

## 2 Vektorräume beliebiger Dimensionen

In diesem Abschnitt werden zunächst die Vektorräume  $\mathbb{R}^n$  und  $\mathbb{C}^n$  behandelt, einschließlich linearer Gleichungssysteme. Anschließend werden allgemeinere algebraische Strukturen erörtert: Gruppen, Körper und Vektorräume über beliebigen Körpern samt linearen Abbildungen. Diese abstrakteren Teile (ab Abschn. 2.3) können vom anwendungsorientierten Leser zunächst übersprungen werden.

### 2.1 Die Vektorräume $\mathbb{R}^n$ und $\mathbb{C}^n$

#### 2.1.1 Der Raum $\mathbb{R}^n$ und seine Arithmetik

Analog zum  $\mathbb{R}^2$  und  $\mathbb{R}^3$  führt man  $\mathbb{R}^n$  ein: Der  $\mathbb{R}^n$  ist die Menge aller *reellen Spaltenvektoren*

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (x_1, \dots, x_n \in \mathbb{R}).$$

Die reellen Zahlen  $x_1, \dots, x_n$  heißen dabei die *Koordinaten (Komponenten, Einträge)* des Spaltenvektors  $\mathbf{x}^1$ , und  $n$  ist seine *Dimension*. Zwei Spaltenvektoren  $\mathbf{x}$  und  $\mathbf{y}$ , s. (2.1), sind genau dann *gleich*,  $\mathbf{x} = \mathbf{y}$ , wenn ihre entsprechenden Koordinaten übereinstimmen, d.h. wenn  $x_1 = y_1, x_2 = y_2, \dots, x_n = y_n$  gilt. (Spaltenvektoren verschiedener Dimensionen sind natürlich verschieden.)

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{x} \pm \mathbf{y} = \begin{bmatrix} x_1 \pm y_1 \\ x_2 \pm y_2 \\ \vdots \\ x_n \pm y_n \end{bmatrix}, \quad \lambda \mathbf{x} = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{bmatrix}. \quad (2.1)$$

*Addition, Subtraktion, Multiplikation mit einem Skalar*  $\lambda \in \mathbb{R}$  (auch *s-Multiplikation* genannt), werden mit den Vektoren des  $\mathbb{R}^n$  koordinatenweise ausgeführt, wie in (2.1) angegeben. Es gelten alle Regeln des Satzes 1.1 aus Abschn. 1.1.3 entsprechend: Assoziativgesetz für  $+$ , Kommutativgesetz für  $+$ , usw. Auf Grund dieser Gesetze nennt man  $\mathbb{R}^n$  einen *n-dimensionalen reellen Vektorraum*.

Klammern werden bei längeren Summen und Multiplikationen mit Skalaren normalerweise

---

<sup>1</sup> Auch die waagerechte Schreibweise  $[x_1, \dots, x_n]$  (oder  $(x_1, \dots, x_n)$ ) wird viel verwendet. Man spricht dann von *Zeilenvektoren*. Der gemeinsame Ausdruck für Zeilen- und Spaltenvektoren ist *n-Tupel*. Wir bevorzugen beim  $\mathbb{R}^n$  die senkrechte Anordnung, da sie sich später zwangloser in die Matrizenrechnung einordnet.

weggelassen, man schreibt also  $\mathbf{a} + \mathbf{b} + \mathbf{c}$ ,  $\mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d}$  usw.,  $\lambda\mu\mathbf{a}$ ,  $\lambda\mu\nu\mathbf{a}$  usw. Weitere Bräuche:

$$-\mathbf{x} := (-1)\mathbf{x}, \quad \mathbf{x}\lambda := \lambda\mathbf{x}, \quad \frac{\mathbf{x}}{\lambda} := \frac{1}{\lambda}\mathbf{x}, \quad (\lambda \neq 0), \quad \mathbf{0} := \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (2.2)$$

Dies alles hat niemand anders erwartet!

**Bemerkung:**

- (a) Die Vektoren des  $\mathbb{R}^n$  werden auch *Punkte* oder *Elemente* genannt, um sprachliche Eintönigkeit zu vermeiden.
- (b)  $\mathbb{R}^1$  und  $\mathbb{R}$  werden als gleich angesehen, d.h. man setzt einfach  $[x] = x \in \mathbb{R}$ .
- (c) Im  $\mathbb{R}^2$  und  $\mathbb{R}^3$  werden Vektoren gerne durch Pfeile über den Buchstaben symbolisiert.  $\vec{x}$  und  $\mathbf{x}$  bedeuten im  $\mathbb{R}^n$  mit  $n = 2$  oder  $n = 3$  also dasselbe.<sup>2</sup>

Aus schreibtechnischen Gründen notiert man Spaltenvektoren des  $\mathbb{R}^n$  auch in der Form

$$\mathbf{x} = [x_1, x_2, \dots, x_n]^T,$$

wobei <sup>T</sup> die Abkürzung für »transponiert« ist.

**Beispiel 2.1:**

Ein physikalisches Beispiel für einen höherdimensionalen Raum ist der sechsdimensionale Phasenraum  $\mathbb{R}^6$  in der kinetischen Gastheorie. Jedem Punkt (Gasmolekül) ordnet man dabei seine drei Ortskoordinaten und seine drei Impulskoordinaten zu, die zu einem 6-dimensionalen Vektor zusammengefasst werden (s. [72], S. 545ff).

## 2.1.2 Inneres Produkt, Beträge von Vektoren

**Definition 2.1:**

- (a) Das *innere Produkt* (*Skalarprodukt*) zweier Vektoren

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

aus  $\mathbb{R}^n$  ist definiert durch

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \dots + x_n y_n \quad (2.3)$$

<sup>2</sup> Die Kennzeichnung der Vektoren des  $\mathbb{R}^2$  und  $\mathbb{R}^3$  durch Pfeile kommt von der Veranschaulichung durch geometrische Pfeile  $\overrightarrow{AB}$  her. Ab Dimension 4 versagt die Anschauung aber. Hier sind Vektoren nur noch algebraische Objekte. Aus diesem Grunde wählen wir bei allgemeinen Erörterungen des  $\mathbb{R}^n$  die neutrale Schreibweise  $\mathbf{x}$ , also die Darstellung durch fett geschriebene Zeichen.

(b) Der *Betrag* von  $\mathbf{x}$ , auch *Länge* oder *euklidische Norm* genannt, ist

$$|\mathbf{x}| := \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}. \quad (2.4)$$

Abkürzung  $\mathbf{x}^2 = \mathbf{x} \cdot \mathbf{x}$ .

**Satz 2.1:**

(Regeln für das innere Produkt) Für alle Vektoren  $\mathbf{x}, \mathbf{y}, \mathbf{z}$  aus  $\mathbb{R}^n$  und alle  $\lambda \in \mathbb{R}$  gilt:

- |       |  |                             |
|-------|--|-----------------------------|
| (I)   | $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$  | <i>Kommutativgesetz</i>     |
| (II)  | $(\mathbf{x} + \mathbf{y}) \cdot \mathbf{z} = \mathbf{x} \cdot \mathbf{z} + \mathbf{y} \cdot \mathbf{z}$             | <i>Distributivgesetz</i>    |
| (III) | $\lambda(\mathbf{x} \cdot \mathbf{y}) = (\lambda\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (\lambda\mathbf{y})$ | <i>Assoziativgesetz</i>     |
| (IV)  | $\mathbf{x} \neq \mathbf{0} \Leftrightarrow \mathbf{x} \cdot \mathbf{x} > 0$   | <i>positive Definitheit</i> |

Ferner: Regeln für die Beträge von Vektoren:

- |        |  |                                |
|--------|--|--------------------------------|
| (V)    | $ \lambda\mathbf{x}  =  \lambda  \mathbf{x} $                  | <i>Homogenität</i>             |
| (VI)   | $ \mathbf{x} \cdot \mathbf{y}  \leq  \mathbf{x}  \mathbf{y} $  | <i>Schwarzsche Ungleichung</i> |
| (VII)  | $ \mathbf{x} + \mathbf{y}  \leq  \mathbf{x}  +  \mathbf{y} $   | <i>Dreiecksungleichung</i>     |
| (VIII) | $ \mathbf{x} - \mathbf{y}  \geq   \mathbf{x}  -  \mathbf{y}  $ | <i>2. Dreiecksungleichung</i>  |
| (IX)   | $ \mathbf{x}  = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$     | <i>Definitheit</i>             |

**Bemerkung:** Es handelt sich hier um die gleichen Regeln, wie wir sie aus  $\mathbb{R}^2$  und  $\mathbb{R}^3$  kennen.

**Beweis:**

Die Nachweise von (I) bis (V) und (IX) ergeben sich unmittelbar aus den Koordinatendarstellungen der Vektoren. Zum Beweis der *Schwarzschen Ungleichung* (VI) bemerken wir zunächst, dass sie im Falle  $\mathbf{y} = \mathbf{0}$  erfüllt ist. Im Falle  $\mathbf{y} \neq \mathbf{0}$  arbeitet man mit dem Hilfsvektor  $\mathbf{p} = ((\mathbf{x} \cdot \mathbf{y})/\mathbf{y}^2)\mathbf{y}$  (der Projektion von  $\mathbf{x}$  auf  $\mathbf{y}$ , wenn man an  $\mathbb{R}^2$  oder  $\mathbb{R}^3$  denkt). Damit ist

$$\begin{aligned} 0 \leq (\mathbf{x} - \mathbf{p})^2 &= \mathbf{x}^2 - 2\frac{(\mathbf{x} \cdot \mathbf{y})^2}{\mathbf{y}^2} + \frac{(\mathbf{x} \cdot \mathbf{y})^2}{\mathbf{y}^2} = \mathbf{x}^2 - \frac{(\mathbf{x} \cdot \mathbf{y})^2}{\mathbf{y}^2} \\ \Rightarrow 0 \leq \mathbf{x}^2 \mathbf{y}^2 - (\mathbf{x} \cdot \mathbf{y})^2 &\Rightarrow (\mathbf{x} \cdot \mathbf{y})^2 \leq \mathbf{x}^2 \mathbf{y}^2 \Rightarrow |\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}||\mathbf{y}|. \end{aligned}$$

Die *Dreiecksungleichung* ergibt sich nun leicht aus (VI):

$$\begin{aligned} |\mathbf{x} + \mathbf{y}|^2 &= (\mathbf{x} + \mathbf{y})^2 = |\mathbf{x}|^2 + 2\mathbf{x} \cdot \mathbf{y} + |\mathbf{y}|^2, \quad (\text{VI liefert:}) \\ &\leq |\mathbf{x}|^2 + 2|\mathbf{x}||\mathbf{y}| + |\mathbf{y}|^2 = (|\mathbf{x}| + |\mathbf{y}|)^2. \end{aligned}$$

Die 2. Dreiecksungleichung erhält man analog zu Folg. 1.3 in Abschn. 1.1.3. □

Mit dem Distributivgesetz (II) leitet man wieder die *binomischen Formeln* her:

$$\begin{aligned} (a + b)^2 &= a^2 + 2a \cdot b + b^2, \quad (a - b)^2 = a^2 - 2a \cdot b + b^2, \\ (a + b) \cdot (a - b) &= a^2 - b^2. \end{aligned} \quad (2.5)$$

**Bemerkung:** Das innere Produkt ist hier algebraisch eingeführt worden, da wir ja im unanschaulichen Raum  $\mathbb{R}^n$  zunächst keine geometrischen Begriffe wie Winkel, Geraden usw. haben. Diese können wir nun aber analog zum  $\mathbb{R}^2$  oder  $\mathbb{R}^3$  erklären, wobei wir das innere Produkt heranziehen. Dies geschieht in den folgenden Abschnitten.

### Übung 2.1:

Zeige:  $|\mathbf{x} + \mathbf{y}| = |\mathbf{x}| + |\mathbf{y}|$  gilt genau dann, wenn  $|\mathbf{x}|\mathbf{y} = |\mathbf{y}|\mathbf{x}$  gilt.

### 2.1.3 Unterräume, lineare Mannigfaltigkeiten

Die Begriffe *Linearkombination*, *lineare Abhängigkeit* und *Unabhängigkeit* werden wörtlich aus Abschn. 1.2.7 vom  $\mathbb{R}^3$  in den  $\mathbb{R}^n$  übernommen:

- (I)  $\sum_{i=1}^m \lambda_i \mathbf{a}_i$  (mit  $\lambda_i \in \mathbb{R}$ ) heißt eine *Linearkombination* der Vektoren  $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$ .
- (II)  $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$  sind *linear abhängig*, wenn wenigstens einer der Vektoren als Linearkombination der übrigen darstellbar ist, oder einer der Vektoren  $\mathbf{0}$  ist. Andernfalls heißen  $\mathbf{a}_1, \dots, \mathbf{a}_m$  *linear unabhängig*.
- (III) Es folgt:  $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$  sind genau dann linear unabhängig, wenn die Linearkombination

$$\sum_{i=1}^m \lambda_i \mathbf{a}_i = \mathbf{0} \tag{2.6}$$

nur mit  $\lambda_1 = \lambda_2 = \dots = \lambda_m = 0$  möglich ist. (s. Folg. 1.12, Abschn. 1.2.7)

#### Satz 2.2:

(*Fundamentallemma*) Je  $n + 1$  Vektoren des  $\mathbb{R}^n$  sind linear abhängig.

#### Beweis:

Beweis: Sind  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{n+1}$  beliebige  $n + 1$  Vektoren des  $\mathbb{R}^n$ , so betrachten wir die Gleichung  $\sum_{k=1}^{n+1} \lambda_k \mathbf{a}_k = \mathbf{0}$ . Koordinatenweise hingeschrieben, ist dies ein Gleichungssystem von  $n$  Gleichungen für die  $n + 1$  Unbekannten  $\lambda_1, \dots, \lambda_{n+1}$ . Wir denken uns die »Nullzeile«  $0 \cdot \lambda_1 + \dots + 0 \cdot \lambda_{n+1} = 0$  noch darunter geschrieben, um ebenso viele Gleichungen wie Unbekannte zu haben. Die Nullzeile ändert die Lösungsgesamtheit nicht. Mit dem Gaußschen Algorithmus<sup>3</sup> ergibt sich aber, dass dieses Gleichungssystem unendlich viele Lösungen hat. (Es entsteht kein Dreieckssystem und Folgerung 2.4(b) in Abschn. 2.2.3 gilt.) Damit gibt es Lösungen, bei denen die  $\lambda_k$  nicht alle 0 sind, d.h.: Die  $\mathbf{a}_1, \dots, \mathbf{a}_{n+1}$  sind linear abhängig.  $\square$

<sup>3</sup> Hier wird auf die Abschnitte 2.2.1 und 2.2.3 vorgegriffen. Dabei entsteht kein logischer Zirkel, da die dortige Erläuterung des Gaußschen Algorithmus (bis Folg. 2.4) unabhängig vom vorliegenden und nächsten Abschnitt verstanden werden kann.

Da die Koordinateneinheitsvektoren  $e_1, \dots, e_n$  ( $e_i := [0, \dots, 0, 1, 0, \dots, 0]^T$ , 1 an  $i$ -ter Stelle, sonst Nullen) sicherlich linear unabhängig sind, ist  $n$  die maximale Zahl linear unabhängiger Vektoren des  $\mathbb{R}^n$ . Dies motiviert zu folgendem Dimensionsbegriff:

**Definition 2.2:**

- (I) Eine Teilmenge  $U$  von  $\mathbb{R}^n$  heißt ein *Unterraum* oder *Teilraum* von  $\mathbb{R}^n$ , wenn mit je zwei Vektoren  $a, b$  aus  $U$  auch  $a + b$  und  $\lambda a$  (für alle  $\lambda \in \mathbb{R}$ ) in  $U$  liegen.
- (II) Die maximale Anzahl linear unabhängiger Vektoren aus  $U$  heißt die *Dimension* von  $U$ .
- (III) Es sei  $m$  die Dimension von  $U$ . Darin wird jedes  $m$ -Tupel  $(a_1, a_2, \dots, a_m)$  von linear unabhängigen Vektoren aus  $U$  eine *Basis* von  $U$  genannt. Man schreibt symbolisch:

$$\dim U = m.$$

**Bemerkung.** Der kleinste Unterraum von  $\mathbb{R}^n$  ist  $U = \{0\}$ . Er hat die Dimension 0. Der größte Unterraum von  $\mathbb{R}^n$  ist  $\mathbb{R}^n$  selbst, mit der Dimension  $n$ .

Allgemein lässt sich ein Unterraum von  $\mathbb{R}^n$  so beschreiben:

**Satz 2.3:**

- (a) Ist  $(a_1, \dots, a_m)$  eine Basis des  $m$ -dimensionalen Unterraumes  $U$  von  $\mathbb{R}^n$ , so ist  $U$  die Menge aller Linearkombinationen

$$x = \lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_m a_m \quad (\lambda_i \in \mathbb{R}). \quad (2.7)$$

- (b) Die reellen Zahlen  $\lambda_1, \dots, \lambda_m$  sind dabei eindeutig durch  $x$  und die Basis  $(a_1, \dots, a_m)$  bestimmt.

**Beweis:**

- (a) Ist  $x$  ein beliebiger Vektor aus  $U$ , so sind  $x, a_1, \dots, a_m$  linear abhängig, da es nicht mehr als  $m$  linear unabhängige Vektoren in  $U$  gibt (wegen  $\dim U = m$ ). Also gibt es eine Linearkombination

$$\mu_0 x + \mu_1 a_1 + \dots + \mu_m a_m = 0, \quad (\mu_i \in \mathbb{R}). \quad (2.8)$$

in der nicht alle  $\mu_i$  Nullen sind. Dann ist aber  $\mu_0 \neq 0$ , sonst wären die  $(a_1, \dots, a_m)$  linear abhängig. Division der Gl. (2.8) durch  $\mu_0$  und  $\lambda_i := -\mu_i/\mu_0$  liefern Gleichung (2.7).

Umgekehrt muss jedes  $x$  der Form (2.7) zu  $U$  gehören, da nach Definition des Unterraumes folgendes gilt:  $\lambda_i a_i \in U$  für alle  $i$ ,  $\lambda_1 a_1 + \lambda_2 a_2 \in U$ ,  $(\lambda_1 a_1 + \lambda_2 a_2) + \lambda_3 a_3 \in U$  usw.

- (b) Zum Nachweis der Eindeutigkeit der  $\lambda_i$  nehmen wir an, dass es neben (2.7) eine weitere Linearkombination  $x = \sum_{i=1}^m \mu_i a_i$  gibt. Subtraktion von (2.7) liefert  $0 = \sum_{i=1}^m (\lambda_i - \mu_i) a_i$ . Da die  $a_1, \dots, a_m$  linear unabhängig sind, folgt  $\lambda_i - \mu_i = 0$ , also  $\lambda_i = \mu_i$  für alle  $i$ , d.h. die  $\lambda_i$  sind eindeutig bestimmt.  $\square$

Umgekehrt regt (2.7) zur *Konstruktion* von Unterräumen an: Ist  $A$  eine beliebige nichtleere Teilmenge des  $\mathbb{R}^n$ , so bildet die Menge aller Linearkombinationen  $\mathbf{x} = \sum_{i=1}^p \lambda_i \mathbf{a}_i$  mit beliebigen  $\lambda_i \in \mathbb{R}$ ,  $\mathbf{a}_i \in A$  und  $p \in \mathbb{N}$  offenbar einen Unterraum  $U$  von  $\mathbb{R}^n$ . Er wird durch

$$U =: \text{Span } A$$

symbolisiert. Man sagt, » $A$  spannt den Unterraum  $U$  auf«, oder: » $A$  ist ein *Erzeugendensystem* von  $U$ «. Oft ist  $A$  dabei eine endliche Menge  $\{\mathbf{a}_1, \dots, \mathbf{a}_m\}$ . Der von  $\{\mathbf{a}_1, \dots, \mathbf{a}_m\}$  *aufgespannte Unterraum*

$$U = \text{Span}\{\mathbf{a}_1, \dots, \mathbf{a}_m\}$$

besteht also aus allen Linearkombinationen  $\mathbf{x} = \sum_{i=1}^m \lambda_i \mathbf{a}_i$ . Sind die  $\mathbf{a}_1, \dots, \mathbf{a}_m$  hierbei linear unabhängig, so bilden sie eine Basis von  $U$ , wie wir im Folgenden beweisen:

**Satz 2.4:**

(*Konstruktion von Unterräumen im  $\mathbb{R}^n$* ) Sind die Vektoren  $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$  linear unabhängig, so bilden die Vektoren der Form

$$\mathbf{x} = \lambda_1 \mathbf{a}_1 + \dots + \lambda_m \mathbf{a}_m \quad (\lambda_i \in \mathbb{R}) \quad (2.9)$$

einen  $m$ -dimensionalen Unterraum von  $\mathbb{R}^n$ . Die Menge  $\{\mathbf{a}_1, \dots, \mathbf{a}_m\}$  ist also eine Basis des Unterraumes.

**Beweis:**

- (I) Die Vektoren der Form (2.9) bilden einen Unterraum  $U$ , wie oben erläutert. Seine Dimension ist mindestens  $m$ , da er ja die  $m$  Vektoren  $(\mathbf{a}_1, \dots, \mathbf{a}_m)$  enthält. Frage: Gibt es mehr als  $m$  linear unabhängige Vektoren in  $U$ ? Wir zeigen, dass dies nicht der Fall ist, d.h. es wird folgendes bewiesen:
- (II) Je  $m + 1$  Vektoren aus  $U$  sind linear abhängig. Zum Beweis wählt man  $m + 1$  beliebige Vektoren  $\mathbf{b}_1, \dots, \mathbf{b}_{m+1} \in U$  aus. Sie lassen sich in folgender Form darstellen:

$$\mathbf{b}_i = \sum_{k=1}^m \gamma_{ik} \mathbf{a}_k \quad (\gamma_{ik} \in \mathbb{R}), \quad i = 1, \dots, m + 1.$$

Man betrachtet nun die Gleichung

$$\sum_{i=1}^{m+1} \lambda_i \mathbf{b}_i = \mathbf{0}. \quad (2.10)$$

Sie wird umgeformt in

$$\mathbf{0} = \sum_{i=1}^{m+1} \lambda_i \mathbf{b}_i = \sum_{i=1}^{m+1} \lambda_i \sum_{k=1}^m \gamma_{ik} \mathbf{a}_k = \sum_{k=1}^m \mathbf{a}_k \sum_{i=1}^{m+1} \lambda_i \gamma_{ik}. \quad (2.11)$$

Da die  $\mathbf{a}_k$  linear unabhängig sind, muss

$$\sum_{i=1}^{m+1} \lambda_i \gamma_{ik} = 0 \quad \text{für alle } k = 1, \dots, m$$

gelten. Wir fassen die  $\gamma_{ik}$  zu Vektoren  $\mathbf{c}_i = [\gamma_{i1}, \dots, \gamma_{im}]^T$  zusammen. Die letzte Gleichung wird damit zu

$$\sum_{i=1}^{m+1} \lambda_i \mathbf{c}_i = \mathbf{0}, \quad \mathbf{c}_i \in \mathbb{R}^m. \quad (2.12)$$

Da aber je  $m + 1$  Vektoren des  $\mathbb{R}^m$  linear abhängig sind (Satz 2.2), gibt es  $\lambda_i$ , die nicht alle Null sind, und die (2.12) erfüllen. Wegen (2.11) ist damit auch (2.10) mit diesen  $\lambda_i$  richtig, d.h. die  $\mathbf{b}_i$  sind linear abhängig.  $\square$

### Satz 2.5:

(Austausch von Basiselementen) Es sei  $U$  ein  $m$ -dimensionaler Unterraum von  $\mathbb{R}^n$  mit der Basis  $(\mathbf{a}_1, \dots, \mathbf{a}_m)$ . Sind  $\mathbf{b}_1, \dots, \mathbf{b}_k$  (mit  $k < m$ ) beliebige linear unabhängige Vektoren aus  $U$ , so kann man sie zu einer Basis  $(\mathbf{b}_1, \dots, \mathbf{b}_k, \mathbf{b}_{k+1}, \dots, \mathbf{b}_m)$  von  $U$  ergänzen, wobei die  $\mathbf{b}_{k+1}, \dots, \mathbf{b}_m$  aus  $\{\mathbf{a}_1, \dots, \mathbf{a}_m\}$  entnommen sind.

### Beweis:

Es gibt ein  $\mathbf{a}_i$ , das nicht Linearkombination der  $\mathbf{b}_1, \dots, \mathbf{b}_k$  ist. Wären nämlich alle  $\mathbf{a}_1, \dots, \mathbf{a}_m$  Linearkombinationen der  $\mathbf{b}_1, \dots, \mathbf{b}_k$ , so wären damit alle  $\mathbf{x} = \sum_{i=1}^m \lambda_i \mathbf{a}_i \in U$  auch Linearkombination der  $\mathbf{b}_1, \dots, \mathbf{b}_k$ , (nach Ersetzen der  $\mathbf{a}_i$  durch Linearkombinationen der  $\mathbf{b}_1, \dots, \mathbf{b}_k$ ). Damit hätte  $U$  eine Dimension  $\leq k$  (nach Satz 2.4). Wegen  $k < m$  kann dies nicht sein. Also ist wenigstens ein  $\mathbf{a}_i$  keine Linearkombination der  $\mathbf{b}_1, \dots, \mathbf{b}_k$ . Setze  $\mathbf{b}_{k+1} := \mathbf{a}_i$  und wende den gleichen Schluss auf  $\mathbf{b}_1, \dots, \mathbf{b}_{k+1}$  an, (falls  $k + 1 < m$ ). So fortfahrend erhält man die neue Basis  $(\mathbf{b}_1, \dots, \mathbf{b}_m)$  von  $U$ .  $\square$

Geraden und Ebenen im  $\mathbb{R}^2$  und  $\mathbb{R}^3$  hatten wir durch Parameterdarstellungen beschrieben. Ihre Verallgemeinerungen auf den  $\mathbb{R}^n$  heißen »lineare Mannigfaltigkeiten«.

### Definition 2.3:

Es seien  $\mathbf{r}_0, \mathbf{a}_1, \dots, \mathbf{a}_m$  ( $m \leq n$ ) Vektoren des  $\mathbb{R}^n$ , wobei  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m$  linear unabhängig seien. Dann heißt die Menge aller Vektoren

$$\mathbf{x} = \mathbf{r}_0 + \lambda_1 \mathbf{a}_1 + \lambda_2 \mathbf{a}_2 + \dots + \lambda_m \mathbf{a}_m \quad (\lambda_i \in \mathbb{R}) \quad (2.13)$$

eine  $m$ -dimensionale lineare Mannigfaltigkeit. Im Falle  $m = 1$ , also  $\mathbf{x} = \mathbf{r}_0 + \lambda_1 \mathbf{a}_1$  ( $\lambda_1 \in \mathbb{R}$ ) nennt man sie eine Gerade, im Falle  $m = n - 1$  eine Hyperebene im  $\mathbb{R}^n$ .

Gleichung (2.13) wird die *Parameterdarstellung* der linearen Mannigfaltigkeit genannt. Die Vektoren  $\sum_{k=1}^m \lambda_k \mathbf{a}_k$ , mit den  $\mathbf{a}_k$  aus (2.13), bilden einen  $m$ -dimensionalen Unterraum  $U$  (siehe Satz 2.4). Aus diesem Grunde symbolisiert man die Mannigfaltigkeit  $M$  in Def. 2.3 auch durch

$$M = \mathbf{r}_0 + U. \quad (2.14)$$

### Folgerung 2.1:

Ist  $M$  eine  $m$ -dimensionale lineare Mannigfaltigkeit in  $\mathbb{R}^n$  und  $\mathbf{x}_0$  irgendein Vektor aus  $M$ , so ist

$$M - \mathbf{x}_0 := \{\mathbf{x} - \mathbf{x}_0 \mid \mathbf{x} \in M\} \quad (2.15)$$

ein  $m$ -dimensionaler Unterraum von  $\mathbb{R}^n$ .

Eine *Hyperebene*  $H$ , gegeben durch  $\mathbf{x} = \mathbf{r}_0 + \sum_{i=1}^{n-1} \lambda_i \mathbf{a}_i$ , (s. (2.13)) kann auch in der *Hesseschen Normalform*  $\mathbf{x} \cdot \mathbf{n} = \rho$  ( $\rho \geq 0$ ) beschrieben werden (analog zu  $\mathbb{R}^2$  und  $\mathbb{R}^3$ ).  $\mathbf{n} \neq \mathbf{0}$  wird dabei aus dem Gleichungssystem  $\mathbf{a}_i \cdot \mathbf{n} = 0$  ( $i = 1, \dots, n-1$ ) berechnet (eindeutig bis auf einen reellen Faktor  $\neq 0$ , s. Gaußscher Algorithmus für rechteckige Systeme, Abschn. 2.2.5). Durch  $|\mathbf{n}| = 1$  und  $\mathbf{n} \cdot \mathbf{r}_0 > 0$  ist  $\mathbf{n}$  und  $\rho = \mathbf{n} \cdot \mathbf{r}_0$  eindeutig bestimmt (falls  $\rho > 0$ ).

### Übung 2.2\*

Es sei eine Hyperebene im  $\mathbb{R}^4$  durch folgende Parameterdarstellung gegeben

$$\mathbf{x} = \sum_{i=1}^3 \lambda_i \mathbf{a}_i, \quad \text{mit} \quad \mathbf{a}_1 = \begin{bmatrix} 1 \\ 8 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} -1 \\ 2 \\ 7 \\ -2 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 10 \\ -1 \\ 3 \\ 4 \end{bmatrix}.$$

Gib die Hessesche Normalform dazu an. *Hinweis:* Löse zunächst das Gleichungssystem  $\mathbf{a}_1 \cdot \mathbf{n}' = 0$ ,  $\mathbf{a}_2 \cdot \mathbf{n}' = 0$ ,  $\mathbf{a}_3 \cdot \mathbf{n}' = 0$  für einen (unbekannten) Vektor  $\mathbf{n}'$ , dessen letzte Komponente 1 gesetzt wird. Denke an die Cramersche Regel in 1.2.7. Berechne dann  $\mathbf{n}$  und  $\rho$  aus  $\mathbf{n}'$ .

## 2.1.4 Geometrie im $\mathbb{R}^n$ , Winkel, Orthogonalität

### Winkel

#### Definition 2.4:

Der Winkel  $\varphi$  zwischen zwei Vektoren  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$  mit  $\mathbf{a} \neq \mathbf{0}, \mathbf{b} \neq \mathbf{0}$ , ist

$$\varphi = \angle(\mathbf{a}, \mathbf{b}) := \arccos \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}| \cdot |\mathbf{b}|}. \quad (2.16)$$

Ist  $\mathbf{a} = \mathbf{0}$  oder  $\mathbf{b} = \mathbf{0}$ , so kann  $\angle(\mathbf{a}, \mathbf{b})$  jede Zahl aus  $[0, \pi]$  sein.



Man sagt,  $\mathbf{a}$  und  $\mathbf{b}$  aus  $\mathbb{R}^n$  stehen *rechtwinklig* (*senkrecht, orthogonal*) aufeinander — in Zeichen  $\mathbf{a} \perp \mathbf{b}$  — wenn  $\mathbf{a} \cdot \mathbf{b} = 0$  ist (also  $\sphericalangle(\mathbf{a}, \mathbf{b}) = \pi/2$  im Falle  $\mathbf{a} \neq \mathbf{0}$ ,  $\mathbf{b} \neq \mathbf{0}$ ).

Man gewinnt daraus unmittelbar:

**Folgerung 2.2:**

- (a) *Pythagoras im  $\mathbb{R}^n$* . Zwei Vektoren  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$  stehen genau dann rechtwinklig aufeinander, wenn Folgendes erfüllt ist:

$$(\mathbf{a} + \mathbf{b})^2 = \mathbf{a}^2 + \mathbf{b}^2 \quad (\text{wie auch } (\mathbf{a} - \mathbf{b})^2 = \mathbf{a}^2 + \mathbf{b}^2). \quad (2.17)$$

Für alle  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$  gelten die folgenden Gleichungen:

- (b) *Cosinussatz:*

$$(\mathbf{a} - \mathbf{b})^2 = \mathbf{a}^2 + \mathbf{b}^2 - 2|\mathbf{a}||\mathbf{b}| \cos \sphericalangle(\mathbf{a}, \mathbf{b}). \quad (2.18)$$

- (c) *Parallelogrammgleichung:*

$$(\mathbf{a} + \mathbf{b})^2 + (\mathbf{a} - \mathbf{b})^2 = 2(\mathbf{a}^2 + \mathbf{b}^2). \quad (2.19)$$

## Orthonormalbasis

**Definition 2.5:**

- (a) Sind  $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$  Vektoren der Länge 1, die paarweise rechtwinklig aufeinander stehen, so nennt man das  $m$ -Tupel  $(\mathbf{a}_1, \dots, \mathbf{a}_m)$  dieser Vektoren ein Orthonormalsystem. Es gilt also dabei:

$$\mathbf{a}_i \cdot \mathbf{a}_k = \delta_{ik} \quad \text{für alle } i, k \in \{1, \dots, m\} \quad (2.20)$$

$$\text{mit dem Kroneckersymbol } \delta_{ik} := \begin{cases} 1, & \text{falls } i = k \\ 0, & \text{falls } i \neq k. \end{cases} \quad (2.21)$$

- (b) Eine Orthonormalbasis von  $\mathbb{R}^n$ , oder eines Unterraumes von  $\mathbb{R}^n$ , ist eine Basis, die gleichzeitig ein Orthonormalsystem ist.

**Bemerkung:**

- (a) Die Koordinateneinheitsvektoren  $\mathbf{e}_i \in \mathbb{R}^n$  bilden eine Orthonormalbasis von  $\mathbb{R}^n$ .
- (b) Die Vektoren  $\mathbf{a}_i$  eines Orthonormalsystems  $(\mathbf{a}_1, \dots, \mathbf{a}_m)$  sind linear unabhängig, denn aus  $\sum_{i=1}^m \lambda_i \mathbf{a}_i = \mathbf{0}$  folgt nach Multiplikation mit  $\mathbf{a}_k$  sofort:  $\lambda_k \mathbf{a}_k \cdot \mathbf{a}_k = 0$ , also  $\lambda_k = 0$  (wegen  $\mathbf{a}_k \cdot \mathbf{a}_k = |\mathbf{a}_k|^2 = 1$ ), für alle  $k = 1, \dots, m$ .

Ist nun  $(\mathbf{b}_1, \dots, \mathbf{b}_m)$  eine beliebige Basis eines  $m$ -dimensionalen Unterraumes  $U$  von  $\mathbb{R}^n$  (der auch gleich  $\mathbb{R}^n$  sein kann), so können wir sie in eine *Orthonormalbasis* von  $U$  verwandeln. Das gelingt (z.B.) mit dem folgenden *Orthogonalisierungsverfahren von Erhard Schmidt*<sup>4</sup>:

(I) Man setzt

$$\mathbf{a}_1 := \mathbf{b}_1 / |\mathbf{b}_1|. \quad (2.22)$$

(II) Für  $k = 2, 3, \dots, m$  bildet man nacheinander

$$\mathbf{d}_k := \mathbf{b}_k - \sum_{i=1}^{k-1} (\mathbf{b}_k \cdot \mathbf{a}_i) \mathbf{a}_i \quad (2.23)$$

$$\text{und } \mathbf{a}_k := \mathbf{d}_k / |\mathbf{d}_k|.$$

Damit ist für zwei beliebige verschiedene  $\mathbf{a}_j, \mathbf{a}_k$ , wobei wir ohne Beschränkung der Allgemeinheit  $j < k$  annehmen, nach (2.23):

$$\mathbf{a}_k \cdot \mathbf{a}_j = \frac{1}{|\mathbf{d}_k|} (\mathbf{d}_k \cdot \mathbf{a}_j) = \frac{1}{|\mathbf{d}_k|} (\mathbf{b}_k \cdot \mathbf{a}_j - (\mathbf{b}_k \cdot \mathbf{a}_j) 1) = 0.$$

Ferner ist offensichtlich  $|\mathbf{a}_k| = 1$  für alle  $k$ . Damit bilden die  $\mathbf{a}_1, \dots, \mathbf{a}_m$  eine Orthonormalbasis von  $U$ . — Wir haben somit gezeigt:

**Satz 2.6:**

Jeder Unterraum von  $\mathbb{R}^n$  besitzt eine Orthonormalbasis.

Ferner gilt:

**Satz 2.7:**

Ist  $(\mathbf{a}_1, \dots, \mathbf{a}_m)$  ein Orthonormalsystem und  $\mathbf{x}$  eine Linearkombination daraus:

$$\mathbf{x} = \sum_{i=1}^m \lambda_i \mathbf{a}_i,$$

so lassen sich die Komponenten  $\lambda_i$  auf folgende Weise leicht berechnen:

$$\lambda_i = \mathbf{x} \cdot \mathbf{a}_i \text{ für alle } i = 1, \dots, m. \quad (2.24)$$

**Beweis:**

$\mathbf{x} \cdot \mathbf{a}_k = \sum_{i=1}^m \lambda_i \mathbf{a}_i \cdot \mathbf{a}_k = \lambda_k \mathbf{a}_k \cdot \mathbf{a}_k = \lambda_k$ , also  $\mathbf{x} \cdot \mathbf{a}_k = \lambda_k$ . Ersetzt man hier  $k$  durch  $i$ , so folgt (2.24). □

<sup>4</sup> Erhard Schmidt (1876–1959), deutscher Mathematiker

<sup>5</sup> Es ist  $\mathbf{d}_k \neq \mathbf{0}$ . Denn  $\mathbf{d}_k$  ist eine Linearkombination der  $\mathbf{b}_1, \dots, \mathbf{b}_k$  (durch Induktion leicht nachweisbar). Der Koeffizient von  $\mathbf{b}_k$  ist dabei 1, s. (2.23). Wäre  $\mathbf{d}_k = \mathbf{0}$ , so müsste er aber 0 sein, da  $\mathbf{b}_1, \dots, \mathbf{b}_k$  linear unabhängig sind. Also ist  $\mathbf{d}_k \neq \mathbf{0}$ .

## Orthogonales Komplement

### Definition 2.6:

- (a) Ist  $U$  ein Unterraum von  $\mathbb{R}^n$ , so heißt die Menge der Vektoren  $\mathbf{x} \in \mathbb{R}^n$ , die auf jedem Vektor von  $U$  rechtwinklig stehen, das orthogonale Komplement  $U^\perp$  von  $U$ . In Formeln:

$$U^\perp := \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \cdot \mathbf{u} = 0 \text{ für alle } \mathbf{u} \in U\}.$$

- (b) Allgemeiner: Ist  $V$  ein Unterraum von  $\mathbb{R}^n$ , der den Unterraum  $U$  umfasst:  $U \subset V$ , so heißt

$$U_V^\perp := \{\mathbf{x} \in V \mid \mathbf{x} \cdot \mathbf{u} = 0 \text{ für alle } \mathbf{u} \in U\}$$

das *orthogonale Komplement* von  $U$  bzgl.  $V$ .  $U^\perp$  und  $U_V^\perp$  sind offenbar Unterräume von  $\mathbb{R}^n$ .

### Folgerung 2.3:

Es sei  $U \subset V$ , wobei  $U, V$  Unterräume von  $\mathbb{R}^n$  sind. Damit folgt

- (a)  $\dim U + \dim U_V^\perp = \dim V$ .  
 (b) Jedes  $\mathbf{v} \in V$  lässt sich eindeutig als Summe  $\mathbf{v} = \mathbf{u} + \mathbf{u}^*$  mit  $\mathbf{u} \in U, \mathbf{u}^* \in U_V^\perp$  darstellen.  
 Im Fall  $V = \mathbb{R}^n$  folgt:  $\dim U + \dim U^\perp = n$ .

### Beweis:

- (a) Man wähle eine Orthonormalbasis  $(\mathbf{b}_1, \dots, \mathbf{b}_m)$  in  $U$  (nach Satz 2.6 möglich), erweitere sie zu einer Basis in  $V$  (nach Satz 2.5 in Abschn. 2.1.3) und verwandle sie mit dem Schmidtschen Orthogonalisierungsverfahren in eine Orthonormalbasis  $(\mathbf{b}_1, \dots, \mathbf{b}_m, \mathbf{b}_{m+1}, \dots, \mathbf{b}_p)$  in  $V$ . Dann spannen  $\mathbf{b}_{m+1}, \dots, \mathbf{b}_p$  offenbar den Raum  $U_V^\perp$  auf, und es gilt (a).  
 (b)  $\mathbf{v} = \underbrace{\lambda_1 \mathbf{b}_1 + \dots + \lambda_m \mathbf{b}_m}_{\mathbf{u}} + \underbrace{\lambda_{m+1} \mathbf{b}_{m+1} + \dots + \lambda_p \mathbf{b}_p}_{\mathbf{u}^*}.$  □

**Bemerkung:** Dies trockene Zeug erweist sich bei linearen Gleichungssystemen und Eigenwertproblemen später als nützlich.

### 2.1.5 Der Raum $\mathbb{C}^n$

Analog zum Raum  $\mathbb{R}^n$  wird der Raum  $\mathbb{C}^n$  gebildet. Er besteht aus allen *Spaltenvektoren* (*n-Tupeln*)

$$\mathbf{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \quad (z_i \in \mathbb{C}) \quad ^6$$

mit *komplexen Zahlen*  $z_i$ , *Koordinaten* genannt. Addition und Multiplikation mit Skalaren  $\lambda \in \mathbb{C}$  sind, wie im  $\mathbb{R}^n$ , koordinatenweise definiert, womit auch alle Gesetze über Addition und Multiplikation mit Skalaren unverändert gelten.

Das *innere Produkt* aus  $\mathbf{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$ ,  $\mathbf{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} \in \mathbb{C}^n$  ist so erklärt

$$\mathbf{z} \cdot \mathbf{w} := \sum_{i=1}^n z_i \overline{w_i},$$

wobei  $\overline{w_i}$  die zu  $w_i$  konjugiert komplexe Zahl ist. Damit gilt

$$\mathbf{z} \cdot \mathbf{w} = \overline{\mathbf{w} \cdot \mathbf{z}}$$

für alle  $\mathbf{z}, \mathbf{w} \in \mathbb{C}^n$ . Alle anderen Gesetze des inneren Produktes gelten unverändert (s. Satz 2.1 (II), (III), (IV)). Mit der Definition

$$|\mathbf{z}| = \sqrt{\mathbf{z} \cdot \mathbf{z}} = \sqrt{\sum_{i=1}^n |z_i|^2}$$

des Betrages von  $\mathbf{z}$  sind überdies die Regeln (V) bis (IX) in Satz 2.1 auch im  $\mathbb{C}^n$  erfüllt (Beweis analog zum  $\mathbb{R}^n$ ).

Wir kommen auf den Raum  $\mathbb{C}^n$  in Abschn. 2.4.2 kurz zurück. Wesentlich benötigen wir diesen Raum aber in der Eigenwerttheorie von Matrizen in Abschn. 3.

## 2.2 Lineare Gleichungssysteme, Gaußscher Algorithmus

Ein Gleichungssystem der Form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned} \tag{2.25}$$

( $a_{ik}, b_i \in \mathbb{R}$  gegeben,  $x_k \in \mathbb{R}$  gesucht) heißt ein (reelles) *lineares Gleichungssystem von  $m$  Gleichungen mit  $n$  Unbekannten*. Sind alle  $b_i = 0$  ( $i = 1, \dots, m$ ), so liegt ein *homogenes* lineares Gleichungssystem vor, andernfalls ein *inhomogenes*.

Als *Lösung* des Gleichungssystems (2.25) bezeichnet man jeden Vektor  $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$  aus  $\mathbb{R}^n$ , dessen Koordinaten  $x_1, \dots, x_n$  alle Gleichungen in (2.25) erfüllen.

6  $\mathbb{C}$  = Menge der komplexen Zahlen, s. Bd. Burg/Haf/Wille (Analysis) [27]

Wir beschäftigen uns zunächst mit dem Fall  $m = n$ , der für die Praxis am wichtigsten ist. Man spricht hier von *quadratischen* linearen Gleichungssystemen.<sup>7</sup>

### 2.2.1 Lösung quadratischer Gleichungssysteme

Zur Lösung eines quadratischen *linearen Gleichungssystems*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ \vdots & \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= b_n \end{aligned} \quad (2.26)$$

von  $n$  Gleichungen mit  $n$  Unbekannten ( $a_{ik}, b_i \in \mathbb{R}$  gegeben,  $x_k \in \mathbb{R}$  gesucht) kann der *Gaußsche Algorithmus*<sup>8</sup>, auch *Subtraktions-* oder *Eliminationsverfahren* genannt, verwendet werden.

Der *Grundgedanke* ist einfach: Man multipliziert die Gleichungen mit konstanten Faktoren und subtrahiert sie dann so voneinander, dass möglichst viele Unbekannte  $x_k$  dabei verschwinden, um, bei Fortführung dieses Prozesses, zu einer Gleichung mit nur einer Unbekannten zu gelangen. Diese Unbekannte wird berechnet und ihr Wert in die übrigen Gleichungen eingesetzt. Sukzessive ermittelt man aus diesen Gleichungen dann die übrigen Unbekannten. — Wir beschreiben diesen Prozess genauer:

**Gaußscher Algorithmus:** Zunächst reduziert man das System (2.26) zu einem Gleichungssystem von  $n - 1$  Gleichungen mit  $n - 1$  Unbekannten, und zwar auf folgende Weise:

**Erster Reduktionsschritt:** Wir gehen davon aus, dass  $a_{11} \neq 0$  ist.

Multipliziert man nun die erste Gleichung in (2.26) rechts und links mit dem Faktor  $c_{21} := a_{21}/a_{11}$  und subtrahiert die Seiten der so entstandenen Gleichung von den entsprechenden Seiten der zweiten Gleichung, so entsteht eine Gleichung, in der  $x_1$  nicht mehr vorkommt. Auf die gleiche Weise verfährt man mit allen weiteren Gleichungen: Man multipliziert also als nächstes die erste Gleichung mit  $c_{31} := a_{31}/a_{11}$  und subtrahiert sie von der dritten Gleichung; wobei wiederum  $x_1$  herausfällt. Anschließend multipliziert man die erste Gleichung mit  $c_{41} := a_{41}/a_{11}$ , subtrahiert sie von der vierten Gleichung usw. Führt man diesen Prozess bis zur  $n$ -ten Gleichung durch, so entsteht ein Gleichungssystem

$$\begin{aligned} a_{22}^{(2)}x_2 + a_{23}^{(2)}x_3 + \dots + a_{2n}^{(2)}x_n &= b_2^{(2)} \\ a_{32}^{(2)}x_2 + a_{33}^{(2)}x_3 + \dots + a_{3n}^{(2)}x_n &= b_3^{(2)} \\ \vdots & \\ a_{n2}^{(2)}x_2 + a_{n3}^{(2)}x_3 + \dots + a_{nn}^{(2)}x_n &= b_n^{(2)} \end{aligned} \quad (2.27)$$

<sup>7</sup> Wer vorrangig an der *praktischen Lösungsberechnung* interessiert ist, findet alles Notwendige dazu in den Abschn. 2.2.1 bis 2.2.3 (bis Beisp. 2.3). Zum Verständnis wird nur der Begriff des Vektors aus  $\mathbb{R}^n$  vorausgesetzt, s. Abschn. 2.1.1.

<sup>8</sup> Für zwei- und dreireihige Systeme ( $n = 2$  oder  $3$ ) ist die direkte Auflösung mit der Cramerschen Regel ebenfalls eine gute Methode ( $n = 2$ : Abschn. 1.1.7, (1.69),  $n = 3$ : Abschn. 1.2.7, (1.99),  $n \in \mathbb{N}$  beliebig: Abschn. 3.4.6). Für  $n \geq 4$  ist aber in der Praxis der Gaußsche Algorithmus respektive ein iteratives Verfahren vorzuziehen.

dessen Koeffizienten sich auf folgende Art ergeben:

$$a_{ik}^{(2)} = a_{ik} - c_{i1}a_{1k}, \quad b_i^{(2)} = b_i - c_{i1}b_1, \quad \text{mit } c_{i1} = \frac{a_{i1}}{a_{11}} \text{ und } i, k = 2, 3, \dots, n.$$

Jede Lösung  $[x_1, \dots, x_n]^T$  von (2.26) liefert offenbar eine Lösung  $[x_2, \dots, x_n]^T$  von (2.27), während man umgekehrt jede Lösung  $[x_2, \dots, x_n]^T$  von (2.27) zu einer Lösung von (2.26) erweitern kann, wenn man die  $x_2, \dots, x_n$  in die erste Gleichung von (2.26) einsetzt und daraus  $x_1$  berechnet.

Den Reduktionsschritt wendet man nun auf das System (2.27) abermals an. Das dann entstandene System reduziert man abermals usw.

Allgemein geht man folgendermaßen vor: *p-ter Reduktionsschritt*. Es sei

$$\begin{aligned} a_{pp}^{(p)}x_p + \dots + a_{pn}^{(p)}x_n &= b_p^{(p)} \\ \vdots & \quad \quad \quad \vdots \\ a_{np}^{(p)}x_p + \dots + a_{nn}^{(p)}x_n &= b_n^{(p)}, \quad p < n \end{aligned} \tag{2.28}$$

eins der reduzierten Gleichungssysteme. Unter der Voraussetzung, dass  $a_{pp}^{(p)} \neq 0$  gilt, bildet man (nach dem Muster von (2.27)) das neue Gleichungssystem

$$\begin{aligned} a_{p+1,p+1}^{(p+1)}x_{p+1} + \dots + a_{p+1,n}^{(p+1)}x_n &= b_{p+1}^{(p+1)} \\ \vdots & \quad \quad \quad \vdots \\ a_{n,p+1}^{(p+1)}x_{p+1} + \dots + a_{nn}^{(p+1)}x_n &= b_n^{(p+1)} \end{aligned} \tag{2.29}$$

mit  $a_{ik}^{(p+1)} = a_{ik}^{(p)} - c_{ip}a_{pk}^{(p)}$ ,  $b_i^{(p+1)} = b_i^{(p)} - c_{ip}b_p^{(p)}$ , wobei  $c_{ip} = \frac{a_{ip}^{(p)}}{a_{pp}^{(p)}}$ ,  $i, k = p+1, \dots, n$ .

Führt man dies für alle  $p = 2, 3, \dots, n-1$  durch — vorausgesetzt, dass bei jedem der reduzierten Gleichungssysteme  $a_{pp}^{(p)} \neq 0$  ist — und schreibt die ersten Gleichungen aller betrachteten Gleichungssysteme untereinander, so erhält man folgendes *Dreieckssystem*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n &= b_1 \\ &+ a_{22}^{(2)}x_2 + a_{23}^{(2)}x_3 + \dots + a_{2n}^{(2)}x_n = b_2^{(2)} \\ &\quad a_{33}^{(3)}x_3 + \dots + a_{3n}^{(3)}x_n = b_3^{(3)} \\ &\quad \quad \vdots \\ &\quad \quad \quad a_{nn}^{(n)}x_n = b_n^{(n)}. \end{aligned} \tag{2.30}$$

Sei  $a_{nn}^{(n)} \neq 0$ , so wird dieses System von unten her durch die sogenannte *Rückwärtselimination* gelöst: Man gewinnt  $x_n$  aus der letzten Gleichung, dann  $x_{n-1}$  aus der vorletzten usw., d.h., man

berechnet nacheinander

$$x_n = \frac{b_n^{(n)}}{a_{nn}^{(n)}} \quad \text{und} \quad x_i = \frac{1}{a_{ii}^{(i)}} \left( b_i^{(i)} - \sum_{k=i+1}^n a_{ik}^{(i)} x_k \right) \quad (2.31)$$

in der Reihenfolge  $i = n-1, n-2, \dots, 1$  (dabei  $a_{ik}^{(1)} := a_{ik}$  und  $b_i^{(1)} := b_i$  gesetzt). Die so errechnete Lösung  $[x_1, \dots, x_n]^T$  des Dreieckssystems ist *eindeutig bestimmt*. Damit ist  $[x_1, \dots, x_n]^T$  auch die einzige Lösung des ursprünglichen Systems (2.26), denn bei jedem Reduktionsschritt bleibt die Lösungsmenge unverändert, wenn man zum reduzierten System die erste Gleichung des vorangehenden Systems hinzunimmt. Aus dem Dreieckssystem (2.30) gewinnen wir also die eindeutig bestimmte Lösung des ursprünglichen Gleichungssystems (2.26).

### Beispiel 2.2:

Gesucht sind reelle  $x_1, x_2, x_3$  mit

$$4x_1 - 8x_2 + 2x_3 = -8 \quad (\text{G1})$$

$$-2x_1 + 6x_2 - 2x_3 = -1 \quad (\text{G2}) \quad (2.32)$$

$$3x_1 + 6x_2 - \frac{1}{2}x_3 = 4 \quad (\text{G3}).$$

Multipliziert man Gleichung (G1) auf beiden Seiten mit  $-2/4 = -1/2$  und subtrahiert sie dann von (G2), so folgt

$$2x_2 - x_3 = -5. \quad (\text{G4})$$

Entsprechend ergibt die Multiplikation von (G1) mit  $3/4$  und anschließende Subtraktion von (G3):

$$12x_2 - 2x_3 = 10. \quad (\text{G5})$$

Mit dem reduzierten System (G4), (G5) verfahren wir entsprechend. (G4) wird mit  $12/2 = 6$  multipliziert und von (G5) subtrahiert. Es folgt

$$4x_3 = 40. \quad (\text{G6})$$

Aus (G6) berechnet man  $x_3$ , aus (G4) anschließend  $x_2$  und aus (G1)  $x_1$ , d.h. wir lösen das Dreieckssystem

$$4x_1 - 8x_2 + 2x_3 = -8 \quad (\text{G1})$$

$$2x_2 - x_3 = -5 \quad (\text{G4})$$

$$4x_3 = 40 \quad (\text{G6})$$

von unten her auf. Man erhält

$$x_3 = 10, \quad x_2 = \frac{5}{2}, \quad x_1 = -2,$$

und hat damit die eindeutige bestimmte Lösung von (2.32) berechnet.

### Übung 2.3\*

Löse folgendes Gleichungssystem

$$\begin{aligned} 3x_1 + 8x_2 - 2x_3 + 2x_4 &= 3 \\ 6x_1 - x_2 + 2x_3 - 3x_4 &= 14 \\ -2x_1 + 3x_2 + 12x_3 + 5x_4 &= 4 \\ 4x_1 - 5x_2 + 6x_3 - 10x_4 &= 19 \end{aligned}$$

## 2.2.2 Matlab-Programme zur Lösung quadratischer Gleichungssysteme

Für das in Kurzform

$$\sum_{k=1}^n a_{ik}x_k = b_i, \quad i = 1, \dots, n \quad (2.33)$$

geschriebene Gleichungssystem (2.26) liefert uns das in Fig. 2.1 gezeigte Matlab-Programm eine direkte Umsetzung der in Abschnitt 2.2.1 beschriebenen Vorgehensweise.

Nutzen wir das numerische Verfahren zur Lösung des Gleichungssystems

$$\begin{aligned} 4x_1 - 18x_2 - 2x_3 - x_4 &= 5 \\ -12x_1 + 6x_2 - 5x_3 - 2x_4 &= -8 \\ -3x_1 + 7x_2 - 23x_3 + 8x_4 &= -1 \\ 2x_1 + 9x_2 + x_3 - 19x_4 &= 7, \end{aligned} \quad (2.34)$$

so erhalten wir die Lösung

$$x_1 = 0,7678, \quad x_2 = -0,0675, \quad x_3 = -0,1919, \quad x_4 = -0,3297.$$

Es lässt sich hierbei leicht nachprüfen, dass die erfolgreiche Berechnung auf dem Umstand beruht, dass alle Koeffizienten

$$a_{11}^{(1)} = a_{11}, \quad a_{22}^{(2)}, \quad a_{33}^{(3)} \text{ und } a_{44}^{(4)}$$

von Null verschieden sind. Bereits durch das einfache System

$$\begin{aligned} x_2 &= 1 \\ x_1 + x_2 &= 2 \end{aligned} \quad (2.35)$$



```

% Gaußscher Algorithmus ohne Pivotierung
%
function x = Gauss(A,b)
%
% OUTPUT VARIABLES:
% -----
% x: Lösungsvektor
%
% INPUT VARIABLES:
% -----
% A: quadratische n * n Matrix.
% b: rechte Seite.
%
n = size(b,1);
x = zeros(n,1);

% Reduktionsschritte zur Überführung in Dreiecksgestalt
for p=1:n-1
    for i=p+1:n
        if (abs(A(p,p)) < 10^(-25))
            error('Diagonalelement kleiner als 10^(-25)');
        end;
        c_ip = A(i,p) / A(p,p);
        b(i) = b(i) - c_ip * b(p);
        for k=p+1:n
            A(i,k) = A(i,k) - c_ip * A(p,k);
        end;
    end;
end;

% Rueckwärtselimination
for i=n:-1:1
    x(i) = b(i);
    if(i < n)
        for k=i+1:n
            x(i) = x(i) - A(i,k) * x(k);
        end;
    end;
    x(i) = x(i) / A(i,i);
end;

return;

```

Fig. 2.1: MATLAB-Implementierung des Gaußschen Algorithmus ohne Pivotierung

wird deutlich, dass dieser Sachverhalt auch bei Gleichungssystemen mit eindeutiger Lösung, hier  $x_1 = x_2 = 1$ , nicht stets gewährleistet werden kann. Während diese Problematik beim Gleichungssystem (2.35) noch offensichtlich ist, existieren auch Systeme, die trotz der eindeutigen Lösbarkeit zu einem Verfahrensabbruch beim Gaußschen Algorithmus führen, ohne dass dieser im Vorfeld notwendigerweise erwartet werden konnte.

So liefert das Modellproblem

$$\begin{aligned} x_1 + 2x_2 + 5x_3 &= 12 \\ 2x_1 + 4x_2 + 2x_3 &= 16 \\ 3x_1 + 10x_2 + x_3 &= 30 \end{aligned} \quad (2.36)$$

einen Verfahrensabbruch erst innerhalb des zweiten Schleifendurchlaufs, da zwar  $a_{11}^{(1)} = a_{11} \neq 0$  gilt, sich jedoch  $a_{22}^{(2)} = 0$  ergibt.

Wie wir im Folgenden sehen werden, lassen sich derartige Verfahrensabbrüche bei gewissen Matrizen durch eine sogenannte Pivotierung vermeiden.

Bezeichnen wir die Koeffizienten  $a_{ik}$ ,  $i, k = 1, \dots, n$  des Ausgangssystems mit dem Superskript (1), d.h.  $a_{ik}^{(1)}$ , so heißt ein Gleichungssystem regulär, falls für alle  $p = 1, \dots, n$  stets

$$\sum_{i,k=p}^n |a_{ik}^{(p)}| > 0 \quad (2.37)$$

gilt.<sup>9</sup> Anschaulich ausgedrückt bedeutet die Bedingung (2.37), dass in den auftretenden Systemen niemals alle Koeffizienten identisch verschwinden. Aus der Eigenschaft (2.37) lässt sich sowohl

$$\sum_{k=p}^n |a_{pk}^{(p)}| > 0 \quad (2.38)$$

als auch

$$\sum_{i=p}^n |a_{ip}^{(p)}| > 0 \quad (2.39)$$

folgern. Bei einem regulären Gleichungssystem kann daher im Fall  $a_{pp}^{(p)} = 0$  durch einen Zeilentauch der  $p$ -ten mit einer  $j$ -ten Zeile ( $j > p$ ) ein nicht verschwindendes Diagonalelement generiert werden, so dass der Gauß-Algorithmus fortgesetzt werden kann. Hierbei nimmt man Zeilenvertauschungen meistens so vor, dass ein *betragsgrößter Koeffizient der ersten Spalte in die linke obere Ecke* des zu behandelnden (reduzierten) Systems gelangt. Damit werden die Faktoren  $c_{ip}$ , mit denen die erste Gleichung des betrachteten Systems multipliziert wird, betragsmäßig  $\leq 1$ . Nach anschließender Subtraktion von den übrigen Gleichungen des Systems entsteht ein reduziertes System, in dem (normalerweise) die Unterschiede zwischen den Gleichungen nicht

<sup>9</sup> Die Eigenschaft (2.37) bedeutet, dass die Matrix des Gleichungssystems regulär ist.

vom Anteil der ersten Gleichung des vorangehenden Systems »erdrückt« worden sind, so dass das Weiterrechnen Erfolg verspricht. Dieses Verfahren heißt Gaußscher Algorithmus mit *Spalten-Pivotierung*.

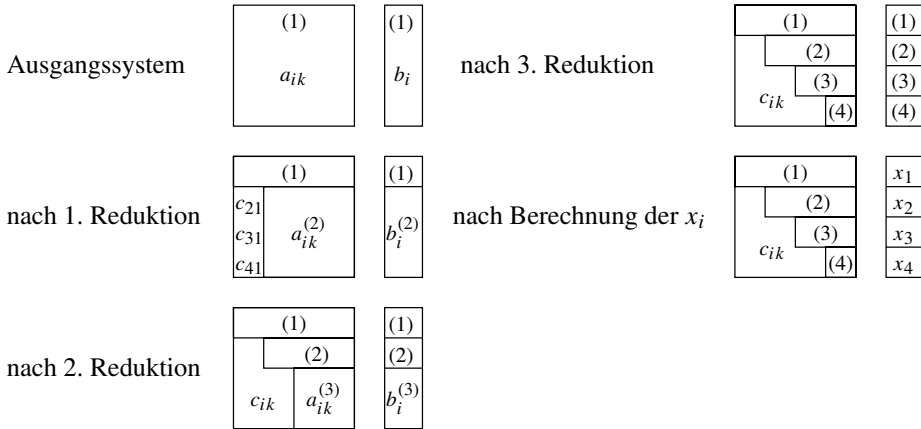


Fig. 2.2: Speicherbelegung für  $n = 4$

Zur ökonomischen Speicherbelegung im Computer verfährt man so, wie es Fig. 2.2 für den Fall  $n = 4$  zeigt: Die ursprünglich vorhandenen Koeffizienten  $a_{ik} =: a_{ik}^{(1)}$ ,  $b_i =: b_i^{(1)}$ , werden nach und nach von den  $c_{ik}$  und den reduzierten Größen  $a_{ik}^{(j)}$ ,  $b_i^{(j)}$  überschrieben. Zum Schluss speichert man die berechneten  $x_n, x_{n-1}, \dots, x_1$  auf den Speicherplätzen ab, auf denen am Anfang die  $b_i$  standen.

Auf diese Weise kommt man mit minimalem Speicherplatzbedarf aus. Man benötigt nur die  $n^2 + n$  Speicherplätze für Koeffizienten  $a_{ik}$  und  $b_i$ , sowie wenige zusätzliche Speicherplätze zur Organisation (Umspeichern, Festhalten eines Zeilenindex usw.).

Wird analog eine Spaltenpivotierung derart durchgeführt, dass ein betragsgrößer Koeffizient der ersten Zeile des (reduzierten) Systems in die linke obere Ecke überführt wird, so spricht man von einer *Zeilen-Pivotierung*. Wird ein betragsmäßig größtes Element des (reduzierten) Systems durch eine Kombination von Zeilen- und Spaltentausch in die linke obere Ecke transformiert, so wird von einer *vollständigen Pivotierung* gesprochen.

Mit dem in Fig. 2.3 dargestellten Gauß-Algorithmus mit Spaltenpivotierung erhalten wir für die Beispielsysteme (2.35) und (2.36) im Gegensatz zum vorherigen Verfahren, die gesuchten Lösungen

$$x_1 = x_2 = 1$$

respektive

$$x_1 = 3, \quad x_2 = 2, \quad x_3 = 1.$$

Neben der grundsätzlichen Durchführbarkeit des Gauß-Verfahrens für reguläre Systeme ergibt sich durch die Pivotierung auch aus Sicht der Rundungsfehlereinflüsse eine Stabilisierung der

```

% Gaußscher Algorithmus mit Spaltenpivotierung
%
function x = Gauss_Pivot(A,b)
%
% OUTPUT VARIABLES:
% -----
% x: Lösungsvektor
%
% INPUT VARIABLES:
% -----
% A: quadratische n * n Matrix.
% b: rechte Seite.
%
n = size(b,1);
x = zeros(n,1);

% Reduktionsschritte zur Überführung in Dreiecksgestalt
for p=1:n-1
    % Spaltenpivotierung
    max = abs(A(p,p));
    index = p;
    for j=p+1:n
        if(max < abs(A(j,p)))
            index = j;
        end;
    end;
    if(index > p)
        tmp = b(p);
        b(p) = b(index);
        b(index) = tmp;
        for j=p:n
            tmp = A(p,j);
            A(p,j) = A(index,j);
            A(index,j) = tmp;
        end;
    end;
    % Ende der Spaltenpivotierung
    for i=p+1:n
        c_ip = A(i,p) / A(p,p);
        b(i) = b(i) - c_ip * b(p);
        for k=p+1:n
            A(i,k) = A(i,k) - c_ip * A(p,k);
        end;
    end;
end;

% Rueckwärtselimination
for i=n:-1:1
    x(i) = b(i);
    if(i < n)
        for k=i+1:n
            x(i) = x(i) - A(i,k) * x(k);
        end;
    end;
    x(i) = x(i) / A(i,i);
end;

return;

```

Fig. 2.3: MATLAB-Implementierung des Gaußschen Algorithmus mit Pivotierung

Methode. Zur Verdeutlichung dieser Auswirkungen betrachten wir das System

$$\varepsilon x_1 + 2x_2 = 1$$

$$x_1 + x_2 = 1$$

mit einem frei wählbaren Parameter  $\varepsilon \in \mathbb{R} \setminus \{2\}$ .

Die exakte Lösung des Systems lautet

$$x_1 = \frac{1}{2 - \varepsilon}, \quad x_2 = \frac{1 - \varepsilon}{2 - \varepsilon}. \quad (2.40)$$

In der Tabelle 2.1 sind die, mit dem vorgestellten Algorithmus erzielten Ergebnisse für  $\varepsilon = 10^{-i}$  mit  $i \in \{1, 2, 10, 12, 13, 15, 16, 20\}$  dargestellt.

Tabelle 2.1: Gauß-Algorithmus mit und ohne Spalten-Pivotierung

$\varepsilon$	Gauß-Algorithmus			
	ohne Pivotierung		mit Pivotierung	
	$x_1$	$x_2$	$x_1$	$x_2$
$10^{-1}$	0,5263	0,4737	0,5263	0,4737
$10^{-2}$	0,5025	0,4975	0,5025	0,4975
$10^{-10}$	0,5000	0,5000	0,5000	0,5000
$10^{-12}$	0,5000	0,5000	0,5000	0,5000
$10^{-13}$	0,4996	0,5000	0,5000	0,5000
$10^{-15}$	0,5551	0,5000	0,5000	0,5000
$10^{-16}$	0	0,5000	0,5000	0,5000
$10^{-20}$	0	0,5000	0,5000	0,5000

Während zunächst für  $\varepsilon = 10^{-i}$  mit  $i = 1, 2, 10, 12$  beide Verfahren identische Resultate liefern, zeigt sich beim Algorithmus ohne Pivotierung ein zunehmend stärker werdender Rundungsfehlereinfluss bei kleiner werdendem Parameter  $\varepsilon$ , der für  $\varepsilon = 10^{-i}$  mit  $i = 13, 15, 16$  und 20 zu unbrauchbaren Ergebnissen führt, obwohl kein Verfahrensabbruch eintritt. Die Spalten-Pivotierung liefert hier eine deutlich erkennbare Stabilisierung der Methode, so dass die im Grenzfalle  $\varepsilon \rightarrow 0$  laut (2.40) vorliegende Lösung  $x_1 = x_2 = \frac{1}{2}$  erzielt wird. Eine mathematische Analyse der Rundungsfehler wird in [96] präsentiert.

**Bemerkung:** Eine formale Spezifizierung der Gleichungssysteme, für die der Gaußsche Algorithmus mit und evtl. sogar ohne Pivotierung bei jeweiliger Vernachlässigung von Rundungsfehlern die eindeutig bestimmte Lösung ermittelt, wird in Abschnitt 3.8.3 präsentiert.

**Übung 2.4\***

Löse das folgende Gleichungssystem durch den Gaußschen Algorithmus mit Spaltenpivotierung:

$$\begin{array}{rrcr} 3078x_1 - & 261x_2 + & 1587x_3 = & -401 \\ 0,126x_1 - & 0,049x_2 - & 3,956x_3 = & 2,647 \\ 1,981 \cdot 10^6 x_1 + & 7,217 \cdot 10^6 x_2 + & 0,221 \cdot 10^6 x_3 = & 0,512 \cdot 10^5. \end{array}$$

**2.2.3 Singuläre lineare Gleichungssysteme**

Es sei ein beliebiges quadratisches lineares Gleichungssystem

$$\sum_{k=1}^n a_{ik} x_k = b_i, \quad i = 1, \dots, n \quad (2.41)$$

vorgelegt, von dem wir zunächst nicht wissen, ob es *regulär* ist (d.h. ob es durch den Gaußschen Algorithmus in ein Dreieckssystem übergeht), oder ob es *singulär* ist (d.h. nicht regulär).

Um das Gleichungssystem zu lösen wenden wir den Gaußschen Algorithmus an, wie er in Abschn. 2.2.1 beschrieben ist, d.h. mit Zeilen- und Spaltenvertauschungen, falls nötig. Hierbei können wir an *Totalpivotierungen* denken, bei denen stets der absolut größte Koeffizient auf der linken Seite eines (reduzierten) Systems in die linke obere Ecke gebracht wird.

Ist das System nicht regulär, so muss schließlich ein (reduziertes) Gleichungssystem entstehen, in dem *alle Koeffizienten auf der linken Seite Null* sind.

Die Reduktion bricht bei diesem Gleichungssystem ab. Schreibt man die jeweils ersten Gleichungen aller bis dahin erhaltenen Gleichungssysteme untereinander und fügt das zuletzt erhaltene System, das links nur Nullen aufweist, hinzu, so ergibt sich

$$\begin{array}{rrrrrr} a_{11}x_1 + a_{12}x_2 + \dots + a_{1p}x_p + a_{1,p+1}x_{p+1} + \dots + a_{1n}x_n = & b_1 \\ a_{22}^{(2)}x_2 + \dots + a_{2p}^{(2)}x_p + a_{2,p+1}^{(2)}x_{p+1} + \dots + a_{2n}^{(2)}x_n = & b_2 \\ \vdots & \vdots \\ a_{pp}^{(p)}x_p + a_{p,p+1}^{(p)}x_{p+1} + \dots + a_{pn}^{(p)}x_n = & b_p^{(p)} \\ 0 + \dots + \dots + 0 = & b_{p+1}^{(p)} \\ \vdots & \vdots \\ 0 + \dots + \dots + 0 = & b_n^{(p)} \end{array} \quad (2.42)$$

mit  $a_{11} \neq 0, a_{22}^{(2)} \neq 0, \dots, a_{pp}^{(p)} \neq 0, (p < n)$ .

Wir wollen ein solches System ein *p-zeiliges Trapezsystem* nennen (wobei dieser Name von den ersten  $p$  Zeilen herrührt). Dieses System besitzt, aus den Abschn. 2.2.1 genannten Gründen, die gleiche Lösungsmenge wie das Ausgangssystem (2.41). Man sieht sofort:

**Folgerung 2.4:**

- (a) Ist in (2.42) eins der  $b_{p+1}^{(p)}, \dots, b_n^{(p)}$  ungleich Null, so ist das Gleichungssystem unlösbar;
- (b) gilt dagegen  $b_{p+1}^{(p)} = b_{p+2}^{(p)} = \dots = b_n^{(p)} = 0$ , so gibt es unendlich viele Lösungen.

**Beweis:**

(a) ist klar. Zu (b): Setzt man, zur besseren Unterscheidung

$$x_{p+1} = t_1, \quad x_{p+2} = t_2, \quad \dots, \quad x_n = t_{n-p} \quad (2.43)$$

und bringt man die zugehörigen Glieder in (2.42) auf die rechte Seite, so kann man (2.42) im Falle (b) wieder »von unten nach oben« auflösen. Die Lösungen hängen von den frei wählbaren Parametern  $t_1, \dots, t_{n-p}$  ab.  $\square$

Das skizzierte Auflösen des Systems (2.42) »von unten nach oben« führt auf

$$x_p = \frac{1}{a_{pp}^{(p)}} \left( b_p^{(p)} - \sum_{j=1}^{n-p} a_{p,p+j}^{(p)} t_j \right)$$

und für  $i = p-1, p-2, \dots, 1$ :

$$x_i = \frac{1}{a_{ii}^{(i)}} \left( b_i^{(i)} - \sum_{j=1}^{n-p} a_{i,p+j}^{(i)} t_j - \sum_{k=i+1}^p a_{i,k}^{(i)} x_k \right). \quad (2.44)$$

In der oberen Gleichung ist  $x_p$  in Abhängigkeit von  $t_1, \dots, t_{n-p}$  gegeben. Die untere Gleichung wird nun in der Reihenfolge  $i = p-1, p-2, \dots, 1$  ausgewertet, wobei man jeweils rechts alle vorher berechneten Ausdrücke für  $x_{i+1}, x_{i+2}, \dots, x_p$  einsetzt. So entstehen Formeln für die Lösungszahlen  $x_i$ , bei denen die rechten Seiten nur noch von den Parametern  $t_1, \dots, t_{n-p}$  abhängen. Sie haben die Form:

$$x_i = u_i + \sum_{j=1}^{n-p} v_{ij} t_j, \quad i = 1, \dots, n. \quad (2.45)$$

Dabei wurden auch die Gleichungen (2.43) mit eingeordnet, denn sie haben die Gestalt (2.45), wenn man den Koeffizienten von  $t_j$  gleich 1 setzt und die übrigen Null, also  $v_{p+s,j} = \delta_{sj}$  (Kronecker-Symbol).

**Beispiel 2.3:**

Es soll die Lösungsmenge  $L$  des folgenden Gleichungssystems berechnet werden

$$\begin{aligned} 8x_1 + 2x_2 - 3x_3 + 5x_4 &= 1 \\ 3x_1 - 7x_2 + 5x_3 - x_4 &= -2 \\ 14x_1 - 12x_2 + 7x_3 + 3x_4 &= -3 \\ -5x_1 - 9x_2 + 8x_3 - 6x_4 &= -3. \end{aligned} \quad (2.46)$$

Der Gaußsche Algorithmus mit Spaltenpivotierung ergibt nach 2 Reduktionen das Trapezsysteem

$$\begin{aligned} 14x_1 - 12x_2 + 7x_3 + 3x_4 &= -3 \\ -13,2857x_2 + 10,5000x_3 - 4,2985x_4 &= -4,0714 \\ 0 &= 0 \\ 0 &= 0. \end{aligned} \quad (2.47)$$

Das Gleichungssystem ist also singulär — und lösbar! Man setzt nun  $x_3 = t_1$ ,  $x_4 = t_2$ , löst dann die zweite Gleichung in (2.47) nach  $x_2$  auf, setzt den gefundenen Ausdruck für  $x_2$  in die erste Gleichung von (2.47) ein und löst nach  $x_1$  auf. Damit erhält man die Lösungsgesamtheit in der Form

$$\left. \begin{aligned} x_1 &\doteq 0,04839 - 0,17742t_1 + 0,53226t_2 \\ x_2 &\doteq 0,30645 - 0,79032t_1 + 0,37097t_2 \\ x_3 &= t_1 \\ x_4 &= t_2 \end{aligned} \right\} (t_1, t_2 \in \mathbb{R}). \quad (2.48)$$

Die Lösungsmenge  $L$  besteht also aus allen Vektoren  $\mathbf{x} = [x_1, x_2, x_3, x_4]^T$  von der Form (2.48), wobei die *Parameter*  $t_1, t_2$  als beliebige reelle Zahlen gewählt werden dürfen. (Die Dezimalzahlen sind gerundet, was durch  $\doteq$  statt  $=$  symbolisiert wird.)

**Algebraische Struktur der Lösungsmenge im singulären Fall**

Die Gleichung (2.45) erhält mit

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}, \quad \mathbf{v}_j = \begin{bmatrix} v_{1j} \\ \vdots \\ v_{nj} \end{bmatrix}, \quad j = 1, \dots, n-p,$$

die vektorielle Gestalt

$$\mathbf{x} = \mathbf{u} + \sum_{j=1}^{n-p} \mathbf{v}_j t_j, \quad (t_j \in \mathbb{R}). \quad (2.49)$$

Die  $\mathbf{v}_j$  sind dabei linear unabhängig, wegen  $v_{p+s,j} = \delta_{sj}$  für  $s, j = 1, \dots, n-p$  (denn lässt man die ersten  $p$  Koordinaten weg, so werden die  $\mathbf{v}_j$  zu Koordinateneinheitsvektoren).



Gleichung (2.49) ist die Parameterdarstellung einer  $(n - p)$ -dimensionalen Mannigfaltigkeit. Also:

**Satz 2.8:**

(a) Ist das lineare Gleichungssystem

$$\sum_{k=1}^n a_{ik}x_k = b_i \quad (i = 1, \dots, n)$$

singulär, so ist die Lösungsmenge entweder leer oder eine  $(n - p)$ -dimensionale lineare Mannigfaltigkeit ( $p < n$  wie in (2.42)).

(b) Ist die rechte Seite dabei Null ( $b_i = 0$  für alle  $i = 1, \dots, n$ ), so ist die Lösungsmenge  $L$  ein  $(n - p)$ -dimensionaler Unterraum (d.h.  $\mathbf{u} = \mathbf{0}$  in (2.49)).

**Beweis:**

(a) ist klar. Zu (b): Im Fall  $b_j = 0$  ( $i = 1, \dots, n$ ) ist  $\mathbf{x}_0 = \mathbf{0}$  eine Lösung. Folgerung 2.1 in Abschn. 2.1.3 ergibt, dass  $L - \mathbf{x}_0 = L - \mathbf{0} = L$  ein  $(n - p)$ -dimensionaler Unterraum ist.  $\square$

**Bemerkung:** Satz 2.8 macht klar, dass der Gaußsche Algorithmus, unabhängig von der speziellen Wahl der Zeilen- und Spaltenvertauschungen, stets auf ein Trapezzsystem mit der gleichen Zeilenzahl  $p$  führt. Denn die Dimension  $n - p$  der Lösungsmenge — und damit die Zahl  $p$  — ist sicherlich unabhängig vom Berechnungsverfahren für die Lösungen.

**Beispiel 2.3:** (Fortsetzung) Mit den Vektoren

$$\mathbf{u} \doteq \begin{bmatrix} 0,04839 \\ 0,30645 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v}_1 \doteq \begin{bmatrix} -0,17742 \\ -0,79032 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 \doteq \begin{bmatrix} 0,53226 \\ 0,37097 \\ 0 \\ 1 \end{bmatrix}$$

gewinnt man die Aussage: Die Lösungsmenge  $L$  des Gleichungssystems (2.46) besteht aus allen Vektoren der Form

$$\mathbf{x} = \mathbf{u} + t_1 \mathbf{v}_1 + t_2 \mathbf{v}_2 \quad (t_1, t_2 \text{ beliebig reell}).$$

**Veranschaulichung:** Im Fall eines dreireihigen quadratischen linearen Gleichungssystems lässt sich die Lösungsstruktur gut anschaulich machen. Es sei das folgende Gleichungssystem gegeben:

$$\begin{array}{ll} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 & \hat{\mathbf{a}}_1^T \cdot \mathbf{x} = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 & \text{kurz } \hat{\mathbf{a}}_2^T \cdot \mathbf{x} = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 & \hat{\mathbf{a}}_3^T \cdot \mathbf{x} = b_3 \end{array}$$

mit den Zeilenvektoren  $\hat{\mathbf{a}}_i = [a_{i1}, a_{i2}, a_{i3}]$  und dem Spaltenvektor  $\mathbf{x} = [x_1, x_2, x_3]^T$ . Die Gleichungen rechts können als Hessesche Normalformen dreier Ebenen  $E_1, E_2, E_3$  aufgefasst werden, da

wir uns die Gleichungen durch Multiplikation mit Faktoren  $\neq 0$  so umgewandelt denken können, dass  $|\hat{a}_i| = 1$  ist, für alle  $i = 1, 2, 3$ .

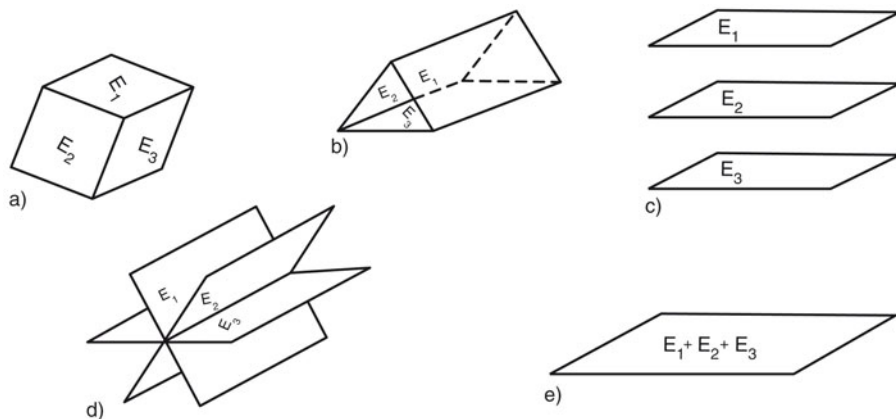


Fig. 2.4: Zur Lösbarkeit linearer Gleichungssysteme

Die gemeinsamen Punkte aller drei Ebenen  $E_1, E_2, E_3$  bilden die Lösungen. Figur 2.4 zeigt fünf charakteristische Möglichkeiten: Im Falle a) haben wir eindeutige Lösbarkeit, da  $E_1 \cap E_2 \cap E_3$  aus einem Punkt besteht. b) und c) zeigen unlösbare Fälle.  $E_1 \cap E_2 \cap E_3 = \emptyset$ , d), e) zeigen Fälle mit unendlich vielen Lösungen: d) Gerade, e) Ebene.

### Übung 2.5:

- (a) Gib die Lösungsmenge des folgenden Gleichungssystems an:

$$\begin{aligned} 3x_1 + 4x_2 - 5x_3 + x_4 &= 1 \\ 7x_1 - 2x_2 + x_3 - 2x_4 &= -2 \\ x_1 - 10x_2 + 11x_3 - 4x_4 &= -4 \\ -4x_1 + 6x_2 - 6x_3 - 3x_4 &= -3. \end{aligned}$$

- (b) Gib die Lösungsmenge für den Fall an, dass die  $-3$  auf der rechten Seite unten in  $-2$  verwandelt wird.
- (c) Streiche aus dem Gleichungssystem in (a) die letzte Zeile heraus, wie auch die Spalte mit  $x_4$ . Es entsteht ein dreireihiges quadratisches Gleichungssystem. Skizziere die drei Ebenen, die durch die Gleichungen gegeben sind, in einem räumlichen Koordinatensystem (man denke sich einen genügend großen »gläsernen« achsenparallelen Quader gezeichnet und veranschauliche die Ebenen durch ihre Schnittlinien mit den Quaderwänden). Wie liegt hier die Lösungsmenge?

### 2.2.4 Allgemeiner Satz über die Lösbarkeit linearer quadratischer Gleichungssysteme

Das Gleichungssystem (2.26) nennen wir ein *quadratisches* System, da ebenso viele Zeilen wie Unbekannte vorkommen. Aus den Spalten des Systems bildet man die *Spaltenvektoren*:

$$\mathbf{a}_1 = \begin{bmatrix} a_{11} \\ \vdots \\ a_{n1} \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} a_{12} \\ \vdots \\ a_{n2} \end{bmatrix}, \quad \dots, \quad \mathbf{a}_n = \begin{bmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}. \quad (2.50)$$

Damit erhält unser Gleichungssystem die knappere vektorielle Gestalt

$$x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \dots + x_n \mathbf{a}_n = \mathbf{b} \quad (\mathbf{a}_k, \mathbf{b} \in \mathbb{R}^n). \quad (2.51)$$

Mit den Zeilen des Systems geht man entsprechend vor: Man kann sie zu *Zeilenvektoren* zusammenfassen:

$$\begin{aligned} \hat{\mathbf{a}}_1 &= [a_{11}, a_{12}, \dots, a_{1n}] \\ &\vdots \\ \hat{\mathbf{a}}_n &= [a_{n1}, a_{n2}, \dots, a_{nn}]. \end{aligned} \quad (2.52)$$

Mit solchen Zeilenvektoren rechnet man völlig analog wie mit Spaltenvektoren: Man kann sie addieren, subtrahieren, mit  $\lambda \in \mathbb{R}$  multiplizieren usw. Führen wir noch den Spaltenvektor

$$\mathbf{x} = [x_1, \dots, x_n]^T$$

ein, so lässt sich unser Gleichungssystem so beschreiben

$$\hat{\mathbf{a}}_i^T \cdot \mathbf{x} = b_i, \quad i = 1, \dots, n. \quad (2.53)$$

Voilà! — Damit gilt der folgende zusammenfassende Satz

#### Satz 2.9:

Bei einem quadratischen linearen Gleichungssystem

$$\sum_{k=1}^n a_{ik} x_k = b_i, \quad i = 1, \dots, n \quad (2.54)$$

tritt genau einer der folgenden Fälle ein:

- 1. Fall: (a)** Das Gleichungssystem ist *regulär*, d.h. der Gaußsche Algorithmus liefert ein *Dreieckssystem*, in dem alle Diagonalkoeffizienten  $a_{11}, a_{22}^{(2)}, \dots, a_{nn}^{(n)}$  ungleich Null sind. Dies bedeutet, dass das Gleichungssystem für jede rechte Seite  $\mathbf{b} \in \mathbb{R}^n$  genau eine Lösung  $\mathbf{x} = [x_1, \dots, x_n]^T \in \mathbb{R}^n$  hat.

(b) Das Gleichungssystem ist genau dann *regulär*, wenn die *Spaltenvektoren*  $\mathbf{a}_1, \dots, \mathbf{a}_n$  *linear unabhängig* sind. Dies ist genau dann der Fall, wenn die *Zeilenvektoren* *linear unabhängig* sind.

**2. Fall: (c)** Das Gleichungssystem ist *singulär*, d.h. der Gaußsche Algorithmus liefert ein  $p$ -zeiliges Trapezsystem, mit  $p < n$ . Das bedeutet: Es gibt *entweder keine Lösung* ( $b_{p+1}^{(p)}, \dots, b_n^{(p)}$  nicht alle Null), *oder* es gibt *unendlich viele Lösungen* (im Falle  $b_{p+1}^{(p)} = \dots = b_n^{(p)} = 0$ ). Im letzteren Fall ist die Lösungsmenge eine *lineare Mannigfaltigkeit* der Dimension  $n - p$  im Raum  $\mathbb{R}^n$ .

(d) Das Gleichungssystem ist genau dann *singulär* (mit  $p$ -zeiligem Trapezsystem), wenn die *Spaltenvektoren*  $\mathbf{a}_1, \dots, \mathbf{a}_n$  *linear abhängig* sind. In diesem Fall ist  $p$  die maximale Anzahl linear unabhängiger Vektoren unter den  $\mathbf{a}_1, \dots, \mathbf{a}_n$ . Für *Zeilenvektoren* gilt die entsprechende Aussage.

### Beweis:

<sup>10</sup>Die Aussagen (a), (c) sind durch die vorangehenden Abschnitte klar. Es müssen nur noch (b) und (d) bewiesen werden.

*Beweis von (d):* Im *singulären* Fall verwandelt der Gaußsche Algorithmus das homogene Gleichungssystem

$$\sum_{k=1}^n x_k \mathbf{a}_k = \mathbf{0}, \quad i = 1, \dots, n \quad (2.55)$$

in ein  $p$ -zeiliges Trapezsystem. Wir denken uns nun alle vorgenommenen Zeilen- und Spaltenvertauschungen schon vor Beginn der Rechnung ausgeführt, und bezeichnen das entstandene System wieder wie in (2.55). Dann liefert der Gaußsche Algorithmus ein  $p$ -zeiliges Trapezsystem auch *ohne* Vertauschungen von Zeilen und Spalten. Von dieser Situation gehen wir aus!

(d<sub>1</sub>) Die Lösungsmenge von (2.55) wird nach Satz 2.8 (b) im vorigen Abschnitt durch

$$\mathbf{x} = \sum_{j=1}^{n-p} \mathbf{v}_j t_j \quad (t_j = x_{p+j}) \quad (2.56)$$

beschrieben. Wählt man alle  $t_j = x_{p+j} = 0$ , so verkürzt sich (2.55) zu  $\sum_{k=1}^p x_k \mathbf{a}_k = \mathbf{0}$ . Nach (2.56) ist dann aber  $\mathbf{x} = \mathbf{0}$ , also auch  $x_1 = \dots = x_p = 0$ , d.h. die  $\mathbf{a}_1, \dots, \mathbf{a}_p$  sind linear unabhängig.

Im Fall  $t_j = x_{p+j} = -1$  und  $t_i = 0$  für alle  $i \neq j$  ( $i \in \{1, \dots, n-p\}$ ) wird (2.55) zu

$$\sum_{k=1}^p x_k \mathbf{a}_k - \mathbf{a}_{p+j} = \mathbf{0},$$

<sup>10</sup> Kann vom anwendungsorientierten Leser überschlagen werden.

wobei die  $x_k$  aus (2.56) gewonnen werden. D.h.:  $\mathbf{a}_{p+j}$  ist eine Linearkombination der  $\mathbf{a}_1, \dots, \mathbf{a}_p$  (für jedes  $j \in \{1, \dots, n+p\}$ ). Folglich liegen alle  $\mathbf{a}_{p+j}$  in dem von  $\mathbf{a}_1, \dots, \mathbf{a}_p$  aufgespannten Unterraum der Dimension  $p$ . Folglich ist  $p$  die maximale Anzahl linear unabhängiger Vektoren unter den Spaltenvektoren  $\mathbf{a}_1, \dots, \mathbf{a}_p$ .

(d<sub>2</sub>) Beschreibt man das Gleichungssystem (2.55) mit den Zeilenvektoren  $\hat{\mathbf{a}}_1, \dots, \hat{\mathbf{a}}_n$ , so erhält es die Gestalt

$$\hat{\mathbf{a}}_i^T \cdot \mathbf{x} = 0, \quad \mathbf{x} = [x_1, \dots, x_n]^T. \quad (2.57)$$

Alle Lösungen  $\mathbf{x}$  stehen also rechtwinklig auf den  $\hat{\mathbf{a}}_i^T$ , und damit auf jedem Vektor des von  $\hat{\mathbf{a}}_1^T, \dots, \hat{\mathbf{a}}_n^T$  aufgespannten Unterraumes  $U \subset \mathbb{R}^n$ . Kurz: Der Lösungsraum  $L$  ist das orthogonale Komplement von  $U$ . Mit  $\dim L = n - p$  und der Gleichung  $\dim U + \dim L = n$  (s. Folg. 2.3, Abschn. 2.1.4) folgt  $\dim U = p$ . Da die  $\hat{\mathbf{a}}_1^T, \dots, \hat{\mathbf{a}}_n^T$  den Raum  $U$  aufspannen, ist die maximale Anzahl linear unabhängiger unter ihnen gleich  $\dim U = p$ .

*Beweis* von (b): Ist das Gleichungssystem (2.54) regulär, so hat es nach (a) für jede rechte Seite genau eine Lösung, also hat  $\sum_{i=1}^n \mathbf{a}_i x_k = \mathbf{0}$  ( $i = 1, \dots, n$ ) nur die Lösung  $x_1 = x_2 = \dots = x_n = 0$ , d.h. die  $\mathbf{a}_1, \dots, \mathbf{a}_n$  sind linear unabhängig. — Setzen wir umgekehrt die  $\mathbf{a}_1, \dots, \mathbf{a}_n$  als linear unabhängig voraus, so hat  $\sum_{i=1}^n \mathbf{a}_i x_k = \mathbf{0}$  nur die Lösung  $x_1 = \dots = x_n = 0$ . Damit ist  $\sum_{i=1}^n \mathbf{a}_i x_k = \mathbf{0}$  ein reguläres Gleichungssystem, denn im singulären Fall läge ja keine eindeutige Lösbarkeit vor. Der Gaußsche Algorithmus erzeugt also ein Dreieckssystem, d.h. auch (2.54) ist regulär.

Die lineare Unabhängigkeit der Zeilenvektoren  $\hat{\mathbf{a}}_i$  folgt im regulären Fall wie oben in (d<sub>2</sub>). Man hat  $p = n$  zu setzen.

Da die  $a_{ik}$  ein quadratisches Schema bilden, folgt umgekehrt aus Symmetriegründen aus der linearen Unabhängigkeit der Zeilenvektoren auch die der Spaltenvektoren.  $\square$

### Folgerung 2.5:

Ist ein quadratisches lineares Gleichungssystem

$$\sum_{k=1}^n x_k \mathbf{a}_k = \mathbf{b} \quad (\mathbf{a}_k, \mathbf{b} \in \mathbb{R}^n)$$

für eine rechte Seite  $\mathbf{b}$  eindeutig lösbar, so ist es für jede beliebige rechte Seite eindeutig lösbar.

### Beweis:

Eindeutige Lösbarkeit für eine rechte Seite bedeutet, dass der Gauß-Algorithmus ein Dreieckssystem erzeugt. Für jede andere rechte Seite entsteht aber links vom Gleichheitszeichen die gleiche Dreiecksstruktur, d.h. es liegt bei jeder linken Seite eindeutige Lösbarkeit vor.  $\square$

Es ist klar, dass diese schöne Eigenschaft für beliebige Gleichungen  $f(x) = b$  nicht allgemein gilt. Sie ist eine Besonderheit linearer quadratischer Gleichungssysteme. Sie ist die Kernaussage der sogenannten *Fredholmschen Alternative*: Entweder liegt für alle rechten Seiten eindeutige Lösbarkeit vor, oder für keine!

### 2.2.5 Rechteckige Systeme, Rangkriterium

**Zur praktischen Lösung.** Gleichungssysteme der Form

$$\begin{array}{cccc} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n & = & b_2 \\ \vdots & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n & = & b_m \end{array} \quad (2.58)$$

mit  $n \neq m$ , nennen wir *rechteckige lineare Gleichungssysteme*. Durch Ergänzen von Nullzeilen oder Nullspalten kann man sie stets zu quadratischen Systemen erweitern, womit sie auf den behandelten Fall zurückgeführt sind.

Man kann aber auch den Gaußschen Algorithmus völlig analog zum Vorherigen direkt auf Rechtecksysteme anwenden.

Oft liegt dabei keine eindeutige Lösbarkeit vor. Man gelangt durch die Rekursionsschritte i. Allg. zu einem Trapezsystem mit unterschiedlicher Zeilen- und Spaltenzahl auf der linken Seite. Dann geht man so vor, wie in Abschn. 2.2.3 erläutert.

Entsteht jedoch ein Dreieckssystem mit noch folgenden Nullzeilen (nur im Fall  $m > n$  möglich), so liegt wieder eindeutige Lösbarkeit vor, und wir berechnen die Lösungen wie in Abschn. 2.2.1.

**Rangkriterium, Lösungsstruktur.**<sup>11</sup> Mit den Spaltenvektoren

$$a_k = \begin{bmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{bmatrix}, \quad k = 1, \dots, n \quad \text{und} \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

aus  $\mathbb{R}^m$  bekommt das Gleichungssystem (2.58) die Form

$$\sum_{k=1}^n x_k a_k = b \quad (a_k, b \in \mathbb{R}^m). \quad (2.59)$$

Dabei nehmen wir  $n, m \in \mathbb{N}$  als beliebig an, also gleich oder ungleich. Das Gleichungssystem heißt *homogen*, wenn  $b = 0$  ist, andernfalls *inhomogen*.

<sup>11</sup> Kann vom praxisorientierten Leser beim ersten Lesen übersprungen werden.

**Definition 2.7:**

Es seien  $\mathbf{c}_1, \dots, \mathbf{c}_k$  beliebige Vektoren aus  $\mathbb{R}^n$ . Als *Rang* der Menge  $\{\mathbf{c}_1, \dots, \mathbf{c}_k\}$ ,

$$\text{Rang}\{\mathbf{c}_1, \dots, \mathbf{c}_k\},$$

bezeichnet man die maximale Anzahl linear unabhängiger Vektoren unter ihnen.

Damit gilt für beliebige lineare Gleichungssysteme, rechteckig oder quadratisch, der

**Satz 2.10:**

(*Rangkriterium*) Das lineare Gleichungssystem (2.59) ist genau dann lösbar, wenn folgendes gilt:

$$\text{Rang}\{\mathbf{a}_1, \dots, \mathbf{a}_n\} = \text{Rang}\{\mathbf{a}_1, \dots, \mathbf{a}_n, \mathbf{b}\}. \quad (2.60)$$

Das Gleichungssystem ist genau dann eindeutig lösbar, wenn die Ränge in (2.60) gleich  $n$  sind, also gleich der Anzahl der Unbekannten.

**Beweis:**

Lösbarkeit von (2.59) bedeutet, dass  $\mathbf{b}$  als Linearkombination der  $\mathbf{a}_1, \dots, \mathbf{a}_n$  geschrieben werden kann, d.h., dass  $\mathbf{b}$  in dem von  $\mathbf{a}_1, \dots, \mathbf{a}_n$  aufgespannten Unterraum  $U$  liegt, d.h.:  $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$  und  $\{\mathbf{a}_1, \dots, \mathbf{a}_n, \mathbf{b}\}$  spannen den gleichen Unterraum auf; d.h.: Beide Mengen  $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$  und  $\{\mathbf{a}_1, \dots, \mathbf{a}_n, \mathbf{b}\}$  haben gleiche Maximalzahl linear unabhängiger Vektoren; d.h.: Es gilt (2.60). — Die Eindeutigkeitsaussage folgt aus Satz 2.3 (b) in Abschn. 2.1.3.  $\square$

**Satz 2.11:**

(*Struktur der Lösungsmenge*) Für das lineare Gleichungssystem

$$\sum_{k=1}^n x_k \mathbf{a}_k = \mathbf{b} \quad (\mathbf{a}_k, \mathbf{b} \in \mathbb{R}^m) \quad (2.61)$$

gelte:

$$p := \text{Rang}\{\mathbf{a}_1, \dots, \mathbf{a}_n\} = \text{Rang}\{\mathbf{a}_1, \dots, \mathbf{a}_n, \mathbf{b}\}$$

(d.h. das Rangkriterium ist erfüllt). Dann folgt

- (a) Die Lösungsmenge ist eine  $(n - p)$ -dimensionale lineare Mannigfaltigkeit im  $\mathbb{R}^n$ .
- (b) Die Lösungsmenge des zu (2.61) gehörenden homogenen linearen Gleichungssystems

$$\sum_{k=1}^n x_k \mathbf{a}_k = \mathbf{0} \quad (2.62)$$

ist ein  $(n - p)$ -dimensionaler Unterraum von  $\mathbb{R}^n$ .

- (c) Ist  $\mathbf{x}_0 \in \mathbb{R}^n$  ein beliebiger, aber fest gewählter Lösungsvektor des inhomogenen linearen Gleichungssystems (2.61), so besteht die Lösungsmenge dieses Systems aus allen Vektoren der Form

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_h, \quad \text{wobei } \mathbf{x}_h \text{ Lösung des homogenen Systems (2.62) ist.}$$

### Beweis:

(a) folgt unmittelbar aus Satz 2.9 (b), Abschn. 2.2.4, wenn man (2.61) im Fall  $m \neq n$  durch Nullzeilen oder Nullspalten zu einem quadratischen System ergänzt. (b) ergibt sich entsprechend aus Satz 2.8 (b) in Abschn. 2.2.3.

Zu (c): Alle  $\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_h$  sind offensichtlich Lösungen von (2.61) (einsetzen!). Ist umgekehrt  $\mathbf{x}_1$  eine beliebige Lösung von (2.61), so ist  $\mathbf{x}_1 - \mathbf{x}_0 =: \mathbf{x}'_h$  zweifellos Lösung von (2.62), woraus  $\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{x}'_h$  folgt, d.h.  $\mathbf{x}_1$  hat die behauptete Form.  $\square$

**Bemerkung:** Die Sätze 2.9 und 2.10 haben hauptsächlich theoretische Bedeutung (als Beweishilfsmittel für weitere Sätze). Praktisch wird die Frage nach der Lösbarkeit und der Berechnung durch den Gaußschen Algorithmus angegangen. (Wir bemerken, dass im Fall sehr großer  $n = m$  auch andere — sogenannte iterative — Verfahren ins Bild kommen, s. Abschn. 3.8.4. Als »sehr groß« ist dabei etwa  $n = m \geq 100$  anzusehen).

### Übung 2.6\*

Gib die Lösungsmengen der folgenden beiden Systeme an:

$$\begin{array}{ll} \text{(a)} & \begin{array}{l} 3x_1 + 5x_2 - 7x_3 + x_4 = 0 \\ 8x_1 - 2x_2 + x_3 - 3x_4 = 1 \end{array} \\ \text{(b)} & \begin{array}{l} 9x_1 - 8x_2 = 1 \\ 5x_1 + 2x_2 = 7 \\ x_1 + 12x_2 = 13 \end{array} \end{array}$$

## 2.3 Algebraische Strukturen: Gruppen und Körper

Es werden zwei algebraische Strukturen — Gruppen und Körper — erläutert. Sie bilden das algebraische Fundament für die spätere allgemeine Vektorraumtheorie. Da unmittelbare Anwendungen im Ingenieurbereich jedoch selten sind, mag der anwendungsorientierte Leser den Abschnitt beim ersten Lesen überspringen und später bei Bedarf hier nachschlagen.

### 2.3.1 Einführung: Beispiel einer Gruppe

Wir beginnen mit einer Spielerei: Ein Plättchen von der Form eines gleichseitigen Dreiecks liegt auf dem Tisch. Seine Ecken sind mit 1, 2, 3 markiert. Die Lage des Dreiecks ist durch einen gezeichneten Umriss auf dem Tisch vermerkt, wobei die Ecken des Umrisses mit ①, ②, ③ gekennzeichnet sind (s. Fig. 2.5).



Wir fragen nun nach allen *Bewegungen* des Plättchens, nach deren Ausführung es wieder exakt auf dem Umriss liegt. Auch Umdrehen ist erlaubt, so dass die Unterseite des Plättchens nach oben kommt. Alle diese Bewegungen heißen *Kongruenzabbildungen* des gleichseitigen Dreiecks. Sie werden nach ihren Endlagen unterschieden.

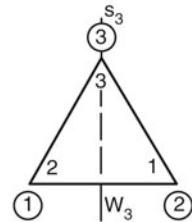
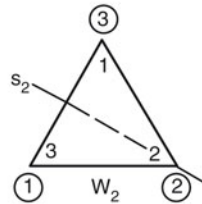
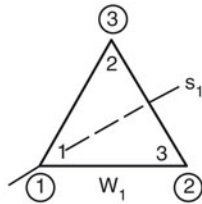
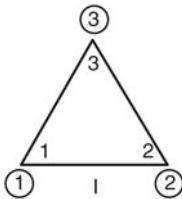
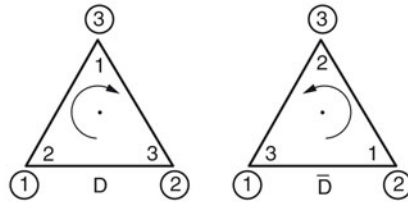


Fig. 2.5: Plättchen in Form eines gleichseitigen Dreiecks.

Fig. 2.6: Kongruenzabbildungen des gleichseitigen Dreiecks.

Dem Leser fallen sicher die folgenden fünf Bewegungen ein (Fig. 2.6):

$D$  = Drehung um  $120^\circ$  im Uhrzeigersinn<sup>12</sup>

$\overline{D}$  = Drehung um  $120^\circ$  gegen den Uhrzeigersinn

$W_1$  = Spiegelung an der (tischfesten) Seitenhalbierenden  $s_1$

$W_2$  = Spiegelung an der (tischfesten) Seitenhalbierenden  $s_2$

$W_3$  = Spiegelung an der (tischfesten) Seitenhalbierenden  $s_3$

( $W$  von »Wenden«). Der Vollständigkeit halber zählen wir noch die »triviale Bewegung« dazu:

$I$  = Identität, d.h. die Lage bleibt unverändert.

Mehr als diese sechs Kongruenzabbildungen des Dreiecks gibt es sicherlich nicht! (Warum?)

<sup>12</sup> eine Achse durch den Dreiecksmittelpunkt, die senkrecht auf dem Dreieck steht.

Wir beschreiben die Bewegungen des Dreiecks kürzer durch:

$$\begin{aligned} I &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, & D &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, & \overline{D} &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \\ W_1 &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, & W_2 &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, & W_3 &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}. \end{aligned} \quad (2.63)$$

Was bedeutet dies? — Nun, ganz einfach: Hier wird angegeben, wie die Ecken des Dreiecks bewegt werden! In  $D = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$  wird z.B. ausgedrückt, dass durch diese Drehung die Ecke 1 des Dreiecks in Ecke ③ des Umrisses überführt wird, ferner Ecke 2 in Ecke ① des Umrisses usw. Man beschreibt dies auch so:

$$D(1) = 3, \quad D(2) = 1, \quad D(3) = 2, \quad \overline{D}(1) = 2 \quad \text{usw.}$$

D.h.: Die Kongruenzabbildungen in (2.63) werden als Funktionen aufgefasst, die die Menge  $\mathbb{N}_3 := \{1, 2, 3\}$  umkehrbar eindeutig auf sich abbilden. Da in den unteren Zeilen der Schemata in (2.63) gerade die Permutationen der Zahlen 1, 2, 3 stehen (d.h. die verschiedenen Reihenfolgen dieser Zahlen), identifiziert man die Kongruenzabbildungen  $I, D, \overline{D}, W_1, W_2, W_3$  auch mit diesen Permutationen. — So weit, so gut! —

Jetzt führen wir zwei unserer Kongruenzabbildungen *nacheinander* aus. Es muss wieder eine der sechs Kongruenzabbildungen entstehen, denn andere gibt es nicht. Wird z.B. zuerst die Drehung  $D$  ausgeführt und dann die Spiegelung  $W_1$ , so ergibt sich insgesamt die Spiegelung  $W_3$  (wie man mit einem ausgeschnittenen Papierdreieck schnell überprüft). Wir beschreiben dies durch

$$W_3 = W_1 * D,$$

gesprochen: » $W_1$  nach  $D$ « oder » $W_1$  verknüpft mit  $D$ «, oder » $W_1$  mal  $D$ «.  $W_1 * D$  nennt man das *Produkt* aus  $W_1$  und  $D$  und bezeichnet die Verknüpfung  $*$  als Multiplikation.

Tabelle 2.2: Multiplikationstabelle der Kongruenzabbildungen des gleichseitigen Dreiecks.

*	$I$	$D$	$\overline{D}$	$W_1$	$W_2$	$W_3$
$I$	$I$	$D$	$\overline{D}$	$W_1$	$W_2$	$W_3$
$D$	$D$	$\overline{D}$	$I$	$W_2$	$W_3$	$W_1$
$\overline{D}$	$\overline{D}$	$I$	$D$	$W_3$	$W_1$	$W_2$
$W_1$	$W_1$	$W_3$	$W_2$	$I$	$\overline{D}$	$D$
$W_2$	$W_2$	$W_1$	$W_3$	$D$	$I$	$\overline{D}$
$W_3$	$W_3$	$W_2$	$W_1$	$\overline{D}$	$D$	$I$

Sämtliche möglichen Produkte zweier Kongruenzabbildungen des gleichseitigen Dreiecks sind in obenstehender Multiplikationstabelle verzeichnet (s. Tab. 2.2).

Sind  $A, B, H$  drei von unseren Kongruenzabbildungen für die  $H = A * B$  gilt, so folgt für jede Ecknummer  $x$  des Dreiecks  $H(x) = A(B(x))$ . Man sieht dies ein, wenn man den Weg einer

Ecke dabei verfolgt (überprüfe dies mit  $W_3 = W_1 * D$ ). Ersetzt man  $H$  durch  $A * B$  in der letzten Gleichung, so folgt

$$(A * B)(x) = A(B(x)). \quad (2.64)$$

für alle Ecknummern  $x$ . Für je drei beliebige Kongruenzabbildungen  $A, B, C$  unseres Dreiecks folgern wir damit

$$(I) \quad (A * B) * C = A * (B * C).$$

Denn für alle Ecken  $x$  gilt wegen (2.64)

$$\begin{aligned} ((A * B) * C)(x) &= (A * B)(C(x)) = A(B(C(x))) \\ &= A((B * C)(x)) = (A * (B * C))(x). \end{aligned}$$

Für jede beliebige Kongruenzabbildung  $A$  des gleichseitigen Dreiecks gilt ferner

$$(II) \quad I * A = A * I = A, \quad \text{und}$$

(III) Es gibt zu  $A$  eine Kongruenzabbildung  $X$  des Dreiecks mit

$$A * X = I = X * A.$$

Man bezeichnet  $X$  durch  $A^{-1}$  und nennt es *Inverses* zu  $A$ .

Man beachte, dass  $A * B = B * A$  nicht allgemein gilt, z.B.  $W_2 * W_1 = D$ ,  $W_1 * W_2 = \overline{D}$ .

Eine Menge mit einer Verknüpfung, in der die drei Gesetze (I), (II), (III) gelten, nennt man eine *Gruppe*. Wir haben es also hier mit der *Gruppe der Kongruenzabbildungen* eines gleichseitigen Dreiecks zu tun.

Im Folgenden wird der Gruppenbegriff allgemein gefasst. Er ist der wichtigste Grundbegriff der modernen Algebra.

### Übung 2.7:

Beschreibe die Gruppe der Kongruenzabbildungen eines Rechtecks (es soll sich um ein »echtes« Rechteck dabei handeln, also kein Quadrat). *Anmerkung* dazu: Nicht alle Permutationen der Ecken entsprechen Kongruenzabbildungen des Rechtecks, sondern nur einige. Die entstehende Gruppe heißt die *Kleinsche*<sup>13</sup> *Vierergruppe*.

## 2.3.2 Gruppen

### Definition 2.8:

Eine *Gruppe*  $(G, *)$  besteht aus einer Menge  $G$  und einer Verknüpfung<sup>14</sup>  $*$ , die jedem Paar  $(x, y)$  von Elementen  $x, y \in G$  genau ein Element  $x * y \in G$  zuordnet, wobei

<sup>13</sup> Felix Christian Klein (1849–1925), deutscher Mathematiker.

die folgenden drei Gesetze erfüllt sind:

- (I) Für alle  $x, y, z \in G$  gilt (Assoziativgesetz)

$$(x * y) * z = x * (y * z) .$$

- (II) Es gibt ein Element  $e \in G$  mit (neutrales Element)

$$x * e = x = e * x$$

für alle  $x \in G$ .  $e$  heißt *neutrales Element* in  $G$ .

- (III) Zu jedem Element  $x \in G$  gibt es ein Element (Inverses)  
 $y \in G$  mit

$$x * y = e = y * x .$$

$y$  heißt *Inverses* zu  $x$ , symbolisiert durch  $x^{-1}$ .

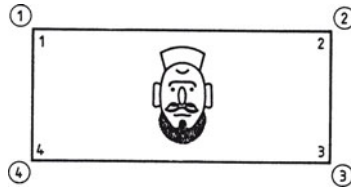


Fig. 2.7: Zu Übung 2.7, Kleinsche Vierergruppe (mit Klein)

Neutrales Element  $e$  und Inverses von  $x$  sind *eindeutig* bestimmt, denn würde auch  $e'$  der Bedingung (II) genügen so folgte  $e = e' * e = e'$ , und wären  $y$  und  $y'$  Inverse von  $x$ , so erhielte man  $y = e * y = (y' * x) * y = y' * (x * y) = y' * e = y'$ .

Wegen (I) schreibt man auch  $x * y * z$  statt  $(x * y) * z$ , da es ja gleichgültig ist, wie man klammert. Entsprechend  $x * y * z * w$  usw.

### Definition 2.9:

Gilt in einer Gruppe  $(G, *)$  zusätzlich

- (IV)  $x * y = y * x$  für alle  $x, y \in G$ , (Kommutativgesetz)

so liegt eine *kommutative* Gruppe vor, auch *abelsche*<sup>15</sup> Gruppe genannt.

<sup>14</sup> Unter einer (binären) *Verknüpfung* aus einer Menge  $M$  verstehen wir eine beliebige Vorschrift, die jedem Paar  $(x, y)$  mit  $x, y \in M$  genau ein Element aus  $M$  zuordnet.

Statt  $*$  werden auch andere Symbole verwandt, oder  $*$  wird ganz weggelassen. Dies ist in der Algebra üblich: Man schreibt einfach  $xy$  statt  $x * y$  und spricht von *multiplikativer Schreibweise* bzw. *multiplikativer Gruppe*  $G$ .

Bei abelschen Gruppen schreibt man häufig  $+$  statt  $*$ , wobei das neutrale Element mit  $0$  bezeichnet wird, und das Inverse zu  $x \in G$  mit  $-x$ . Man spricht von *additiver Schreibweise* bzw. *additiver Gruppe*. Die gerahmten Gleichungen in (I), (II), (III), (IV) bekommen dann die Form

$$\begin{aligned}(x + y) + z &= x + (y + z), & x + 0 &= x = 0 + x, \\ x + (-x) &= 0 = (-x) + x, & x + y &= y + x.\end{aligned}\tag{2.65}$$

Die *Subtraktion* in einer additiven abelschen Gruppe wird folgendermaßen definiert:

$$x - y := x + (-y).$$

Eine *Gruppe*  $(G, *)$  heißt *endlich* wenn  $G$  eine *endliche Menge* ist. Die Anzahl der Elemente von  $G$  heißt die *Ordnung* der Gruppe  $(G, *)$ . Eine nichtendliche Gruppe heißt *unendliche Gruppe*.

Statt von der Gruppe  $(G, *)$  oder  $(G, \cdot)$  oder  $(G, +)$  spricht man auch einfach von der *Gruppe*  $G$ , wenn aus dem Zusammenhang klar ist, mit welcher Verknüpfung gearbeitet wird. Dadurch wird sprachliche Überladung vermieden.

## Beispiele für Gruppen

### Beispiel 2.4:

Die Menge  $\mathbb{Z}$  (der ganzen Zahlen),  $\mathbb{Q}$  (der rationalen Zahlen),  $\mathbb{R}$  (der reellen Zahlen) und  $\mathbb{C}$  (der komplexen Zahlen), sowie  $\mathbb{R}^n$  und  $\mathbb{C}^n$  bilden *additive abelsche Gruppen*:

$$(\mathbb{Z}, +), \quad (\mathbb{Q}, +), \quad (\mathbb{R}, +), \quad (\mathbb{C}, +), \quad (\mathbb{R}^n, +), \quad (\mathbb{C}^n, +).$$

### Beispiel 2.5:

Streicht man aus  $\mathbb{Q}$ ,  $\mathbb{R}$  und  $\mathbb{C}$  jeweils die  $0$  heraus, so liefern die entstehenden Mengen  $\mathbb{Q}'$ ,  $\mathbb{R}'$ ,  $\mathbb{C}'$  *multiplikative abelsche Gruppen*

$$(\mathbb{Q}', \cdot), \quad (\mathbb{R}', \cdot), \quad (\mathbb{C}', \cdot).$$

### Beispiel 2.6:

$G = \{-1, 1\}$  bildet eine multiplikative abelsche Gruppe aus 2 Elementen.

### Beispiel 2.7:

Die Menge  $G$  der komplexen Zahlen mit Absolutbetrag 1 stellt eine multiplikative abelsche Gruppe dar.

---

15 Nach dem norwegischen Mathematiker Niels Henrik Abel (1802–1829).

**Beispiel 2.8:**

Die Menge aller Kongruenzabbildungen einer ebenen oder räumlichen Figur auf sich selbst ergibt eine Gruppe bezüglich der Hintereinanderausführung der Kongruenzabbildungen. Gruppen dieser Art sind meistens nicht kommutativ (s. Beispiel in Abschn. 2.3.1).

**Beispiel 2.9:**

(Gruppe der bijektiven Abbildungen einer Menge auf sich) Wir haben es hier mit dem *Paradebeispiel* für Gruppen zu tun. Zunächst sei kurz wiederholt<sup>16</sup>: Eine Abbildung  $f : M \rightarrow N$  heißt *injektiv* (eindeutig), wenn aus  $x_1, x_2 \in M$  mit  $x_1 \neq x_2$  folgt:  $f(x_1) \neq f(x_2)$ , kurz: »Verschiedene Urbilder haben verschiedene Bilder«.  $f : M \rightarrow N$  heißt *surjektiv* (Abbildung von  $M$  auf  $N$ ), wenn zu jedem  $y \in N$  ein  $x \in M$  existiert mit  $f(x) = y$ , m.a.W.: »Der Bildbereich  $N$  wird durch die Werte  $f(x)$  ausgeschöpft.« (d.h.  $f(M) = N$ ).  $f : M \rightarrow N$  heißt *bijektiv* (umkehrbar eindeutig), wenn  $f$  injektiv und surjektiv ist.

Wir betrachten die Menge  $G$  aller bijektiven Abbildungen  $f : M \rightarrow M$  einer Menge  $M$  auf sich. Zwei solche Abbildungen  $f : M \rightarrow M$  und  $g : M \rightarrow M$  werden durch ihre Hintereinanderausführung (Komposition) verknüpft: Man definiert  $f \circ g$  durch

$$(f \circ g)(x) := f(g(x)) \quad \text{für alle } x \in M.$$

$(f \circ g) : M \rightarrow M$  ist also wieder eine bijektive Abbildung von  $M$  auf sich. Es ist sonnenklar, dass  $(G, \circ)$  eine Gruppe ist, denn es gilt für alle  $f, g, h \in G$ :

$$(I) \quad (f \circ g) \circ h = f \circ (g \circ h) \quad (2.66)$$

wegen

$$\begin{aligned} ((f \circ g) \circ h)(x) &= (f \circ g)(h(x)) = f(g(h(x))) = (f \circ (g \circ h))(x) \\ &= (f \circ (g \circ h))(x) \quad \text{für alle } x \in M. \end{aligned}$$

(II) Die Abbildung  $I : M \rightarrow M$  mit  $I(x) = x$  für alle  $x \in M$  — *Identität* genannt — erfüllt

$$I \circ f = f = f \circ I \quad \text{für alle } f \in G. \quad (2.67)$$

(III) Zu jeder bijektiven Abbildung  $f$  existiert die Umkehrabbildung  $f^{-1}$ , die folgendes erfüllt

$$f \circ f^{-1} = I = f^{-1} \circ f. \quad (2.68)$$

Die Gruppengesetze sind somit erfüllt. Besitzt  $M$  nur ein oder zwei Elemente, so ist die Gruppe  $(G, \circ)$  kommutativ (man prüfe dies nach!), hat  $M$  drei oder mehr Elemente, evtl. sogar unendlich viele, so ist sie nicht kommutativ (s. Abschn. 2.3.1: Die dortige Gruppe kann als Gruppe der bijektiven Abbildungen auf der Eckpunktmenge  $M = \{1, 2, 3\}$  aufgefasst werden. Sie ist nicht kommutativ).

Die beschriebene Gruppe  $G$  aller bijektiven Abbildungen einer Menge  $M$  auf sich wird *Per-*

---

<sup>16</sup> s. Burg/Haf/Wille (Analysis) [27]

*mutationsgruppe von  $M$*  genannt, symbolisiert durch

$$\text{Perm } M := G.$$

Ist  $M$  der Zahlenabschnitt  $\mathbb{N}_n := \{1, 2, \dots, n\}$ , so schreibt man

$$S_n := \text{Perm } \mathbb{N}_n.$$

Diese endlichen Permutationsgruppen werden wir im nächsten Abschnitt genauer betrachten.

**Definition 2.10:**

Es sei  $(G, *)$  eine Gruppe und  $U$  eine nichtleere Teilmenge von  $G$ . Man nennt  $U$  (bzgl.  $*$ ) eine *Untergruppe* von  $(G, *)$ <sup>17</sup>, wenn folgendes gilt

$$(U. I) \quad \text{Wenn } x, y \in U, \text{ so auch } x * y \in U.$$

$$(U. II) \quad \text{Wenn } x \in U, \text{ so auch } x^{-1} \in U.$$

(In additiver Schreibweise:  $x, y \in U \Rightarrow x + y \in U$  und  $x \in U \Rightarrow -x \in U$ .)

Es folgt unmittelbar, dass dann auch gilt

$$(U. III) \quad e \in U, \quad \text{wobei } e \text{ das neutrale Element von } G \text{ ist,}$$

denn: da  $U \neq \emptyset$  so existiert ein  $x \in U$ , damit auch  $x^{-1} \in U$  (nach (U. II)) und somit  $x * x^{-1} = e$  in  $U$  (nach (U. I)).

Die Untergruppenbeziehung beschreiben wir einfach durch

$$(U, *) \subset (G, *).$$

$U$  bildet mit der Verknüpfung  $*$ , eingeschränkt auf die Paare  $(x, y)$  mit  $x, y \in U$ , natürlich wieder eine Gruppe, wie der Name Untergruppe schon sagt.

$G$  und  $\{e\}$  sind Untergruppen von  $(G, *)$ . Sie heißen die *volle* und die *triviale* Untergruppe. Alle anderen Untergruppen von  $G$  werden *echte* Untergruppen genannt.

**Definition 2.11:**

Eine Untergruppe  $U$  von  $(G, *)$  heißt ein *Normalteiler*, wenn für jedes  $u \in U$  und jedes  $a \in G$  gilt:

$$a * u * a^{-1} \in U.$$

(Die Bedeutung der Normalteiler wird später bei Gruppenhomomorphismen klar.)

In kommutativen Gruppen sind alle Untergruppen offenbar auch Normalteiler.

**Beispiel 2.10:**

$$(\mathbb{Z}, +) \subset (\mathbb{Q}, +) \subset (\mathbb{R}, +) \subset (\mathbb{C}, +).$$

---

<sup>17</sup> Oder kurz: Untergruppe von  $G$ .

**Beispiel 2.11:**

$$(\mathbb{Q}', \cdot) \subset (\mathbb{R}', \cdot) \subset (\mathbb{C}', \cdot).$$

**Beispiel 2.12:**

Die in Abschn. 2.3.1 beschriebene Gruppe  $G = \{I, D, \overline{D}, W_1, W_2, W_3\}$  hat die echten Untergruppen  $U_0 = \{I, D, \overline{D}\}$ ,  $U_1 = \{I, W_1\}$ ,  $U_2 = \{I, W_2\}$ ,  $U_3 = \{I, W_3\}$ . Man prüfe an Hand der Gruppentafel (Tab. 2.2) nach:  $U_0$  ist ein Normalteiler,  $U_1, U_2, U_3$  sind es nicht.

**Übung 2.8\***

Welche echten Untergruppen hat die Kleinsche Vierergruppe aus Übung 2.7?

**2.3.3 Endliche Permutationsgruppen**

Wir sehen uns in diesem Abschnitt die Permutationsgruppen  $S_n$  genauer an.  $S_n$  war definiert als die Menge aller bijektiven Abbildungen der Menge  $\mathbb{N}_n = \{1, 2, \dots, n\}$  auf sich. Die Permutationsgruppen  $S_n$  — auch *Symmetriegruppen* genannt — spielen in vielen Bereichen der Mathematik eine Rolle, z.B. in der Kombinatorik, Wahrscheinlichkeitsrechnung, Algebra der Polynomgleichungen und der Behandlung der Determinanten (s. Abschn. 3.4).

Ist  $f: \mathbb{N}_n \rightarrow \mathbb{N}_n$  eine bijektive Abbildung, so wollen wir Funktionswerte  $f(i)$  auch durch  $k_i$  bezeichnen:

$$f(i) =: k_i.$$

Die Abbildung  $f$  kann damit vollständig durch das Schema

$$f = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ k_1 & k_2 & k_3 & \cdots & k_n \end{pmatrix}$$

beschrieben werden, wobei  $(k_1, k_2, \dots, k_n)$  eine *Permutation* von  $(1, 2, \dots, n)$  ist (daher der Name »Permutationsgruppe«). Die Permutationen von  $(1, 2, \dots, n)$  (d.h. die  $n$ -Tupel  $(k_1, \dots, k_n)$ , bestehend aus allen Zahlen  $1, 2, \dots, n$ ) sind also den Abbildungen  $f \in S_n$  umkehrbar eindeutig zugeordnet, so dass wir sie miteinander identifizieren können:

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ k_1 & k_2 & \cdots & k_n \end{pmatrix} = (k_1, \dots, k_n).$$

Die Abbildungen  $f \in S_n$  werden wir also auch als Permutationen bezeichnen. Die *Ordnung* von  $S_n$  ist  $n!$  (s. Burg/Haf/Wille (Analysis) [27]). Die Verknüpfung von  $f, g \in S_n$  bezeichnen wir einfach durch  $fg$ . Als explizites Beispiel siehe die Gruppe  $S_3$  in Abschn. 2.3.1.

**Definition 2.12:**

Als *Transposition*  $t$  bezeichnen wir jede Permutation aus  $S_n$ , die genau zwei Zahlen vertauscht und alle übrigen fest lässt, also

$$t(i) = j, \quad t(j) = i \quad \text{für zwei Zahlen } i, j \in \mathbb{N}_n, i \neq j,$$

und  $t(k) = k$  für alle  $k \neq i$  und  $k \neq j$ .



**Satz 2.12:**

Jede Permutation aus  $S_n$  lässt sich als Produkt von endlich vielen Transpositionen darstellen.

**Beweis:**

Zur Erzeugung von  $(k_1, k_2, \dots, k_n)$  aus  $(1, 2, \dots, n)$  bringt man zuerst  $k_1$  an die erste Stelle, durch Vertauschen von 1 mit  $k_1$  (falls nicht  $k_1 = 1$ ). Dann wird 2 und  $k_2$  vertauscht, falls nicht  $k_2 = 2$ , usw. Diese Folge von Vertauschungen lässt sich als Produkt von Transpositionen schreiben.  $\square$

Für weitere Überlegungen unterscheiden wir *gerade* und *ungerade* Permutationen:

**Definition 2.13:**

Als *Fehlstand* einer Permutation  $(k_1, k_2, \dots, k_n)$  bezeichnet man ein Paar  $k_i, k_j$  mit

$$i < j, \quad \text{aber} \quad k_i > k_j.$$

Eine Permutation heißt *gerade*, wenn die Anzahl ihrer Fehlstände gerade ist. Andernfalls heißt sie *ungerade*. Z.B.: Die Permutation  $(2, 3, 1)$  hat die Fehlstände  $(2, 1)$ ,  $(3, 1)$ , und keine weiteren.  $(2, 3, 1)$  ist also eine gerade Permutation.

Man vereinbart weiter:

$$\text{sgn}(p) := \begin{cases} 1 & \text{falls } p \text{ gerade Permutation} \\ -1 & \text{falls } p \text{ ungerade Permutation} \end{cases}$$

( $\text{sgn}(p)$  wird gesprochen als »Signum  $p$ «).

**Satz 2.13:**

Eine Permutation  $p$  ist genau dann *gerade*, wenn bei jeder Darstellung

$$p = t_1 t_2 \dots t_m, \quad t_i \text{ Transpositionen,}$$

die Anzahl  $m$  der verwendeten Transpositionen gerade ist. Für ungerade Permutationen gilt Entsprechendes.

**Beweis:**

Eine Transposition wird als Vertauschung zweier Zahlen aufgefasst. Wir können also die Permutation  $p = t_1 t_2 \dots t_m e$  (mit dem neutralen Element  $e = (1, 2, \dots, n) = \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}$ ) aus  $S_n$ ) so auffassen, als ob  $p$  aus  $(1, 2, \dots, n)$  durch sukzessives Vertauschen je zweier Zahlen entsteht, und zwar entsprechend den Transpositionen  $t_m, t_{m-1}, \dots, t_1$ , die in dieser Reihenfolge nacheinander angewandt werden. Jede Vertauschung zweier Zahlen erzeugt aber genau einen neuen Fehlstand, oder vernichtet genau einen. Ist  $m$  also gerade, so kommt am Ende eine gerade Anzahl von Fehlständen heraus, anderenfalls eine ungerade.  $\square$

Mit Satz 2.13 beweist man leicht

$$\operatorname{sgn}(p_1 * p_2) = \operatorname{sgn}(p_1) \operatorname{sgn}(p_2), \quad \operatorname{sgn}(p^{-1}) = \operatorname{sgn} p.$$

### Übung 2.9:

(a) Welche der folgenden Permutationen sind gerade und welche ungerade

$$(3, 1, 2), \quad (3, 2, 1), \quad (5, 7, 1, 2, 6, 3, 4)?$$

(b) Schreibe diese Permutationen als Produkte aus Transpositionen.

### 2.3.4 Homomorphismen, Nebenklassen

Dieser Abschnitt rundet das »Einmaleins« über Gruppen ab. Unmittelbare Anwendungen auf die Ingenieurpraxis kommen dabei nicht vor (sehr wohl aber mittelbare für die »Darstellungstheorie von Gruppen« und damit für Kristallographie und Quantenmechanik). Der anwendungsorientierte Leser mag sich zunächst mit »Querlesen« begnügen.

Im vorliegenden Abschnitt schreiben wir i. Allg. Gruppen multiplikativ, d.h. Produkte werden in der Form  $xy$  oder  $x \cdot y$  ausgedrückt. Gruppen  $(G, \cdot)$  werden auch einfach durch  $G$  beschrieben.

#### Definition 2.14:

Es seien  $G, H$  zwei Gruppen. Ein *Homomorphismus* von  $G$  nach  $H$  ist eine Abbildung  $f : G \rightarrow H$ , die folgendes erfüllt:

$$f(xy) = f(x)f(y) \quad \text{für alle } x, y \in G.$$

Diese Gleichung heißt *Homomorphiebedingung*.

Ist  $f$  zusätzlich bijektiv, so heißt  $f$  ein *Isomorphismus* von  $G$  auf  $H$ . Im Falle  $G = H$  nennt man den Isomorphismus  $f : G \rightarrow G$  *Automorphismus*.

#### Beispiel 2.13:

Durch  $f(x) = e^x$  wird ein Isomorphismus von  $(\mathbb{R}, +)$  auf  $(\mathbb{R}^+, \cdot)$  vermittelt, wobei  $\mathbb{R}^+ := (0, \infty)$ . Der Leser überprüfe dies.

#### Beispiel 2.14:

Für alle  $z \in \mathbb{C}$  sei  $f(z) = |z|$ . Mit  $\mathbb{C}' = \mathbb{C} \setminus \{0\}$ , und  $\mathbb{R}^+ = (0, \infty)$  folgt:  $f : \mathbb{C}' \rightarrow \mathbb{R}^+$  ist ein Homomorphismus bezüglich  $\cdot$  von  $(\mathbb{C}, \cdot)$  auf  $(\mathbb{R}^+, \cdot)$ .

#### Beispiel 2.15:

Für unser Beispiel  $G = \{I, D, \overline{D}, W_1, W_2, W_3\}$  aus Abschn. 2.3.1 definieren wir  $1 = f(I) = f(D) = f(\overline{D})$  und  $-1 = f(W_i)$  für alle  $i = 1, 2, 3$ . Damit ist  $f : G \rightarrow \{-1, 1\}$  ein Homomorphismus, wobei  $\{-1, 1\}$  bzgl.  $\cdot$  eine Gruppe ist.

Als *Kern* des Gruppen-Homomorphismus  $f : G \rightarrow H$  — kurz *Kern*  $f$  — bezeichnet man die Menge aller  $x \in G$  mit  $f(x) = e'$ , wobei  $e'$  das neutrale Element von  $H$  ist. Das *Bild* des

Homomorphismus  $f : G \rightarrow H$  ist einfach die Menge  $f(G) = \{f(x) \mid x \in G\}$  (also der Wertebereich von  $f$ .)

**Satz 2.14:**

Eigenschaften von Homomorphismen: Es sei  $f : G \rightarrow H$  ein Homomorphismus der Gruppe  $G$  in die Gruppe  $H$ . Ferner seien  $e \in G$  und  $e' \in H$  die neutralen Elemente. Damit gilt

- (a)  $f(e) = e'$ .
- (b)  $f(x^{-1}) = (f(x))^{-1}$  für alle  $x \in G$ .
- (c) Das Bild von  $f$  ist eine Untergruppe von  $H$ .
- (d) Der Kern von  $f$  ist ein Normalteiler in  $G$ .
- (e) Ist  $g : H \rightarrow K$  ein weiterer Gruppen-Homomorphismus, so ist auch die Komposition  $g \circ f : G \rightarrow K$  ein Gruppen-Homomorphismus.
- (f) Ist  $f : G \rightarrow H$  ein Isomorphismus, so ist auch die Umkehrabbildung  $f^{-1} : H \rightarrow G$  ein Isomorphismus.

Die einfachen Beweise bleiben dem Leser überlassen (s. z.B. [86], S. 38–40).

**Definition 2.15:**

Es sei  $U$  eine Untergruppe der Gruppe  $G$ . Dann heißt jede Menge

$$aU := \{ax \mid x \in U\}, \quad a \in G$$

eine *Linksnebenklasse* von  $U$ , und entsprechend

$$Ua := \{xa \mid x \in U\}$$

eine *Rechtsnebenklasse* von  $U$ . Der gemeinsame Begriff für Links- und Rechtsnebenklasse ist — na was wohl? — *Nebenklasse*.

**Folgerung 2.6:**

Die Gruppe  $G$  wird in Linksnebenklassen von  $U$  zerlegt, d.h. verschiedene Linksnebenklassen sind elementfremd und  $G$  ist die Vereinigung aller Linksnebenklassen. (Für Rechtsnebenklassen gilt Entsprechendes.)

**Beweis:**

Die letzte Aussage ist offensichtlich richtig, da jedes  $x \in G$  ja in »seiner« Nebenklasse  $xU$  liegt. — Es seien  $aU, bU$  zwei verschiedene Linksnebenklassen. Wir können (ohne Beschränkung der Allgemeinheit) annehmen, dass  $au \in aU$  nicht in  $bU$  liegt. Hätten  $aU$  und  $bU$  aber ein gemeinsames Element  $z$ , hätte es die Form  $z = ax = by$  mit passendem  $x, y \in U$ . Es folgte  $a = byx^{-1}$  (durch Rechtsmultiplikation mit  $x^{-1}$ ), also  $au = (byx^{-1})u = b(yx^{-1}u) \in bU$ , im Widerspruch zu  $au \notin bU$ . Daher gilt  $aU \cap bU = \emptyset$ .  $\square$

**Folgerung 2.7:**

Ist  $G$  eine endliche Gruppe, so haben alle Nebenklassen einer Untergruppe  $U$  von  $G$  gleich viele Elemente.

**Beweis:**

Mit  $U = \{a_1, a_2, \dots, a_n\}$  ist  $aU = \{aa_1, aa_2, \dots, aa_n\}$ , wobei  $aa_i \neq aa_j$  für  $j \neq i$  ist, denn aus  $aa_i = aa_j$  folgte nach Linksmultiplikation mit  $a^{-1} : a_i = a_j$ . Entsprechend  $Ua = \{a_1a, \dots, a_na\}$ .  $\square$

**Satz 2.15:**

(von Lagrange) Es sei  $G$  eine endliche Gruppe. Es folgt: Die Ordnung jeder Untergruppe  $U$  von  $G$  ist Teiler der Gruppenordnung.

**Beweis:**

$G$  wird in »gleich große« Linksnebenklassen von  $U$  zerlegt, wobei  $U = eU$  selbst eine Nebenklasse ist. Damit folgt die Behauptung.  $\square$

**Bemerkung:** Dieser Satz ist eine große Hilfe beim Aufsuchen aller Untergruppen einer endlichen Gruppe, denn alle Teilmengen, deren Elementanzahlen die Gruppenordnung nicht teilen, entfallen schon als Kandidaten für Untergruppen.

Eine Untergruppe  $U$  einer Gruppe  $G$  ist genau dann ein *Normalteiler*, wenn  $aU = Ua$  für alle  $a \in G$  gilt, d.h. wenn *jede Linksnebenklasse gleich der entsprechenden Rechtsnebenklasse* ist. Denn  $aU = Ua$  bedeutet: Für jedes  $u \in U$  gibt es ein  $v \in U$  mit  $au = va$ , d.h. für jedes  $u$  gilt  $aua^{-1} = v \in U$ , d.h.  $U$  ist Normalteiler.

**Definition 2.16:**

Ist  $N$  ein Normalteiler einer Gruppe  $G$ , so bildet die Menge aller Nebenklassen von  $N$  eine Gruppe bzgl. der Verknüpfung

$$aN \cdot bN = abN. \quad (2.69)$$

Diese Gruppe heißt *Faktorgruppe von  $G$  nach  $N$*  und wird mit  $G/N$  symbolisiert. (Es ist  $N$  neutrales Element von  $G/N$  und  $(aN)^{-1} = a^{-1}N$ .)

**Folgerung 2.8:**

Durch  $f(x) = xN$  wird ein Homomorphismus von  $G$  auf  $G/N$  vermittelt. Dabei ist  $N$  der Kern von  $f$ .

Jeder Normalteiler einer Gruppe  $G$  ist also Kern eines Homomorphismus und jeder Homomorphismus auf  $G$  hat einen Normalteiler als Kern. Man kann sagen: Die Normalteiler beschreiben alle Homomorphismen. Der folgende Satz fasst dies noch deutlicher zusammen.

**Satz 2.16:**

(Homomorphiesatz für Gruppen) Für jeden Gruppenhomomorphismus  $f : G \rightarrow H$  mit Kern  $N$  gilt: Die Faktorgruppe  $G/N$  ist isomorph zu  $f(G) \subset H$  bezüglich der Abbildung  $\alpha : G/N \rightarrow f(G)$ , definiert durch  $\alpha(aN) = f(a)$ .

**Beweis:**

Es ist hauptsächlich zu zeigen, dass  $f$  alle Elemente aus  $aN$  auf  $f(a)$  abbildet. Man erkennt dies sofort aus  $f(au) = f(a)f(u) = f(a)e' = f(a)$  für alle  $u \in N$  ( $e'$  neutral in  $H$ ). Damit ist  $\alpha$  sinnvoll erklärt, und der Nachweis, dass  $\alpha$  ein Isomorphismus ist, elementar.  $\square$

**Übung 2.10:**

Gib die Faktorgruppe  $G/N$  an für  $G = \{I, D, \overline{D}, W_1, W_2, W_3\}$  und  $N = \{I, D, \overline{D}\}$ , s. Abschn. 2.3.1.

**Übung 2.11:**

Durch  $\text{sgn } p$  ( $p$  Permutation aus  $S_n$ ) ist eine Abbildung  $\text{sgn} : S_n \rightarrow \{-1, 1\}$  gegeben. Zeige, dass sie ein *Homomorphismus* ist ( $G = \{-1, 1\}$  ist eine Gruppe bzgl. der Multiplikation).

*Anleitung:* Stelle zum Nachweis der Homomorphiebedingung die Permutationen als Produkte von Transpositionen dar und verwende Satz 2.13 aus Abschn. 2.3.3.

**2.3.5 Körper****Definition 2.17:**

Ein algebraischer Körper  $(\mathbb{K}, +, \cdot)$  — kurz Körper<sup>18</sup> — besteht aus einer Menge  $\mathbb{K}$  und zwei Verknüpfungen  $+$  (Addition) und  $\cdot$  (Multiplikation), die jedem Paar  $(a, b)$ ,  $a, b \in \mathbb{K}$ , jeweils genau ein Element  $a + b$  bzw.  $a \cdot b$  zuordnen.  $a + b$  heißt die *Summe* von  $a, b$ ,  $a \cdot b$  *Produkt* von  $a, b$ . Dabei müssen folgende Gesetze erfüllt sein.

Für alle  $a, b, c \in \mathbb{K}$  gilt:

(A1)  $a + (b + c) = (a + b) + c.$

(A2)  $a + b = b + a.$

(A3) Es gibt ein Element  $0 \in \mathbb{K}$  mit  $a + 0 = a$  für alle  $a \in \mathbb{K}.$

(A4) Zu jedem  $a \in \mathbb{K}$  gibt es genau ein  $x \in \mathbb{K}$  mit  $a + x = 0$ . Man schreibt dafür  $x =: -a.$

(M1)  $a \cdot (b \cdot c) = (a \cdot b) \cdot c.$

(M2)  $a \cdot b = b \cdot a.$

(M3) Es gibt ein Element  $1 \in \mathbb{K}$  mit  $a \cdot 1 = a$  für alle  $a \in \mathbb{K}.$

(M4) Zu jedem  $a \neq 0$  aus  $\mathbb{K}$  gibt es genau ein  $y \in \mathbb{K}$  mit  $ay = 1$ . Man schreibt dafür  $y =: a^{-1}$  oder  $y = \frac{1}{a}.$

(D1)  $a \cdot (b + c) = a \cdot b + a \cdot c.$

(D2)  $0 \neq 1.$

<sup>18</sup> Da Beiwort *algebraisch* soll zur Unterscheidung von *physikalischen Körpern* im Raum dienen. Das Beiwort wird weggelassen, wenn Irrtümer ausgeschlossen sind.

**Bemerkung:** Die Gesetze (A1) und (M1) heißen *Assoziativgesetze* der *Addition* bzw. *Multiplikation*. (A2) und (M2) werden entsprechend *Kommutativgesetze* genannt, während (D1) *Distributivgesetz* heißt. Der Multiplikationspunkt  $\cdot$  wird auch weggelassen:  $a \cdot b = ab$ .

Die Assoziativgesetze (A1), (M1) bedeuten, wie bei Gruppen, dass es gleichgültig ist, wie man Klammern setzt. Wir lassen sie daher auch einfach weg, sowohl bei dreifachen Summen und Produkten, wie auch bei längeren. Die *Körperaxiome* (A1) bis (D2) zeigen, dass  $(\mathbb{K}, +)$  und  $(\mathbb{K}', \cdot)$  (mit  $\mathbb{K}' = \mathbb{K} \setminus \{0\}$ ) abelsche Gruppen sind; 0 und 1 sind damit eindeutig bestimmt. Die Gruppen sind durch die Gesetze (D1) und (D2) verknüpft.

*Subtraktion* bzw. *Division* werden folgendermaßen definiert

$$a - b := a + (-b); \quad a : b := \frac{a}{b} := a \cdot \frac{1}{b} \quad (b \neq 0). \quad (2.70)$$

Es folgen aus den Axiomen (A1), ..., (D2) alle Regeln der Bruchrechnung, wie in Burg/Haf/Wille (Analysis) [27]:

$$\frac{a}{c} + \frac{b}{d} = \frac{ad + bc}{cd}, \quad \frac{a}{c} \cdot \frac{b}{d} = \frac{ab}{cd}, \quad \frac{a}{c} : \frac{b}{d} = \frac{ad}{cb}, \quad (2.71)$$

$$a(-b) = -ab, \quad (-a)(-b) = ab, \quad (2.72)$$

$$ax = b \quad (a \neq 0) \Leftrightarrow x = \frac{b}{a}; \quad a \cdot 0 = 0. \quad (2.73)$$

Man definiert die *Potenzen* mit natürlichen Zahlen  $n$ :

$$a^n := \underbrace{a \cdot a \cdot a \cdot \dots \cdot a}_{n \text{ Faktoren}}$$

und im Fall  $a \neq 0$ :

$$a^0 = 1, \quad a^{-n} = \frac{1}{a^n} \quad (n \in \mathbb{N}),$$

womit die *Potenzen* für alle ganzen Hochzahlen erklärt sind.

Dies alles hat niemand anders erwartet! Doch nun zu den *Beispielen*:

Körper sind  $\mathbb{Q}$ ,  $\mathbb{R}$  und  $\mathbb{C}$  bezüglich der üblichen Addition und Multiplikation. Diese drei Körper sind unsere Hauptbeispiele.

Doch gibt es auch endliche Körper (bei denen  $\mathbb{K}$  eine endliche Menge ist). Sie sind für die Ingenieurmathematik aber von geringem Interesse.

### Beispiel 2.16:

Der einfachste Körper ist  $\mathbb{K} = \{0, 1\}$  mit  $0 + 0 = 0$ ,  $0 + 1 = 1 + 0 = 1$ ,  $1 + 1 = 0$  und  $0 \cdot 0 = 0 \cdot 1 = 0$ ,  $1 \cdot 1 = 1$ . Ziemlich trivial!

### Beispiel 2.17:

Ist  $p$  eine Primzahl, so kann man mit den Zahlen der Menge  $\mathbb{Z}_p = \{0, 1, \dots, p-1\}$  »zyklisch modulo  $p$ « rechnen. Das geht so: Für  $a, b \in \mathbb{Z}_p$  definiert man  $a \oplus b := a + b$ , falls  $a + b < p$ ,

und  $a \oplus b := a + b - p$ , falls  $a + b \geq p$ . Entsprechend:  $a \odot b := a \cdot b - mp$ , wobei die ganze Zahl  $m \geq 0$  so gewählt wird, dass  $a \cdot b - mp \in \mathbb{Z}_p$ . Damit ist  $(\mathbb{Z}_p, \oplus, \odot)$  ein Körper. (Der Beweis wird hier aus Platzgründen weggelassen.) Man nennt  $\mathbb{Z}_p$  den *Restklassenkörper modulo  $p$* .

**Bemerkung:** Eine Menge  $\mathbb{K}$  mit zwei Verknüpfungen  $+$  und  $\cdot$ , die nur die Gesetze (A1), (A2), (A3), (A4), (M1) und (D1) zuzüglich (D1)':  $(b+c) \cdot a = b \cdot a + c \cdot a$  erfüllen, heißt ein *Ring*. Meist wird auch noch (M3) gefordert, also die Existenz einer  $1 \in \mathbb{K}$  mit  $a \cdot 1 = 1 \cdot a = a$  für alle  $a \in \mathbb{K}$ . Man spricht dann von einem *Ring mit Eins* (-Element). Gilt zusätzlich (M2):  $a \cdot b = b \cdot a$ , so liegt ein *kommutativer Ring* vor. Schließlich heißt ein kommutativer Ring mit 1 ein *Integritätsbereich*, wenn aus  $x \cdot y = 0$  folgt:  $x = 0$  oder  $y = 0$ .

Beispiele für Ringe sind  $(\mathbb{Z}, +, \cdot)$ , der »Ring der ganzen Zahlen« (er ist ein Integritätsbereich), ferner die Menge der *Polynome* bzgl.  $+$ ,  $\cdot$ , und der »Ring der quadratischen  $n$ -reihigen Matrizen«, s. Abschn. 3. Da für die Ingenieurpraxis die Ringtheorie von geringer Bedeutung ist, brechen wir die Erörterung hier ab. Ein einfacher Einstieg ist in [86] zu finden.

### Übung 2.12\*

Zeige, dass  $\mathbb{Z}_3 = \{0, 1, 2\}$  ein Körper wird, wenn man in ihm zyklisch modulo 3 rechnet (s. Beisp. 2.17).

## 2.4 Vektorräume über beliebigen Körpern

Nachdem wir die speziellen Vektorräume  $\mathbb{R}^n$  und  $\mathbb{C}^n$  in Abschn. 2.1 kennen und lieben gelernt haben, werden hier die Vektorräume allgemein als algebraische Strukturen eingeführt,  $\mathbb{R}^n$  und  $\mathbb{C}^n$  erscheinen als Spezialfälle dieser Struktur. Als wichtige weitere Beispiele werden Funktionenräume genannt.<sup>19</sup> Sie spielen bei Differential- und Integralgleichungen eine grundlegende Rolle.

Der praxisorientierte Leser, wie auch der Anfänger, kann diesen Abschnitt zunächst überspringen und mit Abschn. 3 fortfahren. Er mag bei Bedarf hierher zurückkehren und an einem Abend, an dem im Fernsehen nichts Gescheites gesendet wird, diesen Abschnitt genießen.

### 2.4.1 Definition und Grundeigenschaften

Im Folgenden sei  $(\mathbb{K}, +, \cdot)$  ein beliebiger *algebraischer Körper* — kurz *Körper*  $\mathbb{K}$  genannt (s. Abschn. 2.3.5). Man kann sich einfach  $\mathbb{R}$  an Stelle von  $\mathbb{K}$  vorstellen, wenn einem die algebraische Struktur »Körper« noch ungewohnt ist. Die beiden Körper  $\mathbb{R}$  und  $\mathbb{C}$  sind für die Anwendungen sowieso am wichtigsten.

Wir definieren *Vektorräume* über  $\mathbb{K}$ , indem wir uns an  $\mathbb{R}^n$  orientieren.

#### Definition 2.18:

Ein *Vektorraum über einem Körper*  $\mathbb{K}$  besteht aus einer nichtleeren Menge  $V$ , ferner

- (a) einer Vorschrift, die jedem Paar  $(x, y)$  mit  $x, y \in V$  genau ein Element  $x + y \in V$  zuordnet (*Addition*)

<sup>19</sup> s. Abschn. 2.4.9

(b) und einer Vorschrift, die jedem Paar  $(\lambda, \mathbf{x})$  mit  $\lambda \in \mathbb{K}$  und  $\mathbf{x} \in V$  genau ein Element  $\lambda \mathbf{x} \in V$  zuordnet (*Multiplikation mit Skalaren, s-Multiplikation*), wobei für alle  $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$  und  $\lambda, \mu \in \mathbb{K}$  folgende Regeln gelten:

- (A1)  $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$ , (Assoziativgesetz)  
 (A2)  $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ , (Kommutativgesetz)  
 (A3) Es existiert *genau ein Element*  $\mathbf{0}$  in  $V$  mit

$$\mathbf{x} + \mathbf{0} = \mathbf{x} \quad \text{für alle } \mathbf{x} \in V$$

- (A4) Zu jedem  $\mathbf{x} \in V$  existiert genau ein Element  $\mathbf{x}' \in V$  mit

$$\mathbf{x} + \mathbf{x}' = \mathbf{0}.$$

Man bezeichnet  $\mathbf{x}'$  durch  $-\mathbf{x}$  und nennt es das *Negative* zu  $\mathbf{x}$ . Ferner:

- (S1)  $(\lambda + \mu)\mathbf{x} = \lambda\mathbf{x} + \mu\mathbf{x}$  (Distributivgesetze)  
 (S2)  $\lambda(\mathbf{x} + \mathbf{y}) = \lambda\mathbf{x} + \lambda\mathbf{y}$   
 (S3)  $(\lambda\mu)\mathbf{x} = \lambda(\mu\mathbf{x})$  (Assoziativgesetz)  
 (S4)  $1\mathbf{x} = \mathbf{x}$  mit  $1 \in \mathbb{K}$ .

**Bezeichnungen und Erläuterungen:** Statt Vektorraum über  $\mathbb{K}$  sagt man auch  $\mathbb{K}$ -Vektorraum, oder *linearer Raum*<sup>20</sup> über  $\mathbb{K}$ . Der beschriebene Vektorraum  $V$  über  $\mathbb{K}$  wird auch durch das Tripel  $(V, +, \mathbb{K})$  symbolisiert. Die Elemente von  $V$  werden *Vektoren oder Punkte* genannt, die Elemente von  $\mathbb{K}$  *Skalare*.

Die Additionsgesetze (A1) bis (A4) besagen nichts anderes, als dass  $(V, +)$  eine additive abelsche Gruppe ist, s. Abschn. 2.3.2. Wir benötigen aus Abschn. 2.3.2 nicht mehr, als dass die Subtraktion durch

$$\mathbf{x} - \mathbf{y} := \mathbf{x} + (-\mathbf{y})$$

erklärt ist, und dass man Summen  $\mathbf{x} + \mathbf{y} + \mathbf{z}$ ,  $\mathbf{x} + \mathbf{y} + \mathbf{z} + \mathbf{w}$  usw. meistens ohne Klammern schreibt, da es wegen (A1) gleichgültig ist, wie man Klammern setzt.

Dasselbe gilt übrigens für  $\lambda\mu\mathbf{x}$ ,  $\lambda\mu\alpha\mathbf{x}$  usw. auf Grund von Regel (S3). Man schreibt überdies

$$\mathbf{x}\lambda := \lambda\mathbf{x} \quad \text{und} \quad \frac{\mathbf{x}}{\lambda} := \frac{1}{\lambda}\mathbf{x} \quad (\lambda \neq 0).$$

<sup>20</sup> Der Ausdruck »linearer Raum« hat sich insbesondere in der »Funktionalanalysis« eingebürgert, in der die lineare Algebra und Analysis zu einer höheren Einheit verschmelzen.



**Folgerung 2.9:**

Für alle  $x$  aus einem Vektorraum  $V$  über  $\mathbb{K}$  und alle  $\lambda \in \mathbb{K}$  gilt:

- (a)  $0x = \mathbf{0}, \quad \lambda \mathbf{0} = \mathbf{0}$
- (b)  $\lambda x = \mathbf{0} \Rightarrow (\lambda = 0 \text{ oder } x = \mathbf{0})$
- (c)  $(-\lambda)x = \lambda(-x) = -\lambda x, \quad \text{speziell } (-1)x = -x.$

**Beweis:**

- (a)  $0x = 0x + 0x - 0x = (0 + 0)x - 0x = 0x - 0x = \mathbf{0}$ , und analog  $\lambda \mathbf{0} = \mathbf{0}$ .
- (b) Sei  $\lambda x = \mathbf{0}$ . Dann ist  $\lambda = 0$  oder  $\lambda \neq 0$ . Im letzteren Fall gilt  $x = (\lambda^{-1}\lambda)x = \lambda^{-1}(\lambda x) = \lambda^{-1}\mathbf{0} = \mathbf{0}$ .
- (c)  $(-\lambda)x = (-\lambda)x + \lambda x - \lambda x = (-\lambda + \lambda)x - \lambda x = \mathbf{0} - \lambda x = -\lambda x$ . Mit  $\lambda = 1$  folgt  $(-1)x = -1x = -x$  und damit  $\lambda(-x) = (\lambda(-1))x = (-\lambda)x$ .  $\square$

**Übung 2.13\***

Es sei  $(V, +)$  eine nichttriviale abelsche Gruppe und  $\mathbb{K}$  ein Körper.

Wir definieren das Produkt  $\lambda x$  für alle  $\lambda \in \mathbb{K}$  und alle  $x \in V$  durch  $\lambda x := \mathbf{0}$ . Ist damit  $V$  ein Vektorraum über  $\mathbb{K}$ ? Welche Gesetze in Def. 2.18 sind erfüllt und welche evtl. nicht?

**2.4.2 Beispiele für Vektorräume****Beispiel 2.18:**

Die Vektorräume  $\mathbb{R}^n$  über  $\mathbb{R}$ , sowie  $\mathbb{C}^n$  über  $\mathbb{C}$ , sind wohlbekannt.

**Beispiel 2.19:**

Mit einem beliebigen Körper  $\mathbb{K}$  bildet man den *Vektorraum*  $\mathbb{K}^n$  über  $\mathbb{K}$  nach dem gleichen Muster wie  $\mathbb{R}^n$  über  $\mathbb{R}$ . D.h.  $\mathbb{K}^n$  besteht aus allen Spaltenvektoren, die jeweils  $n$  Elemente aus  $\mathbb{K}$  enthalten (die »Koordinaten« des Spaltenvektors). Die Addition  $+$  geschieht koordinatenweise, wie beim  $\mathbb{R}^n$ , und die Multiplikation mit  $\lambda \in \mathbb{K}$  ebenfalls.

Die Kraft des allgemeinen Vektorraumbegriffes entfaltet sich besonders bei Mengen von Funktionen, den sogenannten *Funktionenräumen*. Dazu geben wir einige Beispiele an.

**Beispiel 2.20:**

Die Menge  $C(I)$  aller stetigen reellwertigen Funktionen auf einem Intervall  $I$  bildet bzgl. der üblichen Addition von Funktionen und der üblichen Multiplikation mit reellen Zahlen einen Vektorraum über  $\mathbb{R}$ .

**Beispiel 2.21:**

Analog zu  $C(I)$  kann man die Menge  $C^k(I)$  aller  $k$ -mal stetig differenzierbaren reellwertigen Funktionen auf  $I$  als Vektorraum über  $\mathbb{R}$  auffassen ( $k \in \mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$ ). Dabei identifiziert man  $C^0(I) := C(I)$ ,  $C^\infty(I)$  ist der Vektorraum der beliebig oft stetig differenzierbaren

Funktionen  $f : I \rightarrow \mathbb{R}$ . Es folgt

$$C(I) \supset C^1(I) \supset C^2(I) \supset \dots \supset C^\infty(I).$$

### Beispiel 2.22:

Die Menge  $\text{Pol } \mathbb{R}$  aller Polynome  $p(x) = a_0 + a_1x + \dots + a_nx^n$  ( $x \in \mathbb{R}, a_i \in \mathbb{R}$ ) für beliebige  $n \in \mathbb{N}_0$  bildet bzgl. Addition und Multiplikation mit reellen Zahlen einen Vektorraum über  $\mathbb{R}$ .

Mit  $\text{Pol}_n \mathbb{R}$  bezeichnen wir die Menge aller Polynome aus  $\text{Pol } \mathbb{R}$ , deren Grad höchstens  $n$  ist. Es gilt zweifellos

$$\mathbb{R} = \text{Pol}_0 \mathbb{R} \subset \text{Pol}_1 \mathbb{R} \subset \text{Pol}_2 \mathbb{R} \subset \dots \subset \text{Pol } \mathbb{R}.$$

Auch Zahlenfolgen können sich zu Vektorräumen formieren. So bilden *alle reellen Zahlenfolgen* einen Vektorraum über  $\mathbb{R}$  bzgl. gliedweiser Addition und gliedweiser Multiplikation mit reellen Zahlen. Aber auch *alle konvergenten Folgen*, *alle Nullfolgen* und *alle beschränkten Folgen* stellen je einen Vektorraum dar. Ein für die *Theorie der Fourierreihen wichtiger Folgenraum*<sup>21</sup> sei abschließend erwähnt:

### Beispiel 2.23:

Die Menge  $\ell^2$  aller reellen Zahlenfolgen  $(a_0, a_1, a_2, \dots)$  mit

$$\sum_{k=0}^{\infty} a_k^2 \text{ konvergent}$$

ist ein Vektorraum über  $\mathbb{R}$ . Er wird *Hilbertscher Folgenraum*<sup>22</sup> genannt. Er ist dem  $\mathbb{R}^n$  sehr verwandt. Zwar haben seine Elemente  $\mathbf{a} = (a_0, a_1, a_2, \dots)$  unendlich viele Koordinaten, doch kann man aus  $\mathbf{a}$  und  $\mathbf{b} = (b_0, b_1, b_2, \dots)$  in  $\ell^2$  das *innere Produkt*  $\mathbf{a} \cdot \mathbf{b} = \sum_{i=0}^{\infty} a_i b_i$  bilden, sowie die Länge  $|\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}}$  definieren. Damit lassen sich, wie im  $\mathbb{R}^n$ , Winkel  $\angle(\mathbf{a}, \mathbf{b})$  erklären, kurz, man kann »Geometrie« betreiben. (Die Bedeutung von  $\ell^2$  in der Fourierreihen-Theorie wird in Abschn. 2.4.9 kurz erläutert.)

### Beispiel 2.24:

Allgemeiner als  $\ell^2$  wird  $\ell^p$  eingeführt, der Vektorraum aller reellen Folgen  $\mathbf{a} = (a_0, a_1, a_2, \dots)$  mit konvergenter Summe  $\sum_{i=0}^{\infty} |a_i|^p$  ( $p > 1$ ). Hier arbeitet man mit der »Vektorlänge«  $|\mathbf{a}|_p =$

$$\left( \sum_{i=0}^{\infty} |a_i|^p \right)^{1/p}. \text{ Innere Produkte werden im Falle } p \neq 2 \text{ nicht benutzt.}$$

### Übung 2.14\*

Welche der folgenden Mengen sind Vektorräume über  $\mathbb{R}$ ?

<sup>21</sup> s. Abschn. 2.4.9

<sup>22</sup> Nach dem deutschen Mathematiker David Hilbert (1862–1943).

- (a) Die Menge der reellen Folgen  $(a_n)_{n \in \mathbb{N}}$  mit  $a_n \rightarrow 1$  für  $n \rightarrow \infty$ ?
- (b) Die Menge aller reellen *abbrechenden* Folgen  $(a_n)_{n \in \mathbb{N}}$ ? (Eine *abbrechende* Folge sieht so aus:  $(a_1, a_2, a_3, \dots, a_N, 0, 0, 0, \dots)$ , d.h.  $a_n = 0$  für alle  $n > N$ . Der *Abbrechindex*  $N$  kann für verschiedene Folgen dabei verschieden sein.)
- (c) Die Menge der Funktionen  $f: \mathbb{R} \rightarrow \mathbb{R}$  mit  $|f(x)| \rightarrow 0$  für  $|x| \rightarrow \infty$ ?
- (d) Die Vektoren  $x = \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix} \in \mathbb{R}^3$ ? Überall werden dabei die üblichen Operationen  $+$  und  $\lambda \cdot$  (mit  $\lambda \in \mathbb{R}$ ) verwendet.

### 2.4.3 Unterräume, Basis, Dimension

Wir machen es uns jetzt ziemlich einfach: Die Begriffe und Sätze aus dem entsprechenden Abschn. 2.1.3 über den  $\mathbb{R}^n$  werden auf Vektorräume über  $K$  ausgedehnt. Wir wollen sehen, wie weit das möglich ist.

#### Definition 2.19:

Es sei  $V$  ein Vektorraum über dem Körper  $\mathbb{K}$ . Als *Linearkombination* der Vektoren  $a_1, \dots, a_m \in V$  bezeichnet man jede Summe

$$\sum_{i=1}^m \lambda_i a_i \quad \text{mit } \lambda_i \in \mathbb{K}.$$

Die Vektoren  $a_1, \dots, a_m \in V$  heißen *linear abhängig*, wenn wenigstens einer der Vektoren als Linearkombination der übrigen geschrieben werden kann oder einer der Vektoren  $0$  ist.<sup>23</sup> Andernfalls nennt man die  $a_1, \dots, a_m$  *linear unabhängig*.

Die lineare Unabhängigkeit der  $a_1, \dots, a_m$  ist gleichbedeutend damit, dass die Linearkombination

$$\sum_{i=1}^m \lambda_i a_i = 0 \quad (\lambda_i \in \mathbb{K}) \tag{2.74}$$

nur durch  $\lambda_1 = \lambda_2 = \dots = \lambda_m = 0$  erfüllt werden kann. (Beweis wie in Folg. 1.12, Abschn. 1.2.7.)

#### Satz 2.17:

(*Fundamentallemma in  $\mathbb{K}^n$* ) Je  $n + 1$  Vektoren aus  $\mathbb{K}^n$  sind linear abhängig.

#### Beweis:

Der Gaußsche Algorithmus ohne Pivotierung (s. Abschn. 2.2.1, 2.2.3) verläuft mit Werten aus  $\mathbb{K}$  ganz genau so wie mit Werten aus  $\mathbb{R}$ . Damit lässt sich der Beweis von Satz 2.2, Abschn. 2.1.3 wörtlich übertragen, wenn man  $\mathbb{R}$  durch  $\mathbb{K}$  ersetzt.  $\square$

<sup>23</sup> Letzteres ist hauptsächlich auf den Fall  $m = 1$  gemünzt.

**Definition 2.20:**

$V$  sei ein Vektorraum über  $\mathbb{K}$ . Eine nichtleere Teilmenge  $U$  von  $V$  heißt ein *Unterraum* von  $V$ , wenn mit je zwei Vektoren  $\mathbf{x}, \mathbf{y} \in U$  auch  $\mathbf{x} + \mathbf{y} \in U$  und  $\lambda \mathbf{x} \in U$  (für alle  $\lambda \in \mathbb{K}$ ) ist.

Natürlich bedeutet dies, dass  $U$  selbst ein Vektorraum über  $\mathbb{K}$  ist bzgl. der Addition und  $s$ -Multiplikation in  $V$ , wie man leicht nachweist.

Kleinsten Unterraum von  $V$  ist  $\{\mathbf{0}\}$ , größter  $V$  selbst.

**Definition 2.21:**

- (a) Es sei  $V$  ein Vektorraum über  $\mathbb{K}$ .  $V$  heißt *unendlichdimensional*, wenn es beliebig viele linear unabhängige Vektoren in  $V$  gibt. Symbolisch ausgedrückt:

$$\dim V = \infty.$$

$V$  heißt *endlichdimensional*, wenn  $V$  nicht unendlichdimensional ist.

- (b) Ist  $V$  endlichdimensional, so heißt die maximale Anzahl linear unabhängiger Vektoren in  $V$  die *Dimension* von  $V$ . Diese Zahl sei  $n$ . Man schreibt

$$\dim V = n.$$

Je  $n$  linear unabhängige Vektoren  $\mathbf{a}_1, \dots, \mathbf{a}_n \subset V$  bilden eine *Basis*  $(\mathbf{a}_1, \dots, \mathbf{a}_n)$  von  $V$ .<sup>24</sup>

Die Übertragung der Begriffe und Sätze von  $\mathbb{R}^n$  in  $V$  klappt wie am Schnürchen. Es gilt:

**Satz 2.18:**

- (a) Ist  $(\mathbf{a}_1, \dots, \mathbf{a}_n)$  eine Basis des  $n$ -dimensionalen Vektorraumes  $V$  über  $\mathbb{K}$ , so besteht  $V$  aus allen Linearkombinationen

$$\mathbf{x} = \lambda_1 \mathbf{a}_1 + \lambda_2 \mathbf{a}_2 + \dots + \lambda_n \mathbf{a}_n \quad (\lambda_i \in \mathbb{K}). \quad (2.75)$$

- (b) Die  $\lambda_i$  sind durch  $\mathbf{x}$  und  $\mathbf{a}_1, \dots, \mathbf{a}_n$  eindeutig bestimmt.

Der Beweis verläuft wörtlich wie der Beweis des Satzes 2.3, Abschn. 2.1.3, wenn man dort  $V$  statt  $U$ ,  $\mathbb{K}$  statt  $\mathbb{R}$  und  $n$  statt  $m$  setzt.

Ist  $A$  eine beliebige nichtleere Teilmenge eines Vektorraumes  $V$  über  $\mathbb{K}$ , so bilden alle Linearkombinationen  $\mathbf{x} = \sum_{i=1}^p \lambda_i \mathbf{a}_i$  mit beliebigen  $p \in \mathbb{N}$ ,  $\lambda_i \in \mathbb{K}$ ,  $\mathbf{a}_i \in A$  zweifellos einen Unterraum  $U$  von  $V$ . Er wird

$$\text{Span } A := U$$

<sup>24</sup> Oft wird als Basis auch die *Menge*  $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$  bezeichnet. Wir bevorzugen aber das  $n$ -Tupel  $(\mathbf{a}_1, \dots, \mathbf{a}_n)$  für den Basisbegriff, da man so »Rechts-« und »Linkssysteme« unterscheiden kann.

genannt. Man sagt: » $A$  spannt  $U$  auf« oder » $A$  ist ein *Erzeugendensystem* von  $U$ «. Im Fall  $U = V$  spannt  $A$  den ganzen Raum  $V$  auf:  $\text{Span } A = V$ , und  $A$  ist dann — wen wundert's — ein *Erzeugendensystem* von  $V$ . Es gilt

$$A \subset \text{Span } A.$$

Ist  $A = \{\mathbf{a}_1, \dots, \mathbf{a}_m\}$  dabei endlich, so besteht

$$\text{Span}\{\mathbf{a}_1, \dots, \mathbf{a}_m\}$$

genau aus allen Summen  $\sum_{i=1}^m \lambda_i \mathbf{a}_i$ . Im Fall linear unabhängiger  $\mathbf{a}_i$  folgt:

**Satz 2.19:**

(*Konstruktion von Unterräumen*) Sind die Vektoren  $\mathbf{a}_1, \dots, \mathbf{a}_m$  aus dem Vektorraum  $V$  über  $\mathbb{K}$  linear unabhängig, so hat der von ihnen aufgespannte Unterraum

$$U = \text{Span}\{\mathbf{a}_1, \dots, \mathbf{a}_m\}$$

die Dimension  $m$ . Die Menge  $(\mathbf{a}_1, \dots, \mathbf{a}_m)$  ist also eine Basis von  $U$ .

**Beweis:**

$U$  besteht aus allen Vektoren

$$\mathbf{x} = \sum_{i=1}^m \lambda_i \mathbf{a}_i \quad (\lambda_i \in \mathbb{K}). \quad (2.76)$$

Es ist  $\dim U \geq m$ , da die  $\mathbf{a}_1, \dots, \mathbf{a}_m \in U$  linear unabhängig sind. Es ist daher nachzuweisen, dass je  $m + 1$  Vektoren aus  $U$  linear abhängig sind (woraus  $\dim U = m$  folgt). Dieser Nachweis verläuft völlig analog zum Beweis von Satz 2.4 (ab II), in Abschn. 2.1.3 (unter Benutzung des »Fundamentallemmas« Satz 2.17).  $\square$

Für den Spezialfall  $U = V$  erhält man

**Folgerung 2.10:**

Sind die Vektoren  $\mathbf{a}_1, \dots, \mathbf{a}_n$  aus dem Vektorraum  $V$  über  $\mathbb{K}$  linear unabhängig und spannen sie  $V$  auf, so bilden sie eine Basis von  $V$ . Insbesondere ist dann  $\dim V = n$ .

Auch der Satz über den Austausch von Basiselementen (Satz 2.5, Abschn. 2.1.3) wird samt Beweis übertragen. (Man hat nur  $U$  aus  $\mathbb{R}^n$  durch  $V$  zu ersetzen):

**Satz 2.20:**

(*Austausch von Basiselementen*) Es sei  $V$  ein  $n$ -dimensionaler Vektorraum über  $\mathbb{K}$  mit der Basis  $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ . Sind  $\mathbf{b}_1, \dots, \mathbf{b}_k$  ( $k < n$ ) beliebige linear unabhängige Vektoren aus  $V$ , so kann man sie zu einer Basis  $(\mathbf{b}_1, \dots, \mathbf{b}_k, \mathbf{b}_{k+1}, \dots, \mathbf{b}_n)$  von  $V$  ergänzen, wobei die  $\mathbf{b}_{k+1}, \dots, \mathbf{b}_n$  aus  $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$  entnommen sind.

### Basis-Wechsel

Geht man im Vektorraum  $V$  über  $\mathbb{K}$  von einer Basis  $(\mathbf{a}_1, \dots, \mathbf{a}_n)$  zu einer neuen Basis  $(\mathbf{b}_1, \dots, \mathbf{b}_n)$  über, so verwandelt sich die Darstellung

$$\mathbf{x} = \sum_{i=1}^n \xi_i \mathbf{a}_i \quad (\xi_i \in \mathbb{K}) \quad (2.77)$$

eines beliebigen Punktes  $\mathbf{x} \in V$  folgendermaßen: Zunächst gilt

$$\mathbf{a}_i = \sum_{k=1}^n \alpha_{ki} \mathbf{b}_k, \quad i = 1, \dots, n \quad (2.78)$$

mit bestimmten  $\alpha_{ki} \in \mathbb{K}$ . Damit erhält man

$$\mathbf{x} = \sum_{i=1}^n \xi_i \sum_{k=1}^n \alpha_{ki} \mathbf{b}_k = \sum_{k=1}^n \mathbf{b}_k \sum_{i=1}^n \alpha_{ki} \xi_i,$$

also

$$\mathbf{x} = \sum_{k=1}^n \xi'_k \mathbf{b}_k \quad \text{mit} \quad \xi'_k = \sum_{i=1}^n \alpha_{ki} \xi_i. \quad (2.79)$$

### Lineare Mannigfaltigkeiten

Entsprechend Abschn. 2.1.3 wird definiert: Ist  $U$  ein Unterraum des  $\mathbb{K}$ -Vektorraums  $V$  und  $\mathbf{r}_0$  ein beliebiger Vektor aus  $V$ , dann nennt man

$$M = \mathbf{r}_0 + U := \{\mathbf{r}_0 + \mathbf{x} \mid \mathbf{x} \in U\}$$

eine *lineare Mannigfaltigkeit* in  $V$ . Für die Dimension von  $M$  vereinbart man einfach  $\dim M := \dim U$  (auch im Fall  $\dim U = \infty$ ). Ist  $U$  dabei ein Unterraum von der Art, dass  $U \cup \{\mathbf{x}_0\}$  mit einem  $\mathbf{x}_0 \in V \setminus U$  den Raum  $V$  aufspannt, dann heißt  $M$  eine *Hyperebene* in  $V$ . Ist dagegen  $U = \{\lambda \mathbf{x}_0 \mid \lambda \in \mathbb{K}\}$ , ( $\mathbf{x}_0 \in V$ ,  $\mathbf{x}_0 \neq \mathbf{0}$ ), dann heißt  $M$  eine *Gerade* in  $V$ . Auf diese Weise bekommen wir immer mehr »Geometrie« in unsere staubtrockenen Vektorräume.

Ein wichtiges Anwendungsbeispiel für endlichdimensionale Funktionenräume und Mannigfaltigkeiten tritt bei linearen Differentialgleichungssystemen auf.

#### Beispiel 2.25:

Man betrachte das *lineare Differentialgleichungssystem*

$$y'_i(x) = \sum_{k=1}^n a_{ik}(x) y_k(x) + b_i(x), \quad i = 1, \dots, n, \quad (2.80)$$

$x \in I$  (Intervall), wobei die  $a_{ik} : I \rightarrow \mathbb{R}$  und  $b_i : I \rightarrow \mathbb{R}$  gegebene stetige Funktionen sind und die  $y_i : I \rightarrow \mathbb{R}$  gesuchte stetig differenzierbare Funktionen (s. Burg/Haf/Wille (Band III) [25]).

1. *Fall  $b_i(x) \equiv 0$  für alle  $i$  (homogener Fall):* Fasst man die  $y_i(x)$  zu einem »Funktionsvektor«  $\mathbf{y}(x) = [y_1(x), \dots, y_n(x)]^T$  zusammen, so bilden die Lösungen  $\mathbf{y}(x)$  von (2.80) einen  $n$ -dimensionalen Vektorraum  $V$  über  $\mathbb{R}$  (bzgl. der koordinatenweisen Addition und Multiplikation mit  $\lambda \in \mathbb{R}$ ). Eine Basis dieses Vektorraumes heißt ein *Fundamentalsystem* von Lösungen (s. Burg/Haf/Wille (Band III) [25]).

2. *Fall  $b_i(x)$  nicht alle  $\equiv 0$  (inhomogener Fall):* Ist durch  $\mathbf{y}_0(x) = [y_1(x), \dots, y_n(x)]^T$  irgendeine Lösung gegeben, so wird mit einem Fundamentalsystem  $(\mathbf{y}_1(x), \dots, \mathbf{y}_n(x))$  aus dem homogenen Fall jede Lösung durch

$$\mathbf{y}(x) = \mathbf{y}_0(x) + \sum_{i=1}^n \lambda_i \mathbf{y}_i(x), \quad \lambda_i \in \mathbb{R},$$

beschrieben, und jedes  $\mathbf{y}(x)$  dieser Art ist auch Lösung. Die Lösungsmenge ist damit eine  $n$ -dimensionale Mannigfaltigkeit im Raum  $C^1(I)$  aller stetig differenzierbaren Funktionen  $f : I \rightarrow \mathbb{R}$ . (s. Burg/Haf/Wille (Band III) [25]).

**Bemerkung:** Die Funktionenräume  $C(I)$ ,  $C^k(I)$ ,  $C^\infty(I)$ ,  $\text{Pol } \mathbb{R}$ , sind *unendlichdimensional*, da sie alle  $\text{Pol } \mathbb{R}$  als Unterraum enthalten.<sup>25</sup> Der Raum  $\text{Pol } \mathbb{R}$  ist aber unendlichdimensional, da er alle Potenzfunktionen  $p_k(x) = x^k$  ( $k \in \mathbb{N}_0$ ) enthält, von denen je endlich viele linear unabhängig sind (denn aus  $0 \equiv \sum_{k=0}^m \alpha_k x^k$  folgt  $\alpha_k = 0$  für alle  $k = 0, \dots, m$ , da ein nicht verschwindendes Polynom  $m$ -ten Grades höchstens  $m$  Nullstellen hat.)

$\ell^p$  ist ebenfalls unendlichdimensional. (Man betrachte dazu  $\mathbf{e}_i = (0, \dots, 0, 1, 0, 0, \dots) \in \ell^p$ , 1 an  $i$ -ter Stelle).

### Übung 2.15\*

Zeige, dass die Menge der trigonometrischen Reihen

$$f(x) = a_0 + \sum_{k=1}^n (a_k \cos(kx) + b_k \sin(kx)), \quad x \in \mathbb{R}, \quad (n \text{ fest})$$

einen Vektorraum  $T$  über  $\mathbb{R}$  bildet. Beweise ferner, dass  $(1, \cos x, \sin x, \cos(2x), \sin(2x), \dots, \cos(nx), \sin(nx))$  eine Basis von  $T$  ist, also insbesondere  $\dim T = 2n + 1$  gilt.

### Übung 2.16\*

- Zeige, dass alle reellen Polynome, die höchstens den Grad 10 haben und eine Nullstelle bei  $x = 1$ , einen Vektorraum über  $\mathbb{R}$  bilden. Welche Dimension hat der Raum? Gib eine Basis dazu an!
- Löse das gleiche Problem (wie unter (a)) für Polynome vom Höchstgrad 10, die wenigstens zwei Nullstellen haben, und zwar bei 1 und  $-1$ .

<sup>25</sup> Wobei die Polynome nur auf  $I$  betrachtet werden.

### 2.4.4 Direkte Summen, freie Summen

**Direkte Summen:** Es seien  $U_1, U_2, \dots, U_m$  Unterräume des Vektorraums  $V$  über  $\mathbb{K}$ , deren Vereinigung  $U_1 \cup U_2 \cup \dots \cup U_m$  den Raum  $V$  aufspannt. Man nennt dann  $V$  die *Summe*  $U_1 + U_2 + \dots + U_m$  der *Unterräume*, d.h. es ist

$$U_1 + U_2 + \dots + U_m := \text{Span}(U_1 \cup U_2 \cup \dots \cup U_m).$$

Gilt zusätzlich

$$(U_1 + \dots + U_{k-1} + U_{k+1} + \dots + U_m) \cap U_k = \{\mathbf{0}\} \quad \text{für alle } k = 1, \dots, m, \quad (2.81)$$

so nennt man  $V$  die *direkte Summe*  $U_1, \dots, U_m$ , und beschreibt dies durch

$$V = U_1 \oplus U_2 \oplus \dots \oplus U_m.$$

Im Falle zweier Unterräume, also  $V = U_1 + U_2$ , bedeutet  $V = U_1 \oplus U_2$  einfach  $U_1 \cap U_2 = \{\mathbf{0}\}$ .

#### Folgerung 2.11:

Gilt  $V = U_1 + U_2 + \dots + U_m$  ( $U_i$  Unterräume des Vektorraums  $V$ ), so sind folgende Aussagen äquivalent:

- (a)  $V = U_1 \oplus U_2 \oplus \dots \oplus U_m$ .
- (b) Jedes  $\mathbf{x} \in V$  lässt sich eindeutig darstellen als  $\mathbf{x} = \mathbf{u}_1 + \mathbf{u}_2 + \dots + \mathbf{u}_m$  mit  $\mathbf{u}_i \in U_i$  für alle  $i = 1, \dots, m$ .
- (c)  $\mathbf{0} = \mathbf{u}_1 + \dots + \mathbf{u}_m$  ( $\mathbf{u}_i \in U_i$ )  $\Rightarrow \mathbf{u}_i = \mathbf{0}$  für alle  $i$ .  
Ist  $V$  endlichdimensional und  $V = U_1 + \dots + U_m$ , so ist (a) äquivalent zu
- (d)  $\dim V = \dim U_1 + \dim U_2 + \dots + \dim U_m$ .

#### Beweis:

(a)  $\Rightarrow$  (c): Aus  $\mathbf{u}_1 + \dots + \mathbf{u}_m = \mathbf{0}$  ( $\mathbf{u} \in U_i$ ) folgt  $\mathbf{u}_1 = -(\mathbf{u}_2 + \dots + \mathbf{u}_m) \in U_2 + \dots + U_m$  und  $\in U_1$ . Da wegen (a)  $U_1 \cap (U_2 + \dots + U_m) = \{\mathbf{0}\}$ , so  $\mathbf{u}_1 = \mathbf{0}$ ; entsprechend  $\mathbf{u}_2 = \mathbf{u}_3 = \dots = \mathbf{u}_m = \mathbf{0}$ .

(c)  $\Rightarrow$  (b): Hat  $\mathbf{x}$  zwei Darstellungen  $\mathbf{x} = \sum_{i=1}^m \mathbf{u}_i = \sum_{i=1}^m \mathbf{u}'_i$ , so folgt  $\sum_{i=1}^m (\mathbf{u}_i - \mathbf{u}'_i) = \mathbf{0}$ , also  $\mathbf{u}_i - \mathbf{u}'_i = \mathbf{0}$ , d.h.  $\mathbf{u}_i = \mathbf{u}'_i$  für alle  $i$ .

(b)  $\Rightarrow$  (c): klar.

(c)  $\Rightarrow$  (a): Sei  $\mathbf{u}_1 \in U_1 \cap (U_2 + \dots + U_m)$ , folglich  $\mathbf{u}_1 = \mathbf{u}_2 + \dots + \mathbf{u}_m$  mit  $\mathbf{u}_i \in U_i$ . Somit  $\mathbf{0} = -\mathbf{u}_1 + \mathbf{u}_2 + \dots + \mathbf{u}_m \Rightarrow \mathbf{u}_1 = \mathbf{0} \Rightarrow U_1 \cap (U_2 + \dots + U_m) = \{\mathbf{0}\}$ . Entsprechend gilt (2.81) für alle  $k \Rightarrow$  (a).

(a)  $\Leftrightarrow$  (d): Fall  $m = 2$ :

$$[V = U_1 \oplus U_2 \text{ und } [(a_1, \dots, a_s) \text{ Basis von } U_1, (b_1, \dots, b_t) \text{ Basis von } U_2]]$$

$$\Leftrightarrow [(a_1, \dots, a_s, b_1, \dots, b_t) \text{ Basis von } V (a_i \in U_1, b_i \in U_2)]$$

$$\Leftrightarrow \dim V = \dim U_1 + \dim U_2.$$



Für  $m > 2$  folgt (a)  $\Leftrightarrow$  (d) durch vollständige Induktion.  $\square$

### Beispiel 2.26:

Das folgende Beispiel macht klar, dass es sich bei direkten Summen im Grunde um etwas sehr naheliegendes handelt. Und zwar betrachten wir in  $\mathbb{R}^n$  den Unterraum  $U_1$  der Vektoren  $\mathbf{u}_1 = [x_1, \dots, x_s, 0, 0, \dots, 0]^T$ , deren Koordinaten  $x_{s+1} = \dots = x_n = 0$  sind. Entsprechend  $U_2$ , bestehend aus allen  $\mathbf{u}_2 = [0, 0, \dots, 0, x_{s+1}, \dots, x_n]^T$ . Für  $\mathbf{x} = [x_1, \dots, x_n]^T$  folgt damit die eindeutige Zerlegung

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_s \\ x_{s+1} \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_s \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ x_{s+1} \\ \vdots \\ x_n \end{bmatrix} = \mathbf{u}_1 + \mathbf{u}_2, \quad \text{also } U_1 \oplus U_2 = \mathbb{R}^n.$$

Im  $\mathbb{K}^n$  verläuft dies natürlich genauso.

Dies gibt Anlass zu folgender Konstruktion:

### Freie Summen

Sind  $U_1, \dots, U_m$  beliebige Vektorräume über  $\mathbb{K}$ , so bildet die Menge der  $m$ -Tupel

$$\begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_m \end{bmatrix} \quad (\mathbf{u}_i \in U_i), \quad \text{mit} \quad \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_m \end{bmatrix} + \begin{bmatrix} \mathbf{u}'_1 \\ \vdots \\ \mathbf{u}'_m \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1 + \mathbf{u}'_1 \\ \vdots \\ \mathbf{u}_m + \mathbf{u}'_m \end{bmatrix}, \quad \lambda \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_m \end{bmatrix} = \begin{bmatrix} \lambda \mathbf{u}_1 \\ \vdots \\ \lambda \mathbf{u}_m \end{bmatrix},$$

einen Vektorraum  $V$  über  $\mathbb{K}$ . Er heißt die *freie Summe der Vektorräume*  $U_1, \dots, U_m$ , beschrieben durch

$$V = U_1 \dot{+} U_2 \dot{+} \dots \dot{+} U_m.$$

Auch für die Vektoren  $\mathbf{u}_i = U_i$  verwenden wir eine solche Schreibweise:

$$\mathbf{u}_1 \dot{+} \dots \dot{+} \mathbf{u}_m := \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_m \end{bmatrix}.$$

$U_i$  lässt sich bijektiv auf den folgenden Unterraum  $\overline{U}_i \subset V$  abbilden:

$$\overline{U}_i = \left\{ \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{u}_i \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} \mid \mathbf{u}_i \in U_i \right\}, \quad \text{durch} \quad \mathbf{u}_i \mapsto \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{u}_i \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}.$$

(Man sagt,  $U_i$  wird auf diese Weise in den Vektorraum  $V$  »eingebettet«.) Es folgt  $V = \overline{U}_1 \oplus \dots \oplus \overline{U}_m$ . Damit ist die freie Summe in eine direkte Summe überführt worden, und alle Eigenschaften direkter Summen können auf freie Summen übertragen werden.

### Beispiel 2.27:

Die freie Summe  $\mathbb{R}^3 \dot{+} \mathbb{R}^2$  besteht aus allen Vektoren der Form

$$\mathbf{x} = \begin{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ \begin{bmatrix} x_4 \\ x_5 \end{bmatrix} \end{bmatrix}, \quad \text{mit} \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3, \quad \begin{bmatrix} x_4 \\ x_5 \end{bmatrix} \in \mathbb{R}^2.$$

Wir lassen hier die inneren Klammern weg, schreiben also

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}, \quad \text{und damit} \quad \mathbb{R}^3 \dot{+} \mathbb{R}^2 = \mathbb{R}^5.$$

Allgemein also:

$$\mathbb{R}^{n_1} \dot{+} \mathbb{R}^{n_2} \dot{+} \dots \dot{+} \mathbb{R}^{n_t} = \mathbb{R}^{n_1+n_2+\dots+n_t}.$$

Die freien und direkten Summen von Vektorräumen spielen bei direkten Summen von Matrizen eine Rolle (s. beispielsweise Abschn. 4.1).

### Übung 2.17\*

Es seien  $U, V$  Unterräume des endlichdimensionalen Vektorraumes  $W$  über  $\mathbb{K}$  mit  $U + V = W$ .

Beweise

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

Anleitung: Konstruiere Basen  $B_U, B_V, B_{U \cap V}$ , mit  $B_{U \cap V} = B_U \cap B_V$ .

### 2.4.5 Lineare Abbildungen: Definition und Beispiele

Abbildungen von einem Vektorraum in den anderen, die Summen in Summen überführen und Produkte in Produkte, also die Struktur der Vektorräume berücksichtigen, sind von besonderem Interesse: Mit diesen »linearen Abbildungen« wollen wir uns im Folgenden beschäftigen.

#### Definition 2.22:

Es seien  $V$  und  $W$  zwei Vektorräume über dem gleichen Körper  $\mathbb{K}$ . Eine Abbildung  $f : V \rightarrow W$  heißt eine *lineare Abbildung* von  $V$  in  $W$ , wenn für alle  $\mathbf{x}, \mathbf{y} \in V$  und alle  $\lambda \in \mathbb{K}$  folgendes gilt:

$$\begin{aligned} \text{(H1)} \quad f(\mathbf{x} + \mathbf{y}) &= f(\mathbf{x}) + f(\mathbf{y}) && \text{(Additivität)} \\ \text{(H2)} \quad f(\lambda \mathbf{x}) &= \lambda f(\mathbf{x}) && \text{(Homogenität).} \end{aligned}$$

Man sagt auch:  $f$  ist eine *strukturverträgliche Abbildung*.

Lineare Abbildungen heißen auch *lineare Transformationen*, *lineare Operatoren* oder (*Vektorraum*-)*Homomorphismen*. Alle diese Bezeichnungen bedeuten dasselbe.

Aus (H1), (H2) folgt durch sukzessives Anwenden

$$\text{(H)} \quad f\left(\sum_{k=1}^n \lambda_k \mathbf{x}_k\right) = \sum_{k=1}^n \lambda_k f(\mathbf{x}_k) \quad \text{für alle } \mathbf{x}_k \in V, \lambda_k \in \mathbb{K}. \quad (2.82)$$

Umgekehrt folgt aus (H) sowohl (H1) ( $n = 2, \lambda_1, \lambda_2 = 1$ ) wie (H2) ( $n = 1$ ), also gilt: (H)  $\Leftrightarrow$  ((H1) und (H2)). Die Eigenschaft (H) heißt *Linearität* von  $f$ .

#### Beispiel 2.28:

*Lineare Abbildungen von  $\mathbb{R}^n$  in  $\mathbb{R}^m$ .* Ist  $\mathbf{x} = [x_1, \dots, x_n]^T$  ein Vektor des  $\mathbb{R}^n$ , so kann man ihm den Vektor  $\mathbf{y} = [y_1, \dots, y_m]^T$  aus  $\mathbb{R}^m$  durch die Gleichung

$$y_i = \sum_{k=1}^n a_{ik} x_k, \quad i = 1, \dots, m \quad (2.83)$$

eindeutig zuordnen. Die so erklärte Abbildung  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  ist sicher linear. Sie ist durch das rechteckige Schema

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \quad (2.84)$$

der Zahlen  $a_{ik}$  vollständig bestimmt. Ein solches Schema nennt man eine *reelle Matrix* (symbolisiert durch einen großen, fett geschriebenen Buchstaben, z.B.  $A$ ). Umgekehrt kann jede lineare Abbildung  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$  durch eine Matrix der Form (2.84) nebst Gl. (2.83) beschrieben werden. Bezeichnet man nämlich die Bilder der Koordinateneinheitsvektoren  $\mathbf{e}_k =$

$[0, \dots, 0, 1, 0, \dots, 0]^T \in \mathbb{R}^n$  (mit 1 an  $k$ -ter Stelle) mit  $\mathbf{a}_k = [a_{1k}, a_{2k}, \dots, a_{mk}]^T$  ( $k = 1, \dots, n$ ), also  $g(\mathbf{e}_k) = \mathbf{a}_k$ , so folgt für das Bild  $\mathbf{y}$  eines Vektors  $\mathbf{x} = \sum_{k=1}^n x_k \mathbf{e}_k \in \mathbb{R}^n$  nach (2.82)

$$\mathbf{y} = g\left(\sum_{k=1}^n x_k \mathbf{e}_k\right) = \sum_{k=1}^n x_k g(\mathbf{e}_k) = \sum_{k=1}^n x_k \mathbf{a}_k.$$

Schreibt man die Koordinaten der linken und rechten Seite der Gleichung hin, so entsteht gerade (2.83). Matrizen der Form (2.84) beschreiben also über (2.79) alle linearen Abbildungen von  $\mathbb{R}^n$  in  $\mathbb{R}^m$ .

(Matrizen werden in Abschn. 3 ausführlich behandelt.)

Es ist klar, dass in  $\mathbb{K}^n$  ( $\mathbb{K}$  beliebiger algebraischer Körper) alles genauso verläuft.

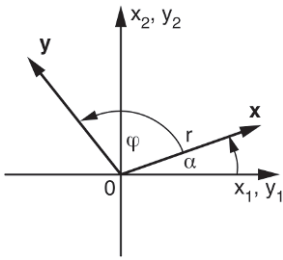


Fig. 2.8: Drehung im  $\mathbb{R}^2$ .

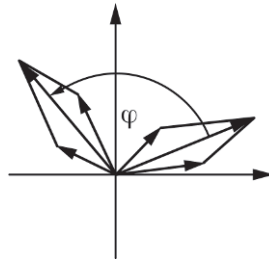


Fig. 2.9: Additivität der Drehung.

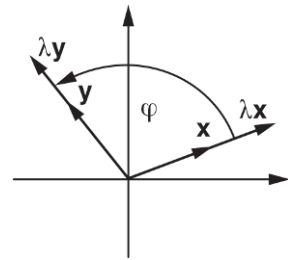


Fig. 2.10: Homogenität der Drehung, mit Fig. 2.9 zusammen also: Linearität.

### Beispiel 2.29:

(Drehungen der Ebene) Die Ebene  $\mathbb{R}^2$  soll um 0 gedreht werden, und zwar um den Winkel  $\varphi$ , s. Fig. 2.8 ( $\varphi > 0$ : Drehung gegen den Uhrzeigersinn;  $\varphi < 0$ : mit dem Uhrzeigersinn). D.h. ein Ortsvektor

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} r \cos \alpha \\ r \sin \alpha \end{bmatrix}$$

( $r, \alpha$  Polarkoordinaten von  $\mathbf{x}$ ) soll in

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} r \cos(\alpha + \varphi) \\ r \sin(\alpha + \varphi) \end{bmatrix}$$

überführt werden. Diese Zuordnung beschreiben wir durch  $\mathbf{y} = f(\mathbf{x})$ . Mit dem Additionstheorem von  $\cos$  und  $\sin$  (Abschn. 1.1.2) folgt

$$\begin{aligned} y_1 &= r(\cos \alpha \cos \varphi - \sin \alpha \sin \varphi) = (\cos \varphi)x_1 - (\sin \varphi)x_2 \\ y_2 &= r(\cos \alpha \sin \varphi + \sin \alpha \cos \varphi) = (\sin \varphi)x_1 + (\cos \varphi)x_2, \end{aligned}$$

$$\begin{aligned} y_1 &= (\cos \varphi)x_1 - (\sin \varphi)x_2 \\ y_2 &= (\sin \varphi)x_1 + (\cos \varphi)x_2 \end{aligned} \quad (2.85)$$

Die *Drehung* wird also durch die *Matrix*

$$\begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix}$$

beschrieben, und ist somit eine lineare Abbildung. Dies lässt sich aber auch unmittelbar geometrisch einsehen, s. Fig. 2.9, Fig. 2.10.

### Beispiel 2.30:

(*Drehungen im dreidimensionalen Raum*) Eine Drehung im Raum  $\mathbb{R}^3$  um den Winkel  $\varphi$  bzgl. einer Achse durch 0, die in Richtung von  $\mathbf{c} \in \mathbb{R}^3$  ( $|\mathbf{c}| = 1$ ) liegt, wird mit einer *Orthogonalbasis*  $(\mathbf{a}, \mathbf{b}, \mathbf{c})$  so beschrieben:

$$\mathbf{x} = \xi_1 \mathbf{a} + \xi_2 \mathbf{b} + \xi_3 \mathbf{c} \quad (\xi_1 = \mathbf{x} \cdot \mathbf{a}, \xi_2 = \mathbf{x} \cdot \mathbf{b}, \xi_3 = \mathbf{x} \cdot \mathbf{c})$$

geht über in

$$\mathbf{y} = f(\mathbf{x}) = (\xi_1 \cos \varphi - \xi_2 \sin \varphi) \mathbf{a} + (\xi_1 \sin \varphi + \xi_2 \cos \varphi) \mathbf{b} + \xi_3 \mathbf{c}. \quad (2.86)$$

Man erkennt die Analogie zum  $\mathbb{R}^2$ , s. (2.85). Natürlich ist diese Drehung eine lineare Abbildung, wie man geometrisch an Hand der Figuren 2.8 bis 2.10 sieht, die den »Blick« in Richtung  $\mathbf{c}$  zeigen.

Lineare Abbildungen auf unendlichdimensionalen Vektorräumen lassen sich nicht so einheitlich beschreiben, wie es bei endlichdimensionalen Räumen möglich ist (s. Beispiel 2.26). Im Folgenden werden einige typische Beispiele angegeben.

### Beispiel 2.31:

Die Differentiation einer Funktion  $f \in C^1(I)$  ordnet ihr die Funktion  $f' \in C(I)$  eindeutig zu. Diese Abbildung symbolisieren wir durch  $D$ , also  $Df = f'$ . Zweifellos ist  $D : C^1(I) \rightarrow C(I)$  eine lineare Abbildung. Sie ist bekanntlich nicht eineindeutig, da  $D(f + c) = Df$ , wenn  $c$  eine konstante Funktion bezeichnet.

Allgemeiner:

### Beispiel 2.32:

Ein »linearer Differentialoperator«  $L$  definiert durch

$$L(f) = a_0 + a_1 f + a_2 f' + a_3 f'' + \dots + a_{n+1} f^{(n)}$$

für alle  $f \in C^n(I)$  (mit  $a_i \in C(I)$ ) ist eine lineare Abbildung  $L : C^n(I) \rightarrow C(I)$ . Hier spielt die lineare Algebra in die Theorie der Differentialgleichungen hinein.

**Beispiel 2.33:**

Durch die beiden »Integraloperatoren«

$$(F(f))(x) := \int_a^b K(x, t) f(t) dt, \quad \text{»Fredholm-Typ«}^{26}$$

$$(T(f))(x) := \int_a^x K(x, t) f(t) dt, \quad \text{»Volterra-Typ«}^{27}$$

werden lineare Abbildungen  $f : C[a, b] \rightarrow C[a, b]$ , und  $T : C[a, b] \rightarrow C[a, b]$  definiert.  $K(x, t)$  wird dabei als (stückweise) stetig auf  $[a, b]^2$  vorausgesetzt.

In der Theorie der Integralgleichungen spielen diese linearen Abbildungen eine große Rolle.

**Übung 2.18\***

Es sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  eine Drehung um  $\varphi = 25,3^\circ$  (d.h.  $f(x)$  geht aus dem Ortsvektor  $x$  durch Drehung um  $25,3^\circ$  gegen den Uhrzeigersinn hervor). Gib die Matrix an, die diese Abbildung beschreibt.

**2.4.6 Isomorphismen, Konstruktion linearer Abbildungen**

Es seien im Folgenden  $V, W$  Vektorräume über demselben Körper  $\mathbb{K}$ .

**Definition 2.23:**

Eine lineare Abbildung  $f : V \rightarrow W$  heißt ein

*Isomorphismus*, wenn  $f$  bijektiv ist<sup>28</sup>

*Endomorphismus*, wenn  $V = W$  ist

*Automorphismus*, wenn  $f$  bijektiv und  $V = W$  ist.

**Bemerkung:** Gelegentlich spricht man auch von *Epimorphismus*, wenn  $f$  surjektiv ist, und *Monomorphismus*, falls  $f$  injektiv ist. Um nicht in einem Sumpf von Begriffen zu ertrinken, werden wir diese Begriffe nicht weiter verwenden.

Zwei Vektorräume  $V, W$  über  $\mathbb{K}$  heißen *isomorph*, in Zeichen  $V \simeq W$ , wenn es einen Isomorphismus  $f : V \rightarrow W$  gibt. Ist  $Z$  ein weiterer Vektorraum über  $\mathbb{K}$ , so gelten die Regeln:

$$V \simeq V, \quad V \simeq W \Leftrightarrow W \simeq V, \quad V \simeq W \text{ und } W \simeq Z \Rightarrow V \simeq Z.$$

Auf Grund dieser Gesetze nennt man  $\simeq$  eine *Äquivalenzrelation*.

<sup>26</sup> Erik Ivar Fredholm (1866–1927), schwedischer Mathematiker.

<sup>27</sup> Vito Volterra (1860–1940), italienischer Mathematiker und Physiker.

<sup>28</sup> Die Begriffe injektiv, surjektiv, bijektiv sind in Beisp. 2.9, Abschn. 2.3.2 erklärt, wie auch ausführlicher in Burg/Haf/Wille (Analysis) [27]

Gilt  $V \simeq W$ , so sagt man auch gelegentlich,  $W$  ist eine (isomorphe) *Kopie* von  $V$ . Dieser Ausdruck weist darauf hin, dass  $V$  und  $W$  nicht wesentlich verschieden sind.

**Satz 2.21:**

(*Konstruktion linearer Abbildungen*) Es seien  $V$  und  $W$  Vektorräume über  $\mathbb{K}$ , wobei  $(\mathbf{a}_1, \dots, \mathbf{a}_n)$  eine Basis von  $V$  ist und  $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$  eine Teilmenge von  $W$ . Damit folgt:

- (I) Es gibt genau eine lineare Abbildung  $f : V \rightarrow W$  mit  $f(\mathbf{a}_k) = \mathbf{b}_k$  für alle  $k = 1, \dots, n$ . Sie wird gebildet durch

$$f\left(\sum_{k=1}^n \lambda_k \mathbf{a}_k\right) := \sum_{k=1}^n \lambda_k \mathbf{b}_k \quad \text{Konstruktion einer linearen Abbildung. (2.87)}$$

- (II)  $f$  ist genau dann injektiv, wenn  $\mathbf{b}_1, \dots, \mathbf{b}_n$  linear unabhängig sind.  
 (III)  $f$  ist genau dann ein Isomorphismus, wenn  $(\mathbf{b}_1, \dots, \mathbf{b}_n)$  eine Basis von  $W$  ist.

**Beweis:**

(I) Man sieht ohne Schwierigkeit, dass  $f$  eine »lineare Abbildung« ist, indem man (H1), (H2) nachweist. Die Eindeutigkeit folgt so: Ist  $g$  ein weiterer Homomorphismus mit  $g(\mathbf{a}_k) = \mathbf{b}_k$  für alle  $k$ , so folgt wegen (H), (letzter Abschnitt), für alle  $\mathbf{x} = \sum_{k=1}^n \lambda_k \mathbf{a}_k \in V$ :

$$g(\mathbf{x}) = g\left(\sum_{k=1}^n \lambda_k \mathbf{a}_k\right) = \sum_{k=1}^n \lambda_k g(\mathbf{a}_k) = \sum_{k=1}^n \lambda_k \mathbf{b}_k = f(\mathbf{x}).$$

(II) Es sei  $f$  nicht injektiv. Dann existieren verschiedene  $\mathbf{x} = \sum_{k=1}^n \lambda_k \mathbf{a}_k$ ,  $\mathbf{x}' = \sum_{k=1}^n \lambda'_k \mathbf{a}_k$  in  $V$  mit  $f(\mathbf{x}) = f(\mathbf{x}')$ , d.h.  $f(\mathbf{x}) - f(\mathbf{x}') = f(\mathbf{x} - \mathbf{x}') = \mathbf{0}$ . Da  $\mathbf{x} \neq \mathbf{x}'$ , also  $\mathbf{z} := \mathbf{x} - \mathbf{x}' \neq \mathbf{0}$  ist, so folgt  $f(\mathbf{z}) = \mathbf{0}$  mit einem  $\mathbf{z} \neq \mathbf{0}$  aus  $V$ . Es sei  $\mathbf{z} = \sum_{k=1}^n \zeta_k \mathbf{a}_k$ , also  $\mathbf{0} = f(\mathbf{z}) = \sum_{k=1}^n \zeta_k \mathbf{b}_k = \mathbf{0}$ . Da nicht alle  $\zeta_k = 0$  sind (wegen  $\mathbf{z} \neq \mathbf{0}$ ), sind damit die  $\mathbf{b}_k$  linear abhängig. — Sind umgekehrt die  $\mathbf{b}_1, \dots, \mathbf{b}_n$  als linear abhängig vorausgesetzt, so gibt es  $\lambda_i \in \mathbb{K}$  (nicht alle 0), mit  $\sum_{i=1}^n \lambda_i \mathbf{b}_i = \mathbf{0}$ . Für  $\mathbf{x} := \sum_{i=1}^n \lambda_i \mathbf{a}_i$  folgt damit  $f(\mathbf{x}) = \sum_{i=1}^n \lambda_i \mathbf{b}_i = \mathbf{0}$ . Es ist also  $f(\mathbf{x}) = \mathbf{0}$  und  $f(\mathbf{0}) = \mathbf{0}$  mit  $\mathbf{x} \neq \mathbf{0}$ . D.h.  $f$  ist nicht injektiv. Damit ist (II) bewiesen. — (III) folgt unmittelbar aus (II).  $\square$

**Bemerkung:** Wir wollen die *Konstruktionsvorschrift* (2.87) für lineare Abbildungen von endlich-dimensionalen Vektorräumen noch einmal hervorheben: Ordnet man den Basisvektoren eines Vektorraums  $V$  eindeutig Vektoren eines anderen Vektorraums  $W$  zu, so ist damit sofort eine lineare Abbildung gegeben (durch (2.87)). Auf diese Weise lassen sich *alle* linearen Abbildungen von  $V$  (mit  $\dim V$  endlich) in  $W$  beschreiben. Die Dimension von  $W$  ist dabei beliebig, sie kann also auch  $\infty$  sein.

Dem Teil (III) in Satz 2.21 können wir noch folgende schöne Formulierung geben:

**Folgerung 2.12:**

Zwei endlichdimensionale Vektorräume  $V, W$  über  $\mathbb{K}$  sind genau dann isomorph, wenn  $\dim V = \dim W$  ist.

Insbesondere gelangen wir damit zu der nützlichen

**Folgerung 2.13:**

Jeder Vektorraum  $V$  über  $\mathbb{K}$  mit endlicher Dimension  $n \in \mathbb{N}$  ist isomorph zu  $\mathbb{K}^n$ .

*Ausführlich:* Ist  $B = (\mathbf{a}_1, \dots, \mathbf{a}_n)$  eine Basis von  $V$ , so wird jedem Vektor  $\mathbf{x} = \sum_{k=1}^n x_k \mathbf{a}_k \in V$  (mit  $x_k \in \mathbb{K}$ ) der Vektor

$$\mathbf{x}_B = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{K}^n$$

zugeordnet. Die dadurch erklärte Abbildung  $f(\mathbf{x}) = \mathbf{x}_B$  ( $f : V \rightarrow \mathbb{K}^n$ ) ist ein Isomorphismus.  $\mathbf{x}_B$  heißt auch der *numerische Vektor* zu  $\mathbf{x}$  bzgl. der Basis  $B$ .

Insbesondere im Fall  $\mathbb{K} = \mathbb{R}$  macht man viel Gebrauch hiervon, da man den  $\mathbb{R}^n$  so gut kennt. Die folgenden Beispiele geben Isomorphismen an, besonders in Verbindung mit Satz 2.21 (III).

**Beispiel 2.34:**

Zwei Ebenen  $E_1, E_2$  im  $\mathbb{R}^3$ , die durch  $\mathbf{0}$  gehen, sind *isomorphe* Vektorräume über  $\mathbb{R}$ . Denn sind

$$\mathbf{x} = \lambda_1 \mathbf{a}_1 + \lambda_2 \mathbf{a}_2, \quad \mathbf{y} = \mu_1 \mathbf{b}_1 + \mu_2 \mathbf{b}_2$$

die Parameterdarstellungen von  $E_1$  und  $E_2$  ( $(\mathbf{a}_1, \mathbf{a}_2)$  Basis von  $E_1$ ,  $(\mathbf{b}_1, \mathbf{b}_2)$  Basis von  $E_2$ ), so wird durch folgende Abbildung eine Isomorphie von  $E_1$  auf  $E_2$  vermittelt:

$$f(\lambda_1 \mathbf{a}_1 + \lambda_2 \mathbf{a}_2) := \lambda_1 \mathbf{b}_1 + \lambda_2 \mathbf{b}_2.$$

**Beispiel 2.35:**

Der Lösungsraum  $V$  der linearen Differentialgleichung  $y^{(n)} = a_1 y + a_2 y' + \dots + a_n y^{(n-1)}$  ist  $n$ -dimensional. Er ist zum  $\mathbb{R}^n$  *isomorph*, denn ist  $y_1, \dots, y_n$  ein Fundamentalsystem<sup>29</sup> von Lösungen (d.h. eine Basis von  $V$ ), so ist folgende Abbildung ein Isomorphismus von  $V$  auf  $\mathbb{R}^n$ :

$$f\left(\sum_{i=1}^n \lambda_i y_i\right) := \sum_{i=1}^n \lambda_i \mathbf{e}_i = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}.$$

---

<sup>29</sup> s. Burg/Haf/Wille (Band III) [25]



### 2.4.7 Kern, Bild, Rang

$V, W$  seien Vektorräume über  $\mathbb{K}$ .

#### Definition 2.24:

Ist  $f : V \rightarrow W$  eine lineare Abbildung, so sind *Kern*, *Bild* und *Rang* dieser Abbildung erklärt durch

$\text{Kern } f = \text{Menge aller } x \in V \text{ mit } f(x) = \mathbf{0}$

$\text{Bild } f = \text{Menge aller } f(x) \text{ mit } x \in V$  <sup>30</sup>

$\text{Rang } f = \dim \text{Bild } f$ .

#### Satz 2.22:

$f : V \rightarrow W$  sei eine lineare Abbildung. Damit gilt

- (a) Kern  $f$  ist ein Unterraum von  $V$
- (b) Bild  $f$  ist ein Unterraum von  $W$
- (c)  $f$  ist genau dann injektiv, wenn Kern  $f = \{\mathbf{0}\}$  ist.

#### Beweis:

(a), (b) überlegt sich der Leser leicht selbst. Zu (c): Es sei Kern  $f = \{\mathbf{0}\}$ . Angenommen,  $f$  ist nicht injektiv. Dann gibt es zwei Vektoren  $x_1, x_2 \in V$  mit  $x_1 \neq x_2$  und  $f(x_1) = f(x_2)$ . Daraus folgt  $f(x_1) - f(x_2) = \mathbf{0}$ , also  $f(x_1 - x_2) = \mathbf{0}$ , also  $f(z) = \mathbf{0}$  mit  $z = x_1 - x_2 \neq \mathbf{0}$ , also  $z \in \text{Kern } f$ , was der Voraussetzung widerspricht. Also ist  $f$  injektiv. — Ist umgekehrt  $f$  als injektiv vorausgesetzt, so kann  $f(x) = \mathbf{0}$  nur die Lösung  $\mathbf{0}$  haben, also folgt Kern  $f = \{\mathbf{0}\}$ .  $\square$

Damit stoßen wir zum krönenden Satz vor:

#### Satz 2.23:

(Dimensionsformel) Für jede lineare Abbildung  $f : V \rightarrow W$  gilt:

$$\dim \text{Kern } f + \dim \text{Bild } f = \dim V \quad ^{31} . \quad (2.88)$$

#### Beweis:

1. Fall: Die Dimensionen von Kern  $f$  und Bild  $f$  sind beide endlich. Ist Kern  $f = \{\mathbf{0}\}$ , so folgt die Behauptung aus Satz 2.22 (c). Ist Bild  $f = \{\mathbf{0}\}$ , so ist Kern  $f = V$ , also (2.88) auch richtig. Wir nehmen darum im Folgenden Kern  $f \neq \{\mathbf{0}\}$  und Bild  $f \neq \{\mathbf{0}\}$  an. Ferner seien  $a_1, \dots, a_q, b_1, \dots, b_p \in V$  gegeben mit:

$$\begin{aligned} (a_1, \dots, a_q) & \text{ ist eine Basis von Kern } f \\ (f(b_1), \dots, f(b_p)) & \text{ ist eine Basis von Bild } f. \end{aligned} \quad (2.89)$$

<sup>30</sup> Bild  $f$  ist also nichts anderes als der Wertebereich  $f(V)$  der Abbildung, s. Burg/Haf/Wille (Analysis) [27]

<sup>31</sup> Mit  $\infty$  wird in diesem Fall so gerechnet:  $\infty + \infty = \infty$ ,  $\infty + n = n + \infty = \infty$  für alle  $n \in \mathbb{N}_0$ .

**Behauptung:**  $\mathbf{a}_1, \dots, \mathbf{a}_q, \mathbf{b}_1, \dots, \mathbf{b}_p$  ist eine Basis von  $V$  (womit die Dimensionsformel gilt).

**Beweis der Behauptung:**  $\mathbf{a}_1, \dots, \mathbf{a}_q, \mathbf{b}_1, \dots, \mathbf{b}_p$  sind linear unabhängig, da  $\sum_{i=1}^q \lambda_i \mathbf{a}_i + \sum_{i=1}^p \mu_i \mathbf{b}_i =$

$\mathbf{0}$  nach Anwendung von  $f$  auf  $\sum_{i=1}^p \mu_i f(\mathbf{b}_i) = \mathbf{0}$  führt, also  $\mu_i = 0$  für alle  $i = 1, \dots, p$ , folglich

$\sum_{i=1}^q \lambda_i \mathbf{a}_i = \mathbf{0}$  und damit auch  $\lambda_i = 0$  für alle  $i = 1, \dots, q$ .

Ist ferner  $\mathbf{x} \in V$  beliebig, so gilt  $f(\mathbf{x}) \in \text{Bild } f$ , also  $f(\mathbf{x}) = \sum_{i=1}^p \beta_i f(\mathbf{b}_i)$  (mit  $\beta_i \in \mathbb{K}$ ) und daher

$$\mathbf{0} = f(\mathbf{x}) - \sum_{i=1}^p \beta_i f(\mathbf{b}_i) = f\left(\mathbf{x} - \sum_{i=1}^p \beta_i \mathbf{b}_i\right),$$

d.h.  $\mathbf{x} - \sum_{i=1}^p \beta_i \mathbf{b}_i \in \text{Kern } f$ , also  $\mathbf{x} - \sum_{i=1}^p \beta_i \mathbf{b}_i = \sum_{k=1}^q \alpha_k \mathbf{a}_k$ , d.h.  $\mathbf{x}$  ist eine Linearkombination der  $\mathbf{a}_i, \mathbf{b}_k$ . Die  $\mathbf{a}_1, \dots, \mathbf{a}_q, \mathbf{b}_1, \dots, \mathbf{b}_p$  spannen also  $V$  auf und bilden somit eine Basis von  $V$ .

2. Fall:  $\dim \text{Kern } f = \infty$ . Wegen  $\text{Kern } f \subset V$  ist dann auch  $\dim V = \infty$  und die Dimensionsformel (2.88) gilt.

3. Fall:  $\dim \text{Bild } f = \infty$ . Angenommen, die Dimension von  $V$  ist endlich, und  $(\mathbf{a}_1, \dots, \mathbf{a}_n)$  sei eine Basis von  $V$ . Dann spannen die Vektoren  $f(\mathbf{a}_i) = \mathbf{b}_i$  ( $i = 1, \dots, n$ )  $\text{Bild } f$  auf (nach Satz 2.21 (I)), also folgt  $\dim \text{Bild } f < \dim V = n$ , im Widerspruch zu  $\dim \text{Bild } f = \infty$ . Folglich gilt  $\dim V = \infty$  und damit die Dimensionsformel (2.88).  $\square$

#### Folgerung 2.14:

(Äquivalenz von Injektivität und Surjektivität) Es sei  $f : V \rightarrow W$  eine lineare Abbildung, wobei die Vektorräume  $V$  und  $W$  die gleiche endliche Dimension  $n$  haben. Damit gilt:

$f$  ist genau dann injektiv, wenn  $f$  surjektiv ist.

#### Beweis:

$f$  injektiv  $\Leftrightarrow \text{Kern } f = \{\mathbf{0}\}$  (s. Satz 2.22 (c))  $\Leftrightarrow \dim \text{Kern } f = 0 \Leftrightarrow \dim \text{Bild } f = n$  (wegen der Dimensionsformel (2.88))  $\Leftrightarrow \text{Bild } f = V \Leftrightarrow f$  surjektiv.  $\square$

**Bemerkung:** Zusammenhang mit linearen Gleichungssystemen. Ein lineares Gleichungssystem

$$y_i = \sum_{k=1}^n a_{ik} x_k, \quad i = 1, \dots, m, \quad (2.90)$$

mit gegebenen  $a_{ik} \in \mathbb{K}$ ,  $y_i \in \mathbb{K}$  und gesuchten  $x_k$  kann als lineare Abbildung  $f : \mathbb{K}^n \rightarrow \mathbb{K}^m$  interpretiert werden, wobei jedem  $\mathbf{x} = [x_1, \dots, x_n]^T$  ein  $\mathbf{y} = [y_1, \dots, y_m]^T$  zugeordnet wird.

*Lösbarkeit liegt genau dann vor, wenn  $y \in \text{Bild } f$  ist, und eindeutige Lösbarkeit genau dann, wenn zusätzlich  $\dim \text{Bild } f = \dim V = n$  (d.h.  $\dim \text{Kern } f = 0$ ). Diese Aussagen spiegeln genau das Rangkriterium Satz 2.10, Abschn. 2.2.5, in anderer Formulierung wieder.*

Die Folgerung 2.14, die besagt: »Entweder bijektiv, oder weder injektiv noch surjektiv«, entspricht im Fall der Gleichungssysteme der Folgerung 2.5, Abschn. 2.2.4: *Im Fall  $n = m$  ist das Gleichungssystem (2.90) entweder für jedes  $y \in \mathbb{K}^n$  eindeutig lösbar, oder für keins.*

Zur praktischen Lösung, auch bei beliebigen Körpern  $\mathbb{K}$ , benutzt man aber in den allermeisten Fällen den guten alten Gauß-Algorithmus, der in  $\mathbb{K}$  genau wie in  $\mathbb{R}$  verläuft (wenn man die Pivotierung nach Größe von Absolutbeträgen außer Acht lässt, sondern sich mit Diagonalelementen  $\neq 0$  begnügt).

### Übung 2.19\*

Durch  $y_1 = 6x_1 + 4x_2 - 10x_3$ ,  $y_2 = -9x_1 - 6x_2 + 15x_3$  ist eine lineare Abbildung  $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$  gegeben. Berechne Kern  $f$  und Bild  $f$ , d.h. gib für beide Räume Basen an. Welche Werte haben  $\dim \text{Kern } f$  und  $\text{Rang } f$ ? Rechne die Dimensionsformel (2.88) für dieses Beispiel nach.

### Übung 2.20\*

Durch  $L(y) = y' - 2y$  ist ein »Differentialoperator« für alle  $y \in C^1(\mathbb{R})$  erklärt.  $L: C^1(\mathbb{R}) \rightarrow C(\mathbb{R})$  ist eine lineare Abbildung (überprüfe das!). Welche Funktionen  $y \in C^1(\mathbb{R})$  liegen im Kern von  $L$ ?

## 2.4.8 Euklidische Vektorräume, Orthogonalität

### Definition 2.25:

Es sei  $V$  ein Vektorraum über  $\mathbb{R}$ . Eine Vorschrift, die jedem Paar  $(x, y)$  mit  $x, y \in V$  genau eine reelle Zahl  $r$  zuordnet, beschrieben durch

$$r = x \cdot y,$$

heißt ein *inneres Produkt* auf  $V$ , wenn folgende Gesetze erfüllt sind: Für alle  $x, y, z \in V$  und alle  $\lambda, \mu \in \mathbb{R}$  gilt

- |       |  |                              |
|-------|--|------------------------------|
| (I)   | $x \cdot y = y \cdot x$  | <i>Kommutativgesetz</i>      |
| (II)  | $(x + y) \cdot z = x \cdot z + y \cdot z$                        | <i>Distributivgesetz</i>     |
| (III) | $\lambda(x \cdot y) = (\lambda x) \cdot y = x \cdot (\lambda y)$ | <i>Assoziativgesetz</i>      |
| (IV)  | $x \neq 0 \Leftrightarrow x \cdot x > 0$                         | <i>positive Definitheit.</i> |

Man nennt  $|x| := \sqrt{x \cdot x}$  die *Länge* (den *Betrag*, die *euklidische Norm*) von  $x$ . Für  $x \cdot x$  schreibt man kürzer  $x^2$ .

Ist auf  $V$  ein inneres Produkt wie oben erklärt, so nennt man  $V$  einen *euklidischen Vektorraum* (oder *Prä-Hilbertraum*).

Wie in Abschn. 2.1.2, Satz 2.1 (V)–(IX) beweist man, nur unter Verwendung von (I)–(IV)

**Folgerung 2.15:**

Ist  $V$  ein euklidischer Vektorraum, so gilt für alle  $\mathbf{x}, \mathbf{y} \in V$  und alle  $\lambda \in \mathbb{R}$

$$\begin{aligned} \text{(V)} \quad & |\lambda \mathbf{x}| = |\lambda| |\mathbf{x}|, & \text{(VI)} \quad & |\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}| |\mathbf{y}|, \\ \text{(VII)} \quad & |\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|, & \text{(VIII)} \quad & |\mathbf{x} - \mathbf{y}| \geq ||\mathbf{x}| - |\mathbf{y}||, \\ \text{(IX)} \quad & |\mathbf{x}| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}. \end{aligned}$$

**Beispiel 2.36:**

Für alle  $f, g \in C([a, b])$  ist folgendermaßen ein inneres Produkt erklärt

$$f \cdot g = \int_a^b f(x)g(x) \, dx. \quad (2.91)$$

**Beispiel 2.37:**

Ist  $(\mathbf{a}_1, \dots, \mathbf{a}_n)$  eine Basis des Vektorraums  $V$  über  $\mathbb{R}$ , und sind

$$\mathbf{x} = \sum_{i=1}^n x_i \mathbf{a}_i, \quad \mathbf{y} = \sum_{i=1}^n y_i \mathbf{a}_i \quad (x_i, y_i \in \mathbb{R})$$

beliebig aus  $V$ , so ist durch  $\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^n x_i y_i$  zweifellos ein inneres Produkt von  $V$  gegeben. Auf diese Weise lassen sich auch im  $\mathbb{R}^n$  verschiedene innere Produkte einführen.

**Definition 2.26:**

Ist  $V$  ein euklidischer Vektorraum, so erklärt man den *Winkel* zwischen zwei Elementen  $\mathbf{a}, \mathbf{b} \in V$ ,  $\mathbf{a} \neq \mathbf{0}$ ,  $\mathbf{b} \neq \mathbf{0}$ , durch

$$\varphi = \angle(\mathbf{a}, \mathbf{b}) := \arccos \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}| \cdot |\mathbf{b}|}. \quad (2.92)$$

Ist  $\mathbf{a} = \mathbf{0}$  oder  $\mathbf{b} = \mathbf{0}$ , so kann  $\angle(\mathbf{a}, \mathbf{b})$  jede beliebige Zahl aus  $[0, \pi]$  bedeuten. Man sagt, die Elemente  $\mathbf{a}, \mathbf{b} \in V$  stehen *rechtwinklig* (*orthogonal*) aufeinander:  $\mathbf{a} \perp \mathbf{b}$ , wenn  $\mathbf{a} \cdot \mathbf{b} = 0$  ist.

Damit folgen *Pythagoras*  $(\mathbf{a} + \mathbf{b})^2 = \mathbf{a}^2 + \mathbf{b}^2 \Leftrightarrow \mathbf{a} \cdot \mathbf{b} = 0$ , und *Cosinussatz*  $(\mathbf{a} - \mathbf{b})^2 = \mathbf{a}^2 + \mathbf{b}^2 - 2|\mathbf{a}| \cdot |\mathbf{b}| \cos \angle(\mathbf{a}, \mathbf{b})$  durch schlichtes Ausmultiplizieren (wie in Abschn. 2.1.4).

*Orthonormalsystem*, *Orthonormalbasis* und *orthogonales Komplement* eines Vektorraumes werden wie in Abschn. 2.1.4 definiert (man hat dort nur  $V$  statt  $\mathbb{R}^n$  zu setzen).

Ebenso funktioniert in einem endlichdimensionalen euklidischen Vektorraum  $V$  das Schmidt'sche Orthogonalisierungsverfahren, und es gelten Satz 2.6, Satz 2.7 und Folgerung 2.3 aus Abschn. 2.1.4 entsprechend (man ersetze dort einfach  $\mathbb{R}^n$  durch  $V$ ).

**Übung 2.21:**

Beweise, dass (2.91) ein inneres Produkt in  $C([a, b])$  beschreibt, d.h. weise nach, dass alle Eigenschaften (I)–(IV) in Definition 2.25 erfüllt sind.

**Übung 2.22\***

- (a) Welchen Winkel bilden  $\sin(kx)$  und  $\cos(nx)$  ( $n, k \in \mathbb{N}, n \neq k$ ) in  $C([0, 2\pi])$  miteinander, wenn das innere Produkt entsprechend (2.91) erklärt ist.
- (b) Welchen Winkel bilden  $f(x) = x$  und  $g(x) = x^2$  in diesem Raum miteinander?

**Übung 2.23\***

Wende auf die Polynome  $f_0(x) = 1$ ,  $f_1(x) = x$ ,  $f_2(x) = x^2$  in  $C([-1, 1])$  das Schmidtsche Orthogonalisierungsverfahren an.

**Übung 2.24\***

Zeige: Zu jeder linearen Abbildung  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  gibt es einen Vektor  $v \in \mathbb{R}^n$  mit  $f(x) = v \cdot x$  für alle  $x \in \mathbb{R}^n$ .

**Übung 2.25\***

Es sei  $|f| = \sqrt{f \cdot f}$  für  $f \in C([a, b])$  (inneres Prod. wie in (2.91)) und  $\|f\|_\infty := \sup_{x \in [a, b]} |f(x)|$ .

- (a) Zeige:

$$\lim_{n \rightarrow \infty} \|f_n\|_\infty = 0 \Rightarrow \lim_{n \rightarrow \infty} |f_n| = 0.$$

- (b) Zeige, dass die Umkehrung nicht gilt. D.h. gib eine Funktionenfolge  $(f_n)$  aus  $C([a, b])$  an mit  $|f_n| \rightarrow 0$  für  $n \rightarrow \infty$ , aber *nicht*  $\|f_n\|_\infty \rightarrow 0$  für  $n \rightarrow \infty$ !

**2.4.9 Ausblick auf die Funktionalanalysis**

<sup>32</sup>Die Funktionalanalysis verknüpft Analysis und lineare Algebra. Insbesondere die Funktionenräume werden dabei wichtig. — Wir beginnen mit einer Verallgemeinerung des euklidischen Raumes:

**Definition 2.27:**

Ein Vektorraum  $V$  über  $\mathbb{R}$  (oder  $\mathbb{C}$ ) heißt ein *normierter linearer Raum*, wenn zu jedem  $x \in V$  eine nichtnegative reelle Zahl erklärt ist, so dass Folgendes für alle  $x, y \in V$  und  $\lambda \in \mathbb{R}$  (oder  $\lambda \in \mathbb{C}$ ) gilt:

$$\|\lambda x\| = |\lambda| \|x\|, \quad \|x + y\| \leq \|x\| + \|y\|, \quad \|x\| = 0 \Leftrightarrow x = \mathbf{0}. \quad (2.93)$$

$\|x\|$  heißt die *Norm* (oder *Länge*) von  $x$ .

<sup>32</sup> Ausführlich beschrieben in Burg/Haf/Wille (Partielle Dgl.) [26]

Man sieht, jeder *euklidische Raum* ist auch ein *normierter linearer Raum* (mit  $\|\mathbf{x}\| := |\mathbf{x}|$ ). Umgekehrtes braucht nicht zu gelten. Das wichtigste Beispiel eines normierten linearen Raumes ist  $C([a, b])$  mit der Norm

$$\|f\| := \|f\|_\infty = \sup_{x \in [a, b]} |f(x)|.$$

Man weist die Normgesetze (2.93) leicht nach. Mit dieser Norm ist  $C([a, b])$  kein euklidischer Raum, d.h. es gibt kein inneres Produkt mit  $\|f\| = \sqrt{f \cdot f}$  in  $C([a, b])$ .

Mit der Analysis wird der Zusammenhang folgendermaßen hergestellt: Man betrachtet in einem normierten Raum  $V$  (unendliche) Folgen  $(\mathbf{a}_n)_{n \in \mathbb{N}}$ , kurz  $(\mathbf{a}_n)$  geschrieben. Das Element  $\mathbf{a}$  heißt *Grenzwert* von  $(\mathbf{a}_n)$ , wenn

$$\|\mathbf{a}_n - \mathbf{a}\| \rightarrow 0 \quad \text{für } n \rightarrow \infty$$

gilt. In diesem Falle sagt man,  $(\mathbf{a}_n)$  konvergiert gegen  $\mathbf{a}$ .

$(\mathbf{a}_n)$  heißt eine *Cauchy-Folge*, wenn zu jedem  $\varepsilon > 0$  ein Index  $n_0 \in \mathbb{N}$  existiert, so dass für alle  $\mathbf{a}_n, \mathbf{a}_m$  mit  $n, m \geq n_0$  gilt:

$$\|\mathbf{a}_n - \mathbf{a}_m\| < \varepsilon.$$

(Vgl. Cauchysches Konvergenzkriterium, Burg/Haf/Wille (Analysis) [27])

**Definition 2.28:**

- (a) Ein normierter linearer Raum heißt *vollständig*, wenn jede Cauchy-Folge aus dem Raum gegen einen Grenzwert in diesem Raum konvergiert.
- (b) Ein vollständiger normierter linearer Raum heißt *Banachraum*.
- (c) Ein vollständiger euklidischer Raum heißt *Hilbertraum* (mit der Norm  $\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$ ).

Natürlich ist  $\mathbb{R}^n$  ein Hilbertraum und damit auch ein Banachraum. Anspruchsvollere Beispiele sind folgende:

**Beispiel 2.38:**

$C([a, b])$  ist ein Banachraum bzgl.  $\|f\|_\infty$ . (Denn ist  $(f_n)$  aus  $C([a, b])$  eine Cauchy-Folge, so ist  $(f_n)$  gleichmäßig konvergent und hat somit einen Grenzwert  $f \in C([a, b])$  (nach Burg/Haf/Wille (Analysis) [27])).

**Beispiel 2.39:**

$C^k([a, b])$ , die Menge der  $k$ -mal stetig differenzierbaren Funktionen  $f : [a, b] \rightarrow \mathbb{R}$  ist bzgl.

$$\|f\| = \sum_{i=0}^k \sup_{x \in [a, b]} |f^{(i)}(x)| \quad (f^{(i)} \text{ } i\text{-te Ableitung, } f^{(0)} = f)$$

ein Banachraum. (Dies folgt aus Burg/Haf/Wille (Analysis) [27] durch vollständige Induktion.)

**Beispiel 2.40:**

Der Folgenraum  $\ell^p$  ( $p > 1$ ), erklärt in Beispiel 2.24, Abschn. 2.4.2, ist ein *Banachraum* mit der in Beispiel 2.24 angegebenen Norm  $|\mathbf{a}|_p$  ( $= \|\mathbf{a}\|$ ). (Für den Nachweis der Normgesetze, insbesondere der Dreiecksungleichung  $|\mathbf{a} + \mathbf{b}|_p \leq |\mathbf{a}|_p + |\mathbf{b}|_p$  und der Vollständigkeit wird auf die Literatur über Funktionalanalysis verwiesen (z.B. [76], Kap. II § 4, S. 56–57).

*Spezialfall:*  $p = 2$ : Der Raum  $\ell^2$  ist ein Hilbertraum mit dem in Beispiel 2.23, Abschn. 2.4.2 erklärten inneren Produkt.  $\ell^2$  kann als der einfachste unendlichdimensionale Hilbertraum angesehen werden. (Entsprechend werden Räume  $\ell^p$  aus komplexen Zahlenfolgen gebildet.)

**Beispiel 2.41:**

Die Menge  $L^p[a, b]$  aller Funktionen  $f : [a, b] \rightarrow \mathbb{R}$ , für die das folgende Integral existiert:

$$\int_a^b |f(x)|^p dx \quad \text{stellt bzgl.} \quad \|f\|_p := \sqrt[p]{\int_a^b |f(x)|^p dx} \quad (p \geq 1)$$

und den üblichen Operatoren  $+$  und  $\lambda \cdot$  ( $\lambda \in \mathbb{R}$ ) stellt bei Funktionen einen Banachraum dar. Hierbei sind die oben auftretenden Integrale im *Sinne von Lebesgue*<sup>33</sup> zu verstehen (s. hierzu beispielsweise [63] oder [75]). Wir heben ausdrücklich hervor: Würde man in  $C[a, b]$  die obige  $\|\cdot\|_p$ -Norm unter Verwendung des bekannten *Riemann-Integrals* (s. Burg/Haf/Wille (Analysis) [27]) einführen, dann käme man zwar zu einem normierten Raum, aber dieser Raum wäre *nicht vollständig*.

Der Spezialfall  $p = 2$  ist von besonderer Bedeutung, weil  $L^2[a, b]$  ein *Hilbertraum* ist. In der Theorie der Fourierreihen spielt der Raum der  $2\pi$ -periodischen und über dem Intervall  $[0, 2\pi]$  im Sinne von Lebesgue quadratisch integrierbaren Funktionen eine besondere Rolle. Ist nämlich  $f$  eine solche Funktion und ist

$$\left[ \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(kx) + b_k \sin(kx)) \right]$$

die Fourier-Reihe von  $f$ , dann kann man ihr die Folge

$$\mathbf{c}_f = (a_0, a_1, b_1, a_2, b_2, \dots)$$

der Fourier-Koeffizienten zuordnen:

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(kt) dt, \quad k = 0, 1, 2, \dots$$

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin(kt) dt, \quad k = 1, 2, \dots$$

33 Henri Lebesgue (1875–1941), französischer Mathematiker

Es lässt sich zeigen, dass die Reihe

$$a_0^2 + \sum_{k=1}^{\infty} (a_k^2 + b_k^2)$$

konvergiert und mit  $\|f\|_2^2$  übereinstimmt. Mehr noch: Die Zuordnung  $f \mapsto \mathbf{c}_f$ , symbolisiert durch  $F(f) := \mathbf{c}_f$ , ist ein *Isomorphismus* (s. Abschn. 2.4.6)

$$F : L^2[0, \pi] \rightarrow \ell^2.$$

Damit spiegelt der Raum  $\ell^2$  alle Fourier-Reihen von  $2\pi$ -periodischen Funktionen aus  $L^2[0, 2\pi]$  wieder und ist somit ein brauchbares Hilfsmittel zum Studium dieser Reihen.

**Bemerkung:** Die Funktionenräume  $C^k(I)$ ,  $L^p(I)$  und andere spielen bei Differentialgleichungen, Integralgleichungen und Fourier-Reihen eine wichtige Rolle. Im Teil »Funktionalanalysis« von Burg/Haf/Wille (Partielle Dgl.) [26] wird ausführlich darauf eingegangen. Ergänzend wird der Leser auf die Literatur über Funktionalanalysis und Differentialgleichungen verwiesen (z.B. [1], [30], [63], [76], [90], [148]).

### Übung 2.26\*

Entspanne dich!



Höhere Mathematik für Ingenieure Band II

Lineare Algebra

Burg, K.; Haf, H.; Wille, F.; Meister, A.

2012, XVII, 417 S. 119 Abb., Softcover

ISBN: 978-3-8348-1853-9