz^n}\\ &= \overline{a}_0+\overline{a}_1\overline{z}+\overline{a}_2\overline{z^2}+\ldots+\overline{a}_n\overline{z^n}\\ &= \overline{a}_0+\overline{a}_1\overline{z}+\overline{a}_2\overline{z}^2+\ldots+\overline{a}_n\overline{z}^n\\ &\overset{2)}{=} a_0+a_1\overline{z}+a_2\overline{z}^2+\ldots+a_n\overline{z}^n\\ &= A(\overline{z})\end{aligned}$$

[24], die wir irgendwann beweisen werden.

Also ist auch $\overline{z}$ Nullstelle von A. Wegen $z \neq \overline{z}$ haben wir nach 7.8 folgende Faktorisierung von A in $\mathbb{C}[X]$

$$A = FB \quad , \quad F := (X - z)(X - \overline{z}).$$

Nun ist

$$F = X^2 - (z + \overline{z})X + z\overline{z} = X^2 - (2\mathrm{Re}\, z)X + |z|^2$$

ein reelles Polynom. Teilt man das reelle Polynom in $\mathbb{C}[X]$ durch das reelle Polynom F, so erhält man ein reelles Polynom, und dieses ist B. Da A irreduzibel in $\mathbb{R}[X]$ ist, folgt $\mathrm{grad}\, B = 0$ und $\mathrm{grad}\, A = \mathrm{grad}\, F = 2$. □

9 Koordinatentransformation und Eigenwerte

Im Folgenden sei V ein Vektorraum über dem Körper K, $\dim V = n$. Es seien $g_1, \ldots, g_n$ und $h_1, \ldots, h_n$ zwei Basen von V, die erste nennen wir g–Basis und die zweite h–Basis. Ist $v \in V$ und

$$v = \sum_{i=1}^{n} x_i g_i \quad , \quad x_i \in K,$$

so ist $x = (x_1, \ldots, x_n) \in K^n$ das Koordinatentupel von v bezüglich der g–Basis; wir schreiben $x = \kappa_g(v)$, siehe 4.11. Analog schreiben wir $y = \kappa_h(v)$, wenn

$$v = \sum_{i=1}^{n} y_i h_i \quad , \quad y_i \in K.$$

Wir erklären das Folgende an einem Beispiel. Sei $V = K^2,\ K := \mathbb{R}$ und

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix} \in K^{2\times 2}$$

sowie $\alpha : V \to V$ die lineare Abbildung definiert durch A, also

$$\alpha(x) = \alpha(x_1, x_2) = Ax = (x_1 + x_2, 2x_1).$$

Die Vektoren $g_1 := e_1,\ g_2 := e_2$ bilden eine g–Basis von V; für $v = x \in A$ ist $\kappa_g(v) = x$. Es ist $a_{*1} = (1, 2) = \alpha(g_1)$ und $a_{*2} = (1, 0) = \alpha(g_2)$.

Allgemein hat man, siehe 2.3

9.1 *Es gelte:*
$\alpha : V \to V$ *sei linear und sei* $v \in V$
$x := \kappa_g(v)$ *und* $y := \kappa_g(\alpha(v))$
$A \in K^{n\times n}$ *definiert durch* $a_{*j} = \kappa_g(g_j),\ j = 1, \ldots, n.$
Dann ist $y = Ax.$ □

Wir führen unser Beispiel weiter und definieren eine Teilmenge Ei von $V = K^2$ als die Menge aller Punkte $x = (x_1, x_2)$, für die gilt

(ei) $5x_1^2 + x_2^2 + 2x_1x_2 = 1$

Wir bestimmen das Bild von Ei unter α, d.h.

$$\alpha(\mathrm{Ei}) = \{\alpha(x) \mid x \in \mathrm{Ei}\}.$$

Mit $y = \alpha(x)$ gilt

$$(y_1, y_2) = y = Ax = (x_1 + x_2, 2x_1).$$

Die Matrix A ist invertierbar; es ist $A^{-1} = \begin{pmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{pmatrix}$. Also ist

$$x = A^{-1}y = \left(\frac{1}{2}y_2, y_1 - \frac{1}{2}y_2\right).$$

Dies setzen wir in (ei) ein und erhalten

$$\begin{aligned} 1 &= 5\left(\frac{1}{4}y_2^2\right) + \left(y_1 - \frac{1}{2}y_2\right)^2 + 2\left(\frac{1}{2}y_2\right)\left(y_1 - \frac{1}{2}y_2\right) \\ &= \frac{5}{4}y_2^2 + y_1^2 + \frac{1}{4}y_2^2 - y_1y_2 + y_2y_1 - \frac{1}{2}y_2^2 \\ &= y_1^2 + y_2^2 \end{aligned}$$

Also ist $\alpha(\mathrm{Ei})$ ein *Kreis* mit *Radius* 1.

Wir interpretieren nun diesen Sachverhalt anders. Die Vektoren

$$h_1 := (0, 1) = \alpha(g_1) \quad , \quad h_2 := \left(\frac{1}{2}, -\frac{1}{2}\right) = \alpha(g_2)$$

bilden eine h–Basis von $V = K^2$. Für $v \in V$ ist

$$\begin{aligned} v &= x_1g_1 + x_2g_2 \quad , \quad x = \kappa_g(v) \\ &= y_1h_1 + y_2h_2 \quad , \quad y = \kappa_h(v) \end{aligned}$$

Dann folgt

$$\alpha^{-1}(v) = \alpha^{-1}(y_1h_1 + y_2h_2) = y_1\alpha^{-1}(h_1) + y_2\alpha^{-1}(h_2) = y_1g_1 + y_2g_2,$$

also $\kappa_g(\alpha^{-1}(v)) = y$. Es ist $\alpha(\alpha^{-1}(v)) = v$. Somit besagt 9.1

$$Ay = x \quad , \quad y = A^{-1}x.$$

Wir setzen nun $T := A^{-1}$ und nennen T *Transformationsmatrix*, welche den Basiswechsel von der g–Basis zur h–Basis beschreibt. Man beachte (wir formulieren allgemein):

(**T**) Die j–te Spalte t_{*j} von $T = (t_{ij})$ ist das Koordinatentupel des Vektors h_j bez. der g–Basis, also $t_{*j} = \kappa_g(h_j), j = 1, \ldots, n$.

Ist wie anfangs $x = \kappa_g(v)$ und $y = \kappa_h(v)$, so hat man das

9.2 Transformationsgesetz: $x = Ty$ *und* $y = T^{-1}x$

Beweis: $\sum_{j=1}^n y_jh_j = \sum_{j=1}^n y_j \left(\sum_{i=1}^n t_{ij}g_i\right) = \sum_{i=1}^n \left(\sum_{j=1}^n t_{ij}y_j\right) g_i = \sum_{i=1}^n x_ig_i$ □

Sei z das Koordinatentupel von $\alpha(v)$ bez. der h–Basis. Wir haben

	g–Basis	h–Basis
Koordinatentupel von v	x	y
Koordinatentupel von $\alpha(v)$	y	z

Bezüglich der g–Basis ist $y = Ax$ nach 9.1; bezüglich der h–Basis existiert ebenfalls eine Matrix B, sodass

$$z = By.$$

Nach 9.2 ist $y = T^{-1}x,\ z = T^{-1}y$. Es folgt

$$T^{-1}y = BT^{-1}x,$$

also $y = TBT^{-1}x$, d.h. $A = TBT^{-1}$. Damit haben wir ein weiteres Transformationsgesetz

9.3 $B = T^{-1}AT$ □

Bezeichnung: Eine Matrix B der Form $B = T^{-1}AT,\ T \in GL(n, K)$ heißt eine zu A *ähnliche* Matrix.

Wir formulieren nocheinmal:

9.4 *Ähnliche Matrizen beschreiben dieselbe lineare Abbildung bez. verschiedener Basen. Sie haben dieselbe Determinante.* □

Letzteres folgt mit 3.9 b):

$$\begin{aligned}\det(T^{-1}AT) &= (\det T^{-1})(\det A)(\det T)\\ &= (\det A)(\det T^{-1})(\det T) = (\det A)(\det(TT^{-1})) = \det A\end{aligned}$$

In unserem Beisiel behalten wir die g–Basis $(g_1 = e_1, g_2 = e_2)$ und nehmen nur eine andere h–Basis, nämlich $h_1 := (1, 1)$ und $h_2 := (1, -2)$. Die Transformationsmatrix ist

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

Eine leichte Rechnung (nach 3.14) ergibt:

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

Also ist

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

Bezüglich der h–Basis gehört in unserem Beispiel zu α eine *Diagonalmatrix*.

Beobachtung: 1) Sei $\lambda \in K$. Wir bilden die Matrix

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

und berechnen ihre Determinante

$$\det(A - \lambda 1) = (1-\lambda)(-\lambda) - 2 = \lambda^2 - \lambda - 2.$$

Wir setzen $\lambda^2 - \lambda - 2 = 0$ und erhalten die Nullstellen $\lambda_1 = 2,\ \lambda_2 = -1$. Es ist

$$\begin{aligned} \alpha(h_1) &= Ah_1 = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 2h_1 \\ \alpha(h_2) &= Ah_2 = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix} = (-1)h_2 \end{aligned}$$

Im Folgenden sei $A \in K^{n\times n}$.

Definition: Eine Zahl $\lambda \in K$ heißt *Eigenwert* von A, wenn $\det(A - \lambda \mathbf{1}_n) = 0$.

9.5 *Sei $\lambda \in K$ und $A \in K^{n\times n}$. Folgende Aussagen sind äquivalent:*

(i) λ *ist Eigenwert von* A.

(ii) $\mathrm{Rang}\,(A - \lambda \mathbf{1}_n) < n$

(iii) *Das homogene lineare Gleichungssystem* $(A - \lambda \mathbf{1}_n)x = 0$ *besitzt eine Lösung* $x \neq 0$.

(iv) *Es existiert ein Vektor* $x \in K^n,\ x \neq 0$, *mit* $Ax = \lambda x$.

Beweis: (i) $\Leftrightarrow$ (ii): Siehe 3.9 a).
(ii) $\Leftrightarrow$ (iii): Siehe 4.21.
(iii) $\Leftrightarrow$ (iv): $Ax = \lambda x \Leftrightarrow Ax - \lambda x = 0 \Leftrightarrow (A - \lambda \mathbf{1}_n)x = 0$ □

Definition: Ist $x \in K^n$ wie in (iii) oder (iv), so heißt x *Eigenvektor* von A zum Eigenwert λ.

Im vorigen Beispiel ist also h_1 Eigenvektor zum Eigenwert $\lambda = 2$ und h_2 Eigenvektor zum Eigenwert $\lambda = -1$.

9.6 *Sei $A \in K^{n\times n}$ in Dreiecksform*

$$A = \begin{pmatrix} \lambda_1 & & & * \\ & \ddots & & \\ & & \lambda_i & \\ & & & \ddots \\ 0 & & & \lambda_n \end{pmatrix}.$$

Dann sind (die nicht notwendig verschiedenen) Zahlen $\lambda_1, \ldots, \lambda_n$ sämtlich Eigenwerte von A.

Beweis: Sei $\lambda \in K$. Dann ist

$$A - \lambda \mathbf{1}_n = \begin{pmatrix} \lambda_1 - \lambda & & & & * \\ & \ddots & & & \\ & & \lambda_i - \lambda & & \\ & & & \ddots & \\ & & & & \lambda_n - \lambda \end{pmatrix}.$$

und nach (unserer) Definition von $\det A$

$$\det(A - \lambda \mathbf{1}) = \prod_{i=1}^{n} (\lambda_i - \lambda).$$

Also gilt

$$\lambda \quad \text{ist Eigenwert} \quad \Leftrightarrow \prod_{i=1}^{n} (\lambda_i - \lambda) = 0 \Leftrightarrow \lambda = \lambda_i \quad \text{für ein} \quad i.$$

□

Beispiel: Sei $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in K^{2\times 2}$. Es ist $\det(A - \lambda \mathbf{1}) = \lambda^2 + 1$. Sei $\alpha : K^2 \to K^2$ definiert durch A, also $\alpha(x) = Ax$.

1) $K = \mathbb{Z}_2$ (also $-1 = 1$). Hier ist $\lambda = 1$ der einzige Eigenwert. Der Vektor $h_1 := (1, 1)$ ist der dazugehörende Eigenvektor, denn $Ah_1 = (1, 1) = \lambda h_1$. Sei $h_2 := (0, 1)$. Dann ist h_1, h_2 Basis von K^2 und

$$Ah_2 = (1, 0) = h_1 + h_2.$$

Bezüglich dieser Basis gehört zu α die Matrix $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$; siehe 9.1.

Nach dem Transformationsgesetz 9.3 ist

$$B = TAT^{-1} \quad \text{mit} \quad T = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Die Matrix B ist keine Diagonalmatrix, aber immerhin Dreiecksmatrix.

2) Sei $K = \mathbb{R}$. Dann hat die Gleichung $\lambda^2 + 1$ keine Lösung in K. Also besitzt A

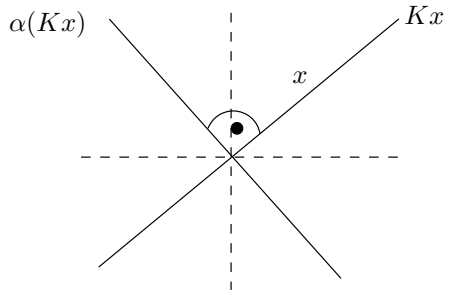
keinen Eigenwert, und darum auch keinen Eigenvektor. *Geometrisch* ist dies klar: Die lineare Abbildung α wirft den Punkt $(1,0)$ auf $(0,1)$ und $(0,1)$ auf $(-1,0)$. Sie ist also eine *Drehung* um den Winkel $\varphi = \frac{\pi}{2}$; siehe Seite 27.

$$A = \begin{pmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix}.$$

Ist nun $0 \neq x \in K^2$ Eigenvektor von A zum Eigenwert λ, so bildet α die Gerade Kx in sich ab, denn

$$\alpha(Kx) = K\alpha(x) = K(\lambda x) \subseteq Kx. \quad ^{25}$$

Aber α ist eine nichttriviale Drehung!



Dasselbe Argument zeigt, dass eine (*Dreh–*) Matrix

$$A = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \quad , \quad 0 \leq \varphi < 2\pi$$

nur dann Eigenwerte besitzt, wenn $\varphi = 0$ oder $\varphi = \pi$ ist. Im Fall $\varphi = 0$ ist $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ und $\lambda = 1$ Eigenwert; im Fall $\varphi = \pi$ ist $A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ und $\lambda = -1$ Eigenwert.

3) Sei $K = \mathbb{C}$. Dann sind $\lambda_1 = i,\ \lambda_2 = -i$ die Eigenwerte von A. Die dazugehörenden Eigenvektoren sind $h_1 = (1,-i)$ und $h_2 = (1,i)$, denn

$$Ah_1 = (i,1) = ih_1 \quad , \quad Ah_2 = (-i,1) = (-i)h_2.$$

[25] Im Fall $\lambda = 0$ ist $K(\lambda x) = 0$ und im Fall $\lambda \neq 0$ ist $K(\lambda x) = Kx$.

Bezüglich der Basis h_1, h_2 gehört zu α die Matrix

$$B = \begin{pmatrix} i & o \\ 0 & -i \end{pmatrix} \qquad \text{(Diagonalmatrix!)} \quad .$$

Also ist $TAT^{-1} = B$, wobei nach 9.2

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

4) $K = \mathbb{Z}_3 = \{0, 1, 2\}$. Die Gleichung $\lambda^2 + 1 = 0$ hat keine Lösung in K; also existieren keine Eigenwerte.

5) $K = \mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$; hier ist $4 = -1 = 2^2$ und $3^2 = 4 = -1$. Also sind die Zahlen 3 und 2 Lösungen von $\lambda^2 + 1$, sie sind also Eigenwerte. Es ist $3 = -2$. Setzen wir also $i = 2$, so ist $-i = 3$. Die Situation ist wie im vorher behandelten komplexen Fall $K = \mathbb{C}$, siehe 3).

Anstatt λ nehmen wir die Variable X und setzen

$$CH_A(X) := \det(A - X\mathbf{1}_n).$$

Aus der Leibnizschen Summenformel 3.15 folgt, dass CH_A ein Polynom $\in K[X]$ vom Grad n ist. Es heißt *charakteristisches Polynom* der Matrix A. Per Definition sind die Nullstellen von CH_A die Eigenwerte von A. Zum Beispiel ist im Fall $n = 2$

$$\begin{aligned} CH_A &= \det \begin{pmatrix} a_{11} - X & a_{12} \\ a_{21} & a_{22} - X \end{pmatrix} = (a_{11} - X)(a_{22} - X) - a_{21}a_{12} \\ &= X^2 - (a_{11} + a_{22})X + (a_{11}a_{22} - a_{21}a_{12}), \end{aligned}$$

und im Fall $n = 3$

$$\begin{aligned} CH_A &= \det \begin{pmatrix} a_{11} - X & a_{12} & a_{13} \\ a_{21} & a_{22} - X & a_{23} \\ a_{31} & a_{32} & a_{33} - X \end{pmatrix} \\ &= (a_{11} - X)(a_{22} - X)(a_{33} - X) + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{31} - a_{31}(a_{22} - X)a_{13} \\ &\quad -a_{32}a_{23}(a_{11} - X) - (a_{33} - X)a_{21}a_{12} \\ &= -X^3 + (a_{11} + a_{22} + a_{33})X^2 \\ &\quad -(a_{22}a_{33} + a_{11}a_{33} + a_{11}a_{22} - a_{31}a_{13} - a_{32}a_{23} - a_{21}a_{12})X \\ &\quad +a_{11}a_{22}a_{31} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{31} - (a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}) \end{aligned}$$

Die Nullstellen von $CH_A(X)$ sind – wie gesagt – die Eigenwerte von A. Wegen $\operatorname{grad} CH_A(X) = n$ folgt aus 7.9:

9.7 *Eine Matrix $A \in K^{n\times n}$ hat höchstens n Eigenwerte.* □

Wir bestimmen das Transformationsverhalten. Sei B wie in 9.3, also

$$B = T^{-1}AT.$$

9.8 *a) Die Matrizen A und B haben dieselben Eigenwerte.*

b) Ist x Eigenvektor von A zum Eigenwert λ, so ist $T^{-1}x$ Eigenvektor von B zum Eigenwert λ.

c) $CH_A = CH_B$

Beweis: c)

$$\begin{aligned} CH_B &= \det(B - X\mathbf{1}) = \det(T^{-1}AT - X\mathbf{1}) \\ &\overset{!}{=} \det(T^{-1}AT - T^{-1}(X\mathbf{1})T) = \det(T^{-1}(A - X\mathbf{1})T) \\ &\overset{3.9\ b)}{=} (\det T^{-1})\det(A - X\mathbf{1})\det T(CH_A)\det T^{-1}\det T \\ &= CH_A \det(T^{-1}T) = CH_A. \end{aligned}$$

a) Da die Eigenwerte die Nullstellen des charakteristischen Polynoms sind, folgt a) aus d).

b) $B(T^{-1}x) = T^{-1}ATT^{-1}x = T^{-1}Ax = T^{-1}\lambda x = \lambda(T^{-1}x)$ □

9.9 Satz: *Sei $K = \mathbb{C}$. Dann besitzt A einen Eigenwert.*

Beweis: Nach dem Fundamentalsatz der Algebra 7.10 besitzt das charakteristische Polynom CH_A eine Nullstelle. □

Wir formulieren nun zwei Sätze, deren Beweis wir in den Fällen $n = 2,\ n = 3$ vorführen. Im allgemeinen Fall argumentiert man analog. Wie vorher sei $A \in K^{n\times n}$.

9.10 *Sei $K = \mathbb{C}$. Dann ist A ähnlich zu einer Dreiecksmatrix (man sagt A ist* triagonalisierbar*).*

9.11 Satz: *Die Matrix A besitze n verschiedene Eigenwerte. Dann ist A ähnlich zu einer Diagonalmatrix (man sagt A ist* diagonalisierbar*).*

Bezüglich der Basis $e_1, \ldots, e_n$ von $V = K^n$ definiert A eine lineare Abbildung α, also $\alpha(x) = Ax$. Wir nennen $e_1, \ldots, e_n$ wieder eine g–Basis, setzen also $g_i := e_i,\ i = 1, \ldots, n$. Im Folgenden konstruieren wir h–Basen. Bezüglich einer solchen Basis wird α durch eine zu A ähnliche Matrix B beschrieben (9.4).

n=2: Sei λ Eigenwert von A; im Fall $K = \mathbb{C}$ besitzt A auf jedem Fall einen solchen (9.9). Sei h_1 Eigenvektor von A und wähle $h_2 \in V$, sodass $h_2 \notin Kh_1$. Dann sind die Vektoren h_1, h_2 linear unabhängig, bilden also eine h–Basis. Es ist

$$\alpha(h_1) = Ah_1 = \lambda h_1$$

und $\alpha(h_2)$ hat die Form $\alpha(h_2) = \mu_1 h_1 + \mu_2 h_2,\ \mu_i \in K$. Nach 9.1 hat B die Form

$$B = \begin{pmatrix} \lambda & \mu_1 \\ 0 & \mu_2 \end{pmatrix}.$$

B ist also Dreiecksmatrix. Damit ist 9.10 im Fall $n = 2$ bewiesen. Beim Beweis von 9.11 können wir annehmen, dass μ ein weiterer Eigenwert von A ist, $\mu \neq \lambda$. Sei h_2 Eigenvektor von μ. Für alle $x \in Kh_1$ gilt $Ax = \lambda x$, denn $Ah_1 = \lambda h_1$. Also ist h_2 nicht aus Kh_1. Die Eigenvektoren h_1, h_2 bilden also eine Basis. Es ist

$$\alpha(h_1) = \lambda h_1 \quad \text{und} \quad \alpha(h_2) = \mu h_2.$$

Danach ist $B = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$ eine Diagonalmatrix (siehe 9.1) und 9.11 ist bewiesen.

n=3: Sei λ Eigenwert von A; im Fall $K = \mathbb{C}$ besitzt A einen solchen. Sei h_1 Eigenvektor zu λ. Wir wählen $h_2 \in V \backslash Kh_1$ und $h_3 \in V \backslash \langle h_1, h_2 \rangle$. Dann ist h_1, h_2, h_3 Basis von V und $\alpha(h_1) = \lambda h_1$. Also hat B die Gestalt

$$B = \begin{pmatrix} \lambda & & * & & *_2 \\ 0 & | & -- & -- & -- \\ & | & & B_1 & \\ 0 & | & & & \end{pmatrix} = T^{-1}AT.$$

Nach dem im Fall $n = 2$ Bewiesenem existiert eine Transformationsmatrix $S_1 \in GL(2, K)$, sodass $S_1^{-1} B_1 S_1$ Dreiecksmatrix ist. Setzt man

$$S := \begin{pmatrix} 1 & & * & & * \\ 0 & | & -- & -- & -- \\ & | & & S_1 & \\ 0 & | & & & \end{pmatrix},$$

so ist

$$S^{-1}BS = \left(\begin{array}{c|ccc} 1 & * & & * \\ 0 & -- & ------- & -- \\ & & S_1^{-1}B_1S_1 & \\ 0 & & & \end{array}\right)$$

eine Dreiecksmatrix. Außerdem ist

$$S^{-1}BS = S^{-1}T^{-1}ATS = (TS)^{-1}A(TS).$$

Damit ist $S^{-1}BS$ ähnlich zu A. Dies beweist 9.10 im betrachteten Fall.

Wir kommen zum Beweis von 9.11. Seien $\lambda_1, \lambda_2, \lambda_3$ die drei verschiedenen Eigenwerte von A und h_1, h_2, h_3 die dazugehörenden Eigenvektoren. Wir argumentieren, dass h_1, h_2, h_3 eine Basis von V bilden. Wegen $\alpha(h_i) = \lambda_i h_i$ ist dann

$$B = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

eine Diagonalmatrix, was 9.11 beweist.

Wir nehmen an, dass h_1, h_2, h_3 keine Basis von V ist und führen dies zu einem Widerspruch. Sei zunächst $h_1 = \mu h_2,\ \mu \neq 0$, angenommen. Dann ist

$$\lambda_1 h_1 = \alpha(h_1) = \alpha(\mu h_2) = \mu\alpha(h_2) = \mu\lambda_2 h_2 = \lambda_2 h_1.$$

Dies widerspricht $\lambda_1 \neq \lambda_2$ (beachte $h_1 \neq 0$). Also sind die Vektoren h_1, h_2 linear unabhängig. Im Fall $h_3 \notin \langle h_1, h_2\rangle$ sind h_1, h_2, h_3 linear unabhängig, also Basis. Sei $h_3 \in \langle h_1, h_2\rangle$ angenommen. Dann existieren $\mu_1, \mu_2 \in K$ mit $h_3 = \mu_1 h_1 + \mu_2 h_2$. Es folgt

$$\begin{aligned} \lambda_3\mu_1 h_1 + \lambda_3\mu_2 h_2 &= \lambda_3 h_3 = \alpha(h_3) = \alpha(\mu_1 h_1 + \mu_2 h_2) \\ &= \mu_1\alpha(h_1) + \mu_2\alpha(h_2) \\ &= \mu_1\lambda_1 h_1 + \mu_2\lambda_2 h_2 \end{aligned}$$

Da h_1, h_2 linear unabhängig, also eine Basis von $\langle h_1, h_2\rangle$ sind, folgt per Koeffizientenvergleich

$$\lambda_3\mu_1 = \mu_1\lambda_1 \quad \text{und} \quad \lambda_3\mu_2 = \mu_2\lambda_2.$$

Im Fall $\mu_1 = 0 = \mu_2$ ist $h_3 = 0$. Aber h_3 ist $\neq 0$, da h_3 Eigenvektor. Also ist $\mu_1 \neq 0$ oder $\mu_2 \neq 0$, d.h. $\lambda_3 = \lambda_1$ oder $\lambda_3 = \lambda_2$ entgegen unserer Annahme.

Damit ist auch 9.11 bewiesen (im Fall $n = 3$).

Für $\lambda \in K$ sei

$$E_\lambda := \{x \in K^n \mid Ax = \lambda x\}.$$

Offenbar ist E_λ Unterraum von K^n und $E_\lambda \neq 0$, wenn λ Eigenwert von A ist. In diesem Fall ist E_λ der Lösungsraum des homogenen linearen Gleichungssystems

$$(A - \lambda \mathbf{1}_n)x = 0.$$

Die Elemente $x \neq 0$ sind die Eigenvektoren von A zum Eigenwert λ; man nennt E_λ den *Eigenraum* zum Eigenwert λ. Die Zahl $\dim E_\lambda$ ist die *geometrische Vielfachheit* des Eigenwerts λ. Folgender Satz ist eine Verallgemeinerung von 9.11:

9.12 Satz: *Seien $\lambda_1, \ldots, \lambda_m$ die Eigenwerte von $A \in K^{n\times n}$. Ist*

$$\sum_{i=1}^{m} \dim E_{\lambda_i} = n,$$

so ist A diagonalisierbar.[26]

Beweis: Wir beschränken uns wieder auf den Fall $n = 3$. Ist $m = 3$, also sind $\lambda_1, \lambda_2, \lambda_3$ verschieden, so folgt die Behauptung aus 9.11. Im Fall $m = 1$ ist $\lambda = \lambda_1$ der einzige Eigenwert von A und $\dim E_\lambda = 3$, also $E_\lambda = K^3$. Es folgt

$$a_{*j} = Ae_j = \lambda e_j \quad , \quad j = 1, 2, 3$$

und damit $A = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}$.

Sei $m = 2$ und $\lambda := \lambda_1$, $\mu := \lambda_2$. Wir können $\dim E_\lambda = 2, \dim E_\mu = 1$ annehmen. Sei h_1, h_2 Basis von E_λ und h_3 Basis von E_μ. Im Fall $h_3 \in E_\lambda$ ist $Ah_3 = \lambda h_3$ im Widerspruch zu $Ah_3 = \mu h_3$. Also ist $h_3 \notin \langle h_1, h_2\rangle$ und h_1, h_2, h_3

[26] Ohne weiteres kann man auch die umgekehrte Aussage beweisen.

Basis von K^3. Bezüglich dieser Basis gehört zu α (definiert durch A) die zu A ähnliche Matrix

$$B = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{pmatrix}. \qquad \square$$

10 Der Abstand und die Orthogonalität

Definition: Für $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ ist

$$|x| := \sqrt[+]{x_1^2 + \cdots + x_n^2}$$

die *Länge* (oder der *Betrag* oder die *Norm*) von x.

Im Falle $n = 1$ ist $|x|$ der Betrag der reellen Zahl x.

Bemerkung: Nach "Pythagoras" ist $|x|$ der Abstand der Punkte $0 = (0, \ldots, 0)$ und x im $\mathbb{R}^n$. Ist $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ ein weiterer Punkt, so heißt die Zahl

$$|x - y| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

der *Abstand* der Punkte x, y.

Offenbar ist:

$$|\lambda x| = |\lambda| \, |x| \, , \quad \lambda \in \mathbb{R}, \; x \in \mathbb{R}^n$$

Definition: Seien $x, y \in \mathbb{R}^n$. Die Zahl

$$[x, y] \; := \; \sum_{i=1}^{n} x_i y_i$$

ist das (*euklidische*) *Skalarprodukt* von x mit y.

Wir haben folgende Rechenregeln:

10.1 *a)* $[\lambda x + \mu x, y] = \lambda [x, y] + \mu [z, y]$, *d.h.* $[\ ,\ ]$ *ist* linear im ersten Argument.

b) $[x, \lambda y + \mu z] = \lambda[x, y] + \mu[y, z]$, *d.h.* [,] *ist* linear im zweiten Argument.

c) $[x, y] = [y, z]$, *d.h.* [,] *ist* symmetrisch.

d) $[x, x] \geq 0$ *und* $[x, x] = 0 \Leftrightarrow x \neq 0$, *d.h.* [,] *ist* positiv definiert.

Bemerkung: Nimmt man den Körper $\mathbb{C}$ anstatt $\mathbb{R}$, so definiert man ähnlich:

$$|x| := \sqrt[+]{|x_1|^2 + \ldots + |x_n|^2},$$

dabei ist $|x_i|^2 = x_i \overline{x}_i$ für $x_i \in \mathbb{C}$, siehe **7**.

$$[x, y] := \sum_{i=1}^{n} x_i \overline{y}_i$$

Dann gelten a) und d). Anstatt c) hat man

$$[x, y] = \overline{[y, x]} \qquad (hermitisch)$$

und anstatt b) hat man

$$[x, \lambda y + \mu z] = \overline{\lambda}[x, y] + \overline{\mu}[x, z] \qquad (semilinear)$$

Alles Folgende gilt auch in diesem komplexen Fall. Wir argumentieren aber vorwiegend im reellen Fall.

10.2 Cauchy-Schwarzsche Ungleichung: *Für alle* $x, y \in \mathbb{R}^n$ *ist*

$$|[x, y]| \ \leq \ |x| \cdot |y|.$$

Beweis: Infolge der Rechengesetze 10.1 gilt für $\lambda, \mu \in \mathbb{R}$

$$\begin{aligned} 0 \ &\leq \ |\lambda x - \mu y|^2 = [\lambda x - \mu x, \ \lambda x - \mu y] \\ &= \ \lambda^2 [x, x] - 2\lambda\mu [x, y] + \mu^2 [y, y], \end{aligned}$$

d h.

$$2\lambda\mu [x, y] \ \leq \ \lambda^2 [x, x] + \mu^2 [y, y].$$

Wir setzen $\lambda = |y|$ und $\mu = |x|$ und erhalten

$$2|x| \ |y| \ [x, y] \ \leq \ 2|x|^2 \, |y|^2.$$

Da die Behauptung für $x = 0$ oder $y = 0$ richtig ist, können wir $x \neq 0 \neq y$ annehmen, und erhalten durch Kürzen mit $|x|$ und $|y|$

$$[x, y] \ \leq \ |x| \cdot |y|.$$

Mit $-x$ statt x gilt dieselbe Ungleichung, also

$$-[x, y] \ = \ [-x, y] \ \leq \ |x| \cdot |y|.$$

Es folgt die Behauptung. □

10.3 Dreiecksungleichung: $|x + y| \ \leq \ |x| + |y|.$

Beweis:

$$\begin{aligned} |x+y|^2 &= [x+y,\ x+y] \\ &= |x|^2 \ + \ |y|^2 \ + \ 2[x,y] \\ &\overset{10.2}{\leq} |x|^2 + |y|^2 + 2|x| \ |y| \\ &= (|x| \ + \ |y|)^2 \end{aligned}$$

□

Bemerkung im $\mathbb{R}^2$*:* Ist $x = (x_1, x_2) \in \mathbb{R}^2$ mit $|x| = 1$, so existiert ein Winkel φ, mit

$$0 \leq \varphi < \pi \quad \text{und} \quad x_1 = \cos\varphi,\ x_2 = \sin\varphi,$$

also

$$x = (\cos\varphi, \sin\varphi).$$

Ist $y \in \mathbb{R}^2$ ein weiterer Punkt vom Betrag 1, also

$$y = (\cos\psi, \sin\psi), \quad 0 \leq \psi \leq \pi,$$

so ist

$$[x, y] = \cos\varphi\cos\psi + \sin\varphi\sin\psi \overset{!}{=} \cos(\varphi - \psi).$$

Definition: Seien $x, y \in \mathbb{R}^n,\ x \neq 0 \neq y$. Nach 10.2 ist

$$-1 \ \leq \ \frac{[x, y]}{|x| \cdot |y|} \ \leq \ 1.$$

Also existiert genau ein Winkel $\varphi, 0 \leq \varphi \leq \pi$ mit

$$\cos\varphi = \frac{[x,y]}{|x|\ |y|}.$$

Wir nennen φ den Winkel zwischen x und y (hängt nicht von der Reihenfolge von x, y ab!).

Hier gilt

$$\varphi = 90^0 \quad \Longleftrightarrow \quad [x,y] = 0.$$

Wir schreiben

$$x \perp y \quad \Longleftrightarrow \quad [x,y] = 0.$$

10.4 (n-dimensionaler Pythagoras) *Seien $v_1, \ldots, v_r \in \mathbb{R}^n$ Vektoren, sodass $v_i \perp v_j$ für $i \neq j$. Dann gilt:*

$$|v_1 + v_2 + \cdots + v_r|^2 \ = \ |v_1|^2 + |v_2|^2 + \cdots + |v_r|^2.$$

Beweis:

$$[v_1 + \cdots + v_r, v_1 + \ldots v_r] = \sum_{i,j=1}^{r} [v_i, v_j] = \sum_{i=1}^{r} [v_i, v_i] \qquad \square$$

10.5 *Seien $v_1, \ldots, v_r$ wie in 10.4. Ist $v_i \neq 0$ für $i = 1, \ldots, r$, so sind $v_1, \ldots v_r$ linear unabhängig.*

Beweis: Seien $\lambda_1, \ldots, \lambda_m \in K$ und

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

Dann folgt für $j = 1, \ldots, m$

$$0 = [v_j, \sum_i \lambda_i v_i] \ = \ \sum_i \lambda_i [v_i, v_j] = \lambda_j [v_j, v_j],$$

also $\lambda_j = 0$, da $[v_j, v_j] \neq 0$. $\square$

Definition: Für eine Menge $U \subseteq \mathbb{R}^n$ sei

$$U^\perp := \{x \in \mathbb{R}^n \mid [u,x] = 0 \quad \text{für alle} \quad u \in U\}.$$

Ist U Unterraum von $\mathbb{R}^n$, so heißt $U^\perp$ *orthogonales Komplement* von U in $\mathbb{R}^n$.

Bemerkungen:

1) $U^\perp$ ist Unterraum von $\mathbb{R}^n$.

2) $U^\perp = \mathbb{R}^n \iff U = \{0\}$

3) Sei $U = \{u_1, \ldots, u_m\}$ eine endliche Menge und

$$u_i = (a_{i1}, \ldots, a_{in}) \in \mathbb{R}^n, \qquad i = 1, \ldots, m.$$

Dann ist $U^\perp$ die Lösungsmenge des homogenen linearen Gleichungssystems

$$(*) \qquad \begin{array}{rcl} a_{11}x_1 & + \cdots + & a_{1n}x_n = 0 \\ \vdots & & \vdots \\ a_{m1}x_1 & + \cdots + & a_{nn}x_n = 0. \end{array}$$

10.6 *Ist U Unterraum von $\mathbb{R}^n$ so gilt* $\dim U^\perp = n - \dim U.$

Beweis: Sei $u_1, \ldots, u_m$ Basis von U, jedes $u \in U$ ist also Linearkombination

$$u = \sum_{i=1}^{m} \lambda_i u_i.$$

Wegen

$$[u, x] = \sum_{i=1}^{n} \lambda_i [u_i, x]$$

gilt $x \in U^\perp \Leftrightarrow [u_i, x] = 0,\ i = 1, \ldots, m.$ Damit ist $U^\perp$ Lösungsraum des Gleichungssystems $(*)$. Da $u_1, \ldots, u_m$ die Zeilen der Koeffizientenmatrix sind, ist Rang $(a_{ij}) = m$. Die Behauptung folgt aus 4.21. □

Definition: Seien $u_1, \ldots, u_m \in \mathbb{R}^n$. Dann heißt $u_1, \ldots, u_m$ ein *Orthonormalsystem* (ON–*System*), wenn gilt:

- $|u_i| = 1 \qquad\qquad i = 1, \ldots, m.$
- $[u_i, u_j] = 0$ für alle $i, j \in \{1, \ldots, m\}$ mit $i \neq j.$

Ist $u_1, \ldots, u_m$ Basis eines Unterraums U von $\mathbb{R}^n$, so heißt $u_1, \ldots, u_m$ ON-Basis von U.

Zum Beispiel ist $e_1, \ldots, e_n$ eine ON-Basis von $U = \mathbb{R}^n$.

Dieselben Vorteile wie die ON-Basis $e_1, \ldots, e_n$ hat jede ON-Basis $h_1, \ldots, h_n$ von $\mathbb{R}^n$: Sei $x \in \mathbb{R}^n$ und

$$x = \sum_{i=1}^{n} y_i h_i \quad , \quad y = \kappa_h(x).$$

Dann folgt aus den Rechengesetzen 10.1

10.7 $|x|^2 = \sum_{i=1}^{n} y_i^2$ *und* $y_i = [h_i, x], i = 1, \ldots, n.$

10.8 *Sei* $u_1, \ldots, u_m$ *eine* ON*-Basis des Unterraums* U *von* $\mathbb{R}^n$ *und sei* $x \in \mathbb{R}^n$. *Setze*

$$\begin{aligned} a &:= \sum_{i=1}^{m} [u_i, x] u_i \ \in U && \textit{(Lotpunkt)} \\ b &:= x - a && \textit{(Lot)} \end{aligned}$$

also $x = a + b$. *Dann gilt*

a) $b \in U^{\perp}$. *Insbesondere ist* $a \perp b$.

b) $|b| < |x - u|$ *für alle* $u \in U$ *mit* $u \neq a$.

Beweis: a) Für $j = 1, \ldots, m$ gilt

$$\begin{aligned} [u_j, b] &= [u_j, x] - [u_j, a] = [u_j, x] - [u_j, \sum_i [u_i, x] u_i] \\ &= [u_j, x] - \sum_i [u_i, x][u_i, u_j] = [u_j, x] - [u_j, x] \\ &= 0. \end{aligned}$$

b) Es ist $x - u = (x - a) + (a - u)$, mit $a - u \in U$. Nach a) ist $[x - a, a - u] = 0$. Mit 10.4 folgt

$$|x - u|^2 = |x - a|^2 + |a - u|^2 > |x - a|^2,$$

da $a - u \neq 0$. □

10.9 Schmidtsches Orthogonaliserungsverfahren: Seien $w_1, \ldots, w_k$ Vektoren $\neq 0$ von $\mathbb{R}^n$ und sei

$$U := < w_1, \ldots, w_k >, \quad m := \dim U.$$

Wir konstruieren induktiv eine ON-Basis $v_1, \ldots, v_m$ von U. Dazu setzen wir für $i = 1, \ldots, k$

$$U_i := < w_1, \ldots, w_i > .$$

Start: Setze

$$v_1 = \frac{1}{|w_1|} w_1.$$

Dann ist v_1 eine ON-Basis von $U_1 = < w_1 >$.

Sei $i \in \{1, \ldots, k-1\}$ und sei bereits eine ON-Basis $v_1, \ldots, v_{m_i}$ von U_i konstruiert $(m_i = \dim U_i)$.

Schritt $i \longrightarrow i+1$: Wir unterscheiden 2 Fälle:

- Sei $w_{i+1} \in U_i$. Dann ist $U_i = U_{i+1}$ und $v_1, \ldots, v_{m_i}$ eine ON-Basis von U_{i+1}.
- Sei $w_{i+1} \notin U_i$. Nach 10.8 sei b das Lot von $x = w_{i+1}$ auf den Unterraum U_i, also

  $$b = w_{i+1} - \sum_{i=1}^{m_i} [v_i, w_{i+1}] v_i.$$

  Setze

  $$v_{m_i+1} := \frac{1}{|b|} b.$$

  Wegen $U_{i+1} = < v_1, \ldots, v_{m_i}, w_{i+1} > = < v_1, \ldots, v_{m_i}, b > = < v_1, \ldots, v_{m_i}, v_{m_i+1} >$ ist dann $v_1, \ldots, v_{m_i}, v_{m_i+1}$ eine ON-Basis von U_{i+1}.

Nach k Schritten erhält man eine ON-Basis von U. □

Sei $U \neq 0$ ein Unterraum von $\mathbb{R}^n$ und $w_1, \ldots, w_m$ eine Basis von U. Dann ist $U = < w_1, \ldots, w_m >$ und das beschriebene Orthogonalisierungsverfahren liefert:

10.10 *Ein Unterraum $U \neq 0$ des $\mathbb{R}^n$ besitzt eine* ON*-Basis.* □

Der Algorithmus 10.9 zeigt auch:

10.11 Satz: *Sei U ein Unterraum von $\mathbb{R}^n$ und $v_1, \ldots, v_m$ eine* ON-Basis *von U. Dann existieren Vektoren $v_{m+1}, \ldots, v_n$ von $\mathbb{R}^n$, sodass $v_1, \ldots, v_m, v_{m+1}, \ldots, v_n$ eine* ON-Basis *von $\mathbb{R}^n$ ist.*

Beweis: Im Fall $U = V$ ist nichts zu beweisen. Sei $U \neq V$. Dann existiert $v_{m+1} \in V \backslash U$. Nach dem Verfahren 10.9 können wir v_{m+1} so wählen, dass $v_1, \ldots, v_m, v_{m+1}$ ein ON–System ist. Sei

$$U_1 := \langle v_1, \ldots, v_m, v_{m+1} \rangle .$$

Im Fall $V = U_1$ sind wir fertig. Sei $V \neq U_1$. Wähle $v_{m+2} \in V \backslash U_1$ und "orthonormalisiere v_{m+2}". so fortfahrend erhält man nach $n-m$ Schritten die gesuchte ON–Basis. □

Bemerkung: In der Situation 10.11 gilt $U^\perp = < v_{m+1}, \ldots, v_n >$.

11 Bewegungen und orthogonale Matrizen

Hier ist wieder $K = \mathbb{R}$; für $K = \mathbb{C}$ gilt Analoges, wenn man das für $\mathbb{C}$ definierte Skalarprodukt nimmt.

Definition: Eine Abbildung $\delta : \mathbb{R}^n \to \mathbb{R}^n$ heißt Bewegung, wenn δ den Abstand erhält, d.h. für alle $x, y \in \mathbb{R}^n$ gilt:

$$|\delta(x) - \delta(y)| \ = \ |x - y|.$$

Sind δ_1, δ_2 zwei Bewegungen, so ist

$$\begin{aligned} |(\delta_1 \circ \delta_2)(x) - (\delta_1 \circ \delta_2)(y)| &= |\delta_1(\delta_2(x) - \delta_1(\delta_2(y))| \\ &= |\delta_2(x) - \delta_2(y)| \\ &= |x - y|, \end{aligned}$$

d.h.

- *Die Komposition $\delta_1 \circ \delta_2$ von zwei Bewegungen ist wieder eine Bewegung.*[27]

Eine Abbildung $\tau : \mathbb{R}^n \to \mathbb{R}^n$ heißt eine *Verschiebung* (*Translation*), wenn es ein $a \in \mathbb{R}^n$ gibt mit

$$\tau(x) = x + a \quad \text{für alle} \quad x \in \mathbb{R}^n;$$

Wir schreiben $\tau = \tau_a$.

Wegen

$$|\tau(x) - \tau(y)| = |x + a - (y + a)| = |x - y|$$

können wir feststellen:

- *Verschiebungen sind Bewegungen.*

Sei δ eine Bewegung und sei $a \in \mathbb{R}^n$ definiert durch

$$\delta(0) = a.$$

Dann ist $\tau_{-a}(a) = 0$, also

$$(\tau_{-a} \circ \delta)(0) = 0.$$

Setzt man $\alpha := \tau_{-a} \circ \delta$, so ist $\alpha(0) = 0$ und $\delta = \tau_a \circ \alpha$.

- *Zu jeder Bewegung δ gibt es eine Verschiebung τ und eine Bewegung α, sodass*

$$\delta = \tau \circ \alpha \quad \textit{und} \quad \alpha(0) = 0.$$

Wir bezeichnen mit $\mathbf{O}_n$ die Menge aller Bewegungen α, für die $\alpha(0) = 0$ gilt (*orthogonale Gruppe*).

11.1 *Sei $\alpha \in \mathbf{O}_n$. Dann gilt für alle $x, y \in \mathbb{R}^n$*

$$[\alpha(x),\ \alpha(y)] = [x, y].$$

[27] Man überzeugt sich leicht, dass die Bewegungen bez. der Komposition eine Gruppe bilden.

Beweis: Zunächst gilt

$$|\alpha(x)| = |\alpha(x) - 0| \stackrel{!}{=} |\alpha(x) - \alpha(0)| = |x - 0| = |x|.$$

Es ist

$$|x - y|^2 = [x - y, x - y] = |x|^2 + |y|^2 - 2[x, y],$$

also

$$[x, y] = \frac{1}{2}(|x|^2 + |y|^2 - |x - y|^2).$$

Genauso folgt mit $\alpha(x), \alpha(y)$ anstatt x, y

$$[\alpha(x), \alpha(y)] = \frac{1}{2}\Big(|\alpha(x)|^2 + |\alpha(y)|^2 - |\alpha(x) - \alpha(y)|^2\Big) = [x, y],$$

denn $|\alpha(x)| = |x|$ und $|\alpha(y)| = |y|$. □

Aus 11.1 folgt (beachte 10.5):

11.2 *Sei $\alpha \in \mathbf{O}_n$ und $v_1, \ldots v_n$ eine* ON*-Basis von $\mathbb{R}^n$. Dann ist auch $\alpha(v_1), \ldots, \alpha(v_n)$ eine* ON*-Basis von $\mathbb{R}^n$.* □

11.3 *Sei $\alpha \in \mathbf{O}_n$. Dann ist $\alpha : \mathbb{R}^n \to \mathbb{R}^n$ eine lineare Abbildung, die bijektiv ist.*

Beweis: Sei $v_i := \alpha(e_i)$ für $i = 1, \ldots n$. Nach 11.2 ist $v_1, \ldots, v_n$ eine ON-Basis, da $e_1, \ldots, e_n$ eine ist. Ist

$$x = (x_1, \ldots, x_n) = \sum_i x_i e_i \in \mathbb{R}^n,$$

so ist für $i = 1, \ldots, n$

$$x_i = [x, e_i] \stackrel{11.1}{=} [\alpha(x), \alpha(e_i)] = [\alpha(x), v_i].$$

Nach 10.7 ist $(x_1, \ldots, x_n)$ das Koordinatentupel von $\alpha(x)$ bez. der Basis $v_1, \ldots, v_n$. Es gilt also

$$\alpha\Big(\sum_i x_i e_i\Big) = \sum_i x_i \alpha(v_i),$$

woraus die Behauptung folgt. □

Definition: Eine Matrix $A \in \mathbb{R}^{n \times n}$ heißt *orthogonal*, wenn

- $A\,A^T = \mathbf{1}_n$ (A^T ist die transponierte)

oder äquivalent

- Die Zeilen von A bilden eine Orthonormalbasis des $\mathbb{R}^n$

oder äquivalent

- Die Spalten von A bilden eine Orthonormalbasis des $\mathbb{R}^n$.

Zu Letzterem beachte:

$$A\,A^T = \mathbf{1}_n \iff A^T = A^{-1} \iff A^T A = 1.$$

Bemerkung: Wir notieren den Fall $K = \mathbb{C}$. Eine Matrix $A \in \mathbb{C}^{n\times n}$ heißt *unitär*, wenn

$$\overline{A}A^T = \mathbf{1}_n.$$

Dabei ist $\overline{A} := (\overline{a_{ij}})$, wenn $A = (a_{ij})$. In diesem Fall bilden die Zeilen und die Spalten eine ON–Basis des $\mathbb{C}^n$. Hier ist

$$A^{-1} = \overline{A}^T \qquad \text{(adjungierte Matrix)} \quad .$$

11.4 *Zu einer orthogonalen Transformation $\alpha : \mathbb{R}^n \to \mathbb{R}^n$ aus $\mathbf{O}_n$ gehört bez. einer ON-Basis eine orthogonale Matrix.*

Beweis: Sei $v_1, \ldots, v_n$ eine ON-Basis. Für $x \in V$ sei

$$\begin{aligned} x &= \sum_{i=1}^{n} x_i v_i \\ \alpha(x) &= \sum_{i=1}^{n} y_i v_i. \end{aligned}$$

Dann ist

$$(y_1, \ldots, y_n) = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

wobei die j-te Spalte von $A \in \mathbb{R}^{n\times n}$ das Koordinatentupel von $\alpha(v_j)$ bez. der Basis $v_1, \ldots, v_n$ ist (siehe 9.1). Nach 11.2 ist aber $\alpha(v_1), \ldots, \alpha(v_n)$ eine Orthonormalbasis. Damit bilden die Spalten von A eine ON-Basis, d.h. A ist orthogonal. □

Aufgrund von 2.4 bezeichnen wir mit $\mathbf{O}_n$ auch die Menge aller orthogonalen $n \times n$ Matrizen.

11.5 *Sei $A \in \mathbf{O}_n$. Dann gilt:*

a) *Ist $\lambda \in \mathbb{R}$ Eigenwert von A, so ist $\lambda = \pm 1$*

b) $\det A = \pm 1$.

Beweis: a) Sei x Eigenvektor zu λ. Dann ist

$$[x, x] = [Ax, Ax] = [\lambda x, \lambda x] = \lambda^2 [x, x].$$

Wegen $x \neq 0$ ist $[x, x] \neq 0$, es kann also durch $[x, x]$ gekürzt werden. Es folgt $\lambda^2 = 1$ und damit die Behauptung.

b) Nach 3.9 und 3.10 ist

$$1 = \det(A\,A^T) = (\det A)(\det A^T) = (\det A)(\det A),$$

also $\det A = \pm 1$. □

Wir diskutieren den Fall $\boldsymbol{n = 2}$: Sei

$$\alpha : \mathbb{R}^2 \to \mathbb{R}^2$$

eine orthogonale Transformation; bezüglich der ON-Basis gehöre zu α die orthogonale Matrix $A = (a_{ij})$. Nach 11.2 ist auch

$$v_1 := \alpha(e_1) \quad \text{und} \quad v_2 := \alpha(e_2)$$

eine ON-Basis des $\mathbb{R}^2$. Dabei sind v_1, v_2 die zwei Spalten von A. Sei

$$v_1 = (\cos\varphi,\, \sin\varphi),\ 0 \leq \varphi < 2\pi.$$

Da v_2 senkrecht auf v_1 steht, hat v_2 die Gestalt D) oder S).

D) $v_2 = \Big(\cos(\varphi + \tfrac{\pi}{2}), \sin\varphi + \tfrac{\pi}{2})\Big) = (-\sin\varphi,\, \cos\varphi).$

S) $v_2 = \Big(\cos(\varphi - \frac{\pi}{2}),\, \sin(\varphi - \frac{\pi}{2})\Big) = (\sin\varphi, -\cos\varphi)$.

Im Falle D) ist

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

und α eine *Drehung* um den Winkel φ; siehe Seite 27. Es ist $\det A = 1$ und A besitzt nur in den Fällen $\varphi = 0$ und $\varphi = \pi$ Eigenwerte in $\mathbb{R}$.

Es gelte S). Dann ist

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

eine symmetrische Matrix mit $\det A = -1$ und den Eigenwerten $1, -1$. Wegen

$$A = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

ist

$$\alpha = \beta \circ \sigma$$

wobei β eine Drehung um den Winkel φ und σ die Spiegelung an der x-Achse ist. Aus elementargeometrischen Überlegungen folgt nun, dass α eine Spiegelung an den Geraden G ist, die man erhält, wenn man die x-Achse um den Winkel $\frac{\varphi}{2}$ dreht. Auf G liegt der Vektor

$$h_1 = (\cos\frac{\varphi}{2}, \sin\frac{\varphi}{2})$$

und

$$h_2 = (\sin\frac{\varphi}{2}, -\cos\frac{\varphi}{2})$$

ist senkrecht zu h_1. Dann ist h_1, h_2 eine ON-Basis mit

$$\alpha(h_1) = h_1,\ \alpha(h_2) = -h_2.$$

Bezüglich dieser Basis gehört zu α die Matrix

$$B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Ist also

$$T = \begin{pmatrix} \cos\frac{\varphi}{2} & \sin\frac{\varphi}{2} \\ \sin\frac{\varphi}{2} & -\cos\frac{\varphi}{2} \end{pmatrix}$$

die dazugehörige Transformationsmatrix, so ist

$$T^{-1}AT = B,$$

und diese Beziehung kann man auch durch Ausrechnen verifizieren.

12 Quadratische Formen und die Hauptachsentransformation

Im Folgenden sei $K = \mathbb{R}$.

Definition: Eine Abbildung $q : K^n \to K^n$ heißt quadratische Form, wenn Zahlen $a_{ij} \in K$ existieren mit $a_{ji} = a_{ij}$, sodass

$$q(x) = \sum_{i,j=1}^{m} a_{ij} x_i x_j \quad , \quad x = (x_1, \ldots, x_n).$$

Wir machen Beispiele:

1) $q(x) = |x|^2 = \sum_{i=1}^{n} x_i^2$, hier ist $a_{ii} = 1$ und $a_{ij} = 0$ für $i \neq j$.

2) $n = 2$ und wie nach 9.1

$$q(x) = 5x_1^2 + x_2^2 + 2x_1x_2.$$

Hier ist $a_{11} = 5,\ a_{22} = 1,\ a_{12} = 1.$

3) $n = 3$ und $q(x) = x_1x_3$. Hier ist $a_{11} = a_{22} = a_{33} = 0 = a_{12} = a_{21}$ und $a_{13} = a_{31} = \frac{1}{2}$.

Sei $q(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$ eine quadratische Form. Wir bilden die Matrix $A = (a_{ij}) \in K^{n\times n}$. Wegen $a_{ij} = a_{ji}$ ist

$$A = A^T$$

eine *symmetrische* Matrix. An Beispielen mache man sich klar:

$$\begin{aligned} q(x) &= (x_1,\ldots,x_n)\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= xAx^T \qquad\qquad \text{(Matrixprodukt)} \end{aligned}$$

Im Beispiel 1 ist $A = \mathbf{1}_n$, im Beispiel 2 ist $A = \begin{pmatrix} 5 & 1 \\ 1 & 1 \end{pmatrix}$ und im Beispiel 3 ist $A = \begin{pmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix}$.

In Kapitel **10** haben wir die Zahl $|x|^2$ eines Vektors $x \in K^n$ in das euklidische Skalarprodukt eingebettet

$$[x,y] = \sum_{i=1}^{n} x_i y_i \quad , \quad |x|^2 = [x,x].$$

Analog gehen wir auch jetzt vor. Einer symmetrischen Matrix $A \in K^{n\times n}$ ordnen wir die Zahl $[x,y] \in K$ zu:

$$\begin{aligned} [x,y] &\overset{\text{def}}{=} xAy^T \\ &= (x_1,\ldots,x_n)\begin{pmatrix} \sum_{j=1}^n a_{1j}y_j \\ \vdots \\ \sum_{j=1}^n a_{nj}y_j \end{pmatrix} \\ &= \sum_{i,j=1}^{n} a_{ij}x_i y_j \end{aligned}$$

Also ist $[x,x] = q(x)$. Die Abbildung

$$[\ ,\ ] : K^n \times K^n \to K$$

heißt *symmetrische Bilinearform*, denn sie ist linear im ersten und zweiten Argument und es ist $[x,y] = [y,x]$; vergleiche 10.1.

Wir kommen wieder auf die quadratische Form $q(x) = xAx^T$ zurück. Sei wieder (wie in **9**) die Basis $e_1,\ldots,e_n$ von K^n unsere g–Basis und sei $h_1,\ldots,h_n$

eine weitere Basis (h–Basis) sowie $T \in GL(n,K)$ die Transformationsmatrix. Die Spalten von T sind die Vektoren $h_1, \ldots, h_n$. Wir schreiben v für x, also ist $\kappa_g(v) = x$; sei $y = \kappa_h(v)$ (siehe Kapitel **9**). Nach 9.2 haben wir $x = Ty$, d.h. in Matrixschreibweise

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = T \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad \text{, d.h.} \quad x^T = Ty^T.$$

Es folgt

$$q(v) = xAx^T = (Ty^T)^T A(Ty^T) = yT^T ATy^T = y(T^T AT)y.$$

Wir haben also das Transformationsgesetz.

12.1 *Gehört zu q die Matrix A bez. der g–Basis, so gehört zu q die Matrix $T^T AT$ bez. der h–Basis.* □

Im Beispiel 1) war $A = \begin{pmatrix} 5 & 1 \\ 1 & 1 \end{pmatrix}$ und $q(v) = xAx^T$.

Wir nehmen die h–Basis $h_1 = (0,1),\ h_2 = \left(\frac{1}{2}, -\frac{1}{2}\right)$. Dann ist $T = \begin{pmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{pmatrix}$ und

$$\begin{aligned} T^T AT &= \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 5 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Wir definieren nun eine andere h–Basis. Zunächst berechnen wir die Eigenwerte von A, also die Nullstellen von

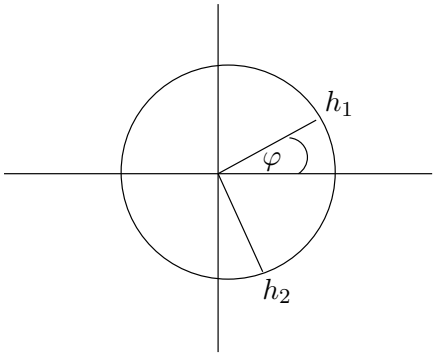
$$CH_A(X) = (5 - X)(1 - X) - 1 = X^2 - 6X + 4$$

zu $\lambda_1 = 3 + \sqrt{5},\ \lambda_2 = 3 - \sqrt{5}$. Zugehörige Eigenvektoren sind

$$\begin{aligned} h_1 &= \frac{1}{\sqrt{10 - 4\sqrt{5}}}(1, \sqrt{5} - 2) \\ h_2 &= \frac{1}{\sqrt{10 + 4\sqrt{5}}}(1, -\sqrt{5} - 2) \end{aligned}$$

Es ist $|h_1| = |h_2| = 1$ und $[h_1, h_2] = 0$, die h–Basis h_1, h_2 ist also ON–Basis. Die dazugehörende Transformationsmatrix T hat die Spalten h_1, h_2, ist also eine orthogonale Matrix. Sie hat Determinante -1. Also existiert ein Winkel φ, sodass (siehe Diskussion nach 11.5)

$$h_1 = (\cos\varphi, \sin\varphi) \quad \text{und} \quad h_2 = (\sin\varphi, -\cos\varphi).$$



Für die orthogonale Matrix T ist $T^T = T^{-1}$. Also ist

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

Sei $\alpha : K^2 \to K^2$ die durch A erklärte Abbildung, also $\alpha(x) = Ax$. Nach dem Transformationsgesetz 9.2 gehört zu α bez. der h–Basis die Matrix $B = T^{-1}AT$. Dabei sind die Spalten von B die Koordinaten von $\alpha(h_1),\ \alpha(h_2)$ bezüglich der h–Basis. Es ist $\alpha(h_1) = \lambda_1 h_1,\ \alpha(h_2) = \lambda_2 h_2$. Damit ist

$$T^{-1}AT = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = T^T A T.$$

Sei $v = x$, also $\kappa_g(v) = x$ und sei $y = \kappa_h(v)$. Nach dem Transformationsgesetz 11.1 ist

$$q(v) = (y_1, y_2) \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \lambda_1 y_1^2 + \lambda_2 y_2^2.$$

Unser $\mathrm{Ei} = \{v \in K^2 \mid q(v) = 1\}$ hat damit bez. der h–Basis die Gleichung

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 = 1 \quad \text{, oder}$$

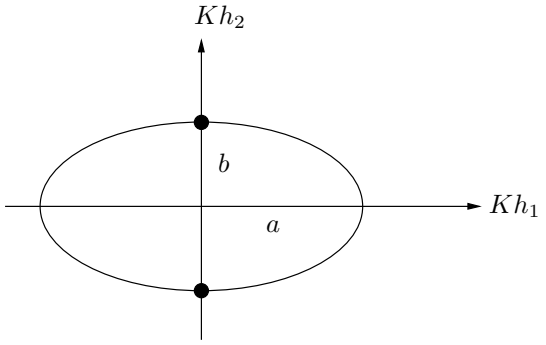
$$\frac{y_1^2}{a^2} + \frac{y_2^2}{b^2} = 1 \quad , \quad a := \frac{1}{\sqrt[+]{\lambda_1}} \quad , \quad b := \frac{1}{\sqrt[+]{\lambda_2}}.$$

Die Geraden Kh_1 und Kh_2 heißen *Hauptachsen* (des $\mathrm{Ei}\,s$ bzw. von A). Seien $\mu_1, \mu_2 \in K$ mit $\mu_1 h_1,\ \mu_2 h_2 \in \mathrm{Ei}$. Es ist $\kappa_h(\mu_1 h_1) = (\mu_1, 0)$ und $\kappa_h(\mu_2 h_2) = (0, \mu_2)$. Also ist

$$\lambda_1 \mu_1^2 = 1 = \lambda_2 \mu_2^2.$$

Es folgt

$$\mu_1 = \pm\frac{1}{\sqrt{\lambda_1}} = \pm a \quad , \quad \mu_2 = \pm\frac{1}{\sqrt{\lambda_2}} = \pm b.$$



Wir haben also die *Ellipse* in ihrer *Normalform* $\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1$, dabei nehmen wir wieder x_i statt y_i.

Im Folgenden sei $A \in K^{n\times n}$ eine symmetrische Matrix und $[\ ,\ ]$ das euklidische Skalarprodukt.

12.2 $[Ax, y] = [x, Ay]$ *für* $x, y \in K^n$.

Beweis: Wir nehmen die Matrixschreibweise

$$[x, y] = (x_1, \ldots, x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = xy^T.$$

Da A symmetrisch ist, gilt

$$Ax = Ax^T \stackrel{!}{=} xA.$$

Da das Matrixprodukt assoziativ ist, folgt

$$[Ax, y] = [xA, y] = (x_1, \ldots, x_n)A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = [x, Ay]. \qquad \square$$

12.3 *Seieh λ, μ verschiedene Eigenwerte von A und x, y die dazugehörigen Eigenvektoren. Dann ist $x \perp y$.*

Beweis: $\lambda[x, y] = [\lambda x, y] = [Ax, y] \stackrel{12.1}{=} [x, Ay] = [x, \mu y] = \mu[x, y]$. Wegen $\lambda \neq \mu$ folgt $[x, y] = 0$. $\square$

Beispiel: Sei $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \in K^{2\times 2}$ symmetrisch. Wir berechnen die Eigenwerte von A, also die Nullstellen des charakteristischen Polynoms

$$(a - X)(c - X) - b^2 = X^2 - (a + b)X + ac - b^2.$$

Wir erhalten

$$\lambda_{1,2} = \frac{1}{2}(a + c \pm \sqrt{(a + c)^2 - 4(ac - b^2)}).$$

Wegen $(a + c)^2 - 4ac = (a - c)^2$ ist der Ausdruck unter der Wurzel positiv. Die symmetrische Matrix A hat also reelle Eigenwerte.

Allgemein gilt:

12.4 Satz: *Eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ besitzt reelle Eigenwerte.*

Beweis: Die reelle Matrix A kann auch als komplexe Matrix aufgefasst werden. Es existiert also eine komplexe Zahl $\lambda \in \mathbb{C}$, welche Eigenwert von A ist (9.9). Wir zeigen $\lambda \in \mathbb{R}$. Dazu bilden wir das komplexe Skalarprodukt

$$[x, y] = \sum_{i=1}^{n} x_i \overline{y}_i$$

für $x, y \in \mathbb{C}^n$. Sei $x \in \mathbb{C}^n$ Eigenvektor von A zum Eigenwert λ. Es folgt

$$\lambda[x,x] = [\lambda x, x] = [Ax, x] = [xA, x] = (x_n, \ldots, x_n)A \begin{pmatrix} \overline{x}_1 \\ \vdots \\ \overline{x}_n \end{pmatrix}.$$

Da $A = (a_{ij})$ eine reelle Matrix ist, gilt $\overline{A} = (\overline{a_{ij}}) = A$. Also ist

$$A\overline{x} = \overline{A}\overline{x} \stackrel{!}{=} \overline{Ax} = \overline{\lambda x} = \overline{\lambda}\overline{x}.$$

Dabei beachte man Rechenregel **4**) in **8**. Es folgt

$$\lambda[x,x] = (x_1, \ldots, x_n) \begin{pmatrix} \overline{\lambda}\overline{x_1} \\ \vdots \\ \overline{\lambda}x_n \end{pmatrix} = \overline{\lambda}(x_1, \ldots, x_n) \begin{pmatrix} \overline{x_1} \\ \vdots \\ x_n \end{pmatrix} = \overline{\lambda}[x,x].$$

Wegen $x \neq 0$ ist $[x, \overline{x}] \neq 0$ und darum $\lambda = \overline{\lambda}$, d.h. $\lambda \in \mathbb{R}$. □

Wir nehmen nun an, dass die symmetrische Matrix n verschiedene Eigenwerte besitzt; seien $h_1, \ldots, h_n$ die dazugehörenden Eigenvektoren. Wir können $|h_1| = \ldots = |h_n| = 1$ annehmen. Nach 12.3 und 10.5 ist dann $h_1, \ldots, h_n$ eine ON–Basis. Die lineare Abbildung $\alpha : K^n \to K^n$ mit $\alpha(x) = Ax$ wird bez. der h–Basis durch eine Matrix $B = T^{-1}AT$ beschrieben. Dabei ist $\alpha(h_i) = \lambda_i h_i$, also

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

eine Diagonalmatrix. Da $h_1, \ldots, h_n$ die Spalten der Transformationsmatrix T sind, ist T eine orthogonale Matrix

$$B = T^{-1}AT = T^TAT,$$

vergleiche 12.2.

Allgemein gilt:

12.5 Hauptachsentransformation: *Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch. Dann existiert eine orthogonale Matrix T, sodass*

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

Dabei sind $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von A.[28]

Beweis: Nach 12.4 besitzt A einen Eigenwert λ_1; sei h_1 ein dazugehörender Eigenvektor, wir können $|h_1| = 1$ annehmen. Nach 10.11 können wir h_1 zu einer ON–Basis $h_1, h_2, \ldots, h_n$ erweitern. Bezüglich dieser h–Basis gehört zu α (definiert durch $\alpha(x) = Ax$) eine Matrix der Form

$$B_1 := T_1^{-1}AT_1 = \begin{pmatrix} \lambda_1 & x \ldots x \\ 0 & \\ \vdots & A_1 \\ 0 & \end{pmatrix} \quad , \quad A_1 \in K^{(n-1)\times(n-1)}.$$

Dabei ist T_1 eine orthogonale Matrix, also $T_1^{-1} = T_1^T$. Es folgt

$$B_1^T = (T_1^T A T_1)^T \overset{3.1}{=} T_1^T A^T (T_1^T)^T = T_1^{-1} A T_1 = B_1.$$

Also ist B_1 symmetrisch. Damit hat die erste Zeile von B_1 die Form $(\lambda_1, 0, \ldots, 0)$. Außerdem folgt, dass auch A_1 symmetrisch ist. Per Induktion nach n existiert eine orthogonale Matrix $S_1 \in K^{(n-1)\times(n-1)}$, sodass $S_1^{-1}A_1S$ eine Diagonalmatrix ist. Sei

$$S := \begin{pmatrix} 1 & 0 \ldots 0 \\ 0 & \\ \vdots & S_1 \\ 0 & \end{pmatrix}$$

und $T := T_1S$. Dann sind S und T orthogonale Matrizen, denn

$$T^{-1} = (T_1S)^{-1} = S^{-1}T_1^{-1} = S^T T_1^T = (T_1S)^T = T^T.$$

[28] Nicht notwendig verschieden!

Überdies ist

$$T^{-1}AT = S^{-1}B_1S = \begin{pmatrix} 1 & 0\dots0 \\ 0 & \\ \vdots & S_1^{-1}A_1S_1 \\ 0 & \end{pmatrix}$$

eine Diagonalmatrix; dabei stehen nach 9.6 die Eigenwerte von A auf der Diagonalen. □

Nach dem Transformationsgesetz 12.1 haben wir das

12.6 Korollar: *Sei q die quadratische Form, definiert durch die symmetrische Matrix $A \in \mathbb{R}^{n\times n}$. Dann existiert eine ON–Basis $h_1, \dots, h_n$ von K^n, sodass*

$$q(x) = \sum_{i=1}^{n} \lambda_i x_i^2.$$

Dabei haben wir den Vektor $x \in K^n$ mit dem Koordinatenvektor $\kappa_h(x)$ identifiziert.

12.7 Korollar:

a) $\lambda_1, \dots, \lambda_n > 0 \;\Leftrightarrow\; q(x) > 0$ *für* $x \neq 0$ *(positiv definiert).*

b) $\lambda_1, \dots, \lambda_n < 0 \;\Leftrightarrow\; q(x) < 0$ *für* $x \neq 0$ *(negativ definiert)*

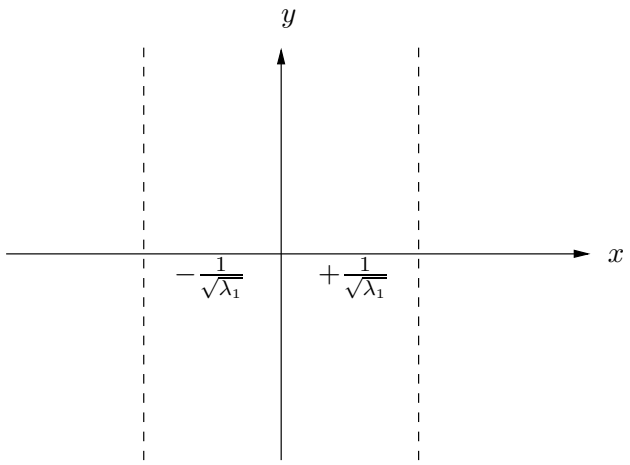
□

Die quadratische Form $q(x)$ heißt *indefiniert*, wenn es Vektoren x, y gibt mit $q(x) < 0 < q(y)$. Dies ist genau dann der Fall, wenn positive und negative Eigenwerte existieren.

Zum Schluss diskutieren wir wieder die Gleichung $q(x) = 1$ im Fall $n = 2$. Sei Q die Mange aller $x = (x_1, x_2)$, für die

$$\lambda_1 x_1^2 + \lambda_2 x_2^2 = 1$$

gilt. Sind $\lambda_1, \lambda_2 > 0$, so ist Q eine *Ellipse*. Ist $\lambda_1, \lambda_2 < 0$, so ist $Q = \emptyset$. Ist $\lambda_1 > 0,\ \lambda_2 < 0$, so ist Q eine *Hyperbel*. Übrig bleiben die *ausgearteten* Fälle $0 \in \{\lambda_1, \lambda_2\}$. Im Fall $\lambda_1 = 0 = \lambda_2$ ist $Q = \emptyset$. Im Fall $\lambda_1 \neq 0,\ \lambda_2 = 0$ ist Q die *Doppelgerade* $\left(\pm\sqrt{\frac{1}{\sqrt{\lambda_1}}}, \mu\right),\ \mu \in \mathbb{R}$.


y
x
$-\frac{1}{\sqrt{\lambda_1}}$
$+\frac{1}{\sqrt{\lambda_1}}$