ulm university universität uulm

Vorabskript zur Vorlesung

Lineare Algebra I

Sommersemester 2014

Prof. Dr. Helmut Maier
Dr. Hans- Peter Reck

Institut für Zahlentheorie und Wahrscheinlichkeitstheorie
Universität Ulm

Inhaltsverzeichnis

1 Einleitung **3**

1.1 Vorbemerkung . . . 3

1.2 Ebene und Raum . . . 4

1.3 Der $\mathbb{R}^n$ . . . 9

1.4 Lineare Gleichungssysteme . . . 10

1.5 Verknüpfungen und Gruppen . . . 14

1.6 Ebene und Raum als reelle Vektorräume, Unterraum . . . 19

1.7 Lineare Unabhängigkeit, Basis, Dimension . . . 25

2 Vektorräume **28**

2.1 Ringe und Körper . . . 28

2.2 Der Körper der komplexen Zahlen . . . 29

2.3 Der allgemeine Begriff des Vektorraums . . . 31

2.4 Lineare Abhängigkeit, Basis, Dimension . . . 34

3 Lineare Abbildungen und Matrizen **41**

3.1 Lineare Abbildungen . . . 41

3.2 Kern und Bild . . . 43

3.3 Lineare Fortsetzung . . . 46

3.4 Isomorphismen . . . 47

3.5 Lineare Abbildungen und Matrizen . . . 49

3.6 Berechnung des Rangs einer Matrix . . . 58

3.7 Basiswechsel . . . 60

4 Lineare Gleichungen **65**

4.1 Theorie der Linearen Gleichungen . . . 65

Kapitel 1

Einleitung

1.1 Vorbemerkung

Die Lineare Algebra ist zusammen mit der Analysis eine der mathematischen Grunddisziplinen. Beide sind am Aufbau aller weiterführender Disziplinen in unterschiedlichem Maße beteiligt. Die Lineare Algebra hat sich aus der Elementargeometrie, in der Figuren der Ebene und des dreidimensionalen Raumes untersucht werden, entwickelt. Sie spielen jedoch auch in Gebieten eine Rolle, die auf den ersten Blick keinen Zusammenhang zur Elementargeometrie zu besitzen scheinen. Viele zentralen Ideen und Begriffe der Linearen Algebra treten schon in einfacher Form in der Geometrie der Ebene und des Raumes auf. Hier ist die geometrische Anschauung sehr hilfreich. Sie hilft beim Verständnis der Begriffe und beim Finden von Beweisen. Die geometriSche AnschAuung selbst hat jedoch keinerlei Beweiskraft.

Wir werden den Aufbau der Theorie nach den Prinzipien der modernen Mathematik vornehmen. Eine mathematische Theorie besteht aus folgenden Bestandteilen:

- Axiome:
  Dies sind (im allgemeinen wenige) Grundtatsachen, die ohne Beweis als wahr angenommen werden. Auch die Grundbegriffe, von denen in den Axiomen die Rede ist, werden nicht weiter erklärt.

- Definitionen:
  Es ist notwendig, Namen für die Objekte der Theorie und die Eigenschaften dieser Objekte einzuführen. Definitionen sind keine Aussagen, die wahr oder falsch sein können, sondern Vereinbarungen über diese Namensgebung. So ist zum Beispiel 2 ein Name für die Summe $1+1$.

- Lehrsätze:
  Dies sind Aussagen, die aus den Axiomen und schon bekannten Lehrsätzen durch eine Kette von logischen Schlüssen, Beweisen genannt, hergeleitet werden.

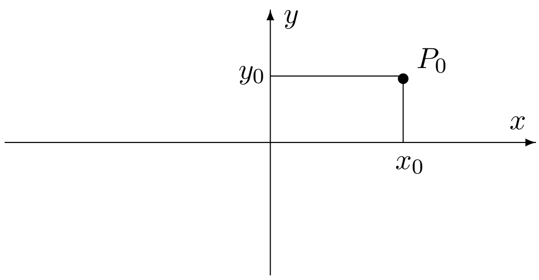
Das Konzept, das ganz am Anfang der Linearen Algebra steht, ist der Begriff des Vektorraumes. Auch in dieser Einleitung steht dieser Begriff im Mittelpunkt. Wir werden hier jedoch nur Vektorräume über dem Körper der reellen Zahlen- reelle Vektorräume- betrachten, die viele Aspekte, mit den aus der Elementargeometrie bekannten Spezialfällen Ebene und Raum gemeinsam haben. Manche der Definitionen und Sätze der späteren Kapitel werden die der Einleitung nur leicht verallgemeinern.

1.2 Ebene und Raum

Wir setzen die Menge $\mathbb{R}$ der reellen Zahlen mit den Rechenoperationen Addition, Subtraktion, Multiplikation und Division als bekannt voraus. Unter $\mathbb{R}^2$ verstehen wir die Menge aller Paare reeller Zahlen:

$$\mathbb{R}^2 = \{(x, y) \colon x, y \in \mathbb{R}\}.$$

Die Ebene E, die wir uns zum Beispiel als Zeichenebene vorstellen, kann durch Einführung eines kartesischen Koordinatensystems mit dem $\mathbb{R}^2$ identifiziert werden. Ein kartesisches Koordinatensystem entsteht durch Vorgabe eines Punktes 0 und einer Zahlengeraden, die wir x-Achse nennen, mit dem Nullpunkt 0. Die y-Achse entsteht durch eine positive Drehung (gegen den Uhrzeigersinn) um $90°$ um den Punkt 0 aus der x-Achse. Fällt man für einen (beliebigen) Punkt $P_0 \in E$ die Lote auf die Achsen, so bestimmen die beiden Fußpunkte die x- bzw. y-Koordinate x_0 bzw. y_0 von P_0, und man schreibt $P_0 = (x_0, y_0)$.



Der Punkt $0 = (0, 0)$ heißt Nullpunkt oder Ursprung des Koordinatensystems. Nach der Festlegung eines kartesischen Koordinatensystems gibt es zu jedem Zahlenpaar $(x, y) \in \mathbb{R}^2$ genau einen Punkt $X \in E$ mit $X = (x, y)$, und umgekehrt. Zu je zwei Punkten P und Q der Ebene gibt es genau eine Parallelverschiebung (der Ebene), die P nach Q bringt. Diese Verschiebung wird mit $\overrightarrow{PQ}$ bezeichnet, und heißt "Vektor von P nach Q". Der Vektor $\vec{v} = \overrightarrow{PQ}$ wird durch einen Pfeil, der von P nach Q zeigt, dargestellt.
Wird unter $\overrightarrow{PQ}$ ein anderer Punkt R nach S verschoben, dann hat offenbar $\overrightarrow{RS}$ die gleiche Wirkung wie $\overrightarrow{PQ}$, das heißt $\overrightarrow{PQ} = \overrightarrow{RS}$. Zwei gleich lange und gleichgerichtete Pfeile im Raum stellen somit den selben Vektor dar.
Offenbar gibt es zu einem Vektor $\overrightarrow{PQ}$ genau einen Punkt S, so dass $\overrightarrow{PQ} = \overrightarrow{0S}$ ist. Wir können den Vektor $\overrightarrow{PQ}$ dann mit dem Punkt S der Ebene identifizieren. Jeder Punkt der Ebene kann also ein Vektor gedeutet werden und umgekehrt, insbesondere kann jeder Vektor der Ebene durch ein Zahlenpaar beschrieben werden. Wir schreiben dieses Paar als Spalte

$$\vec{v} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Die Zahlen $x, y \in \mathbb{R}$ heißen die Komponenten von $\vec{v}$.

Die Addition von Vektoren

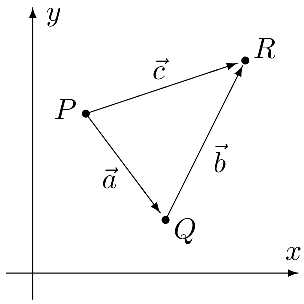
Den zu $\vec{v}$ gleichgerichteten, aber entgegengesetzten Vektor bezeichnen wir mit $-\vec{v}$ mit

$$-\vec{v} = \begin{pmatrix} -x \\ -y \end{pmatrix},$$

insbesondere ist $-\overrightarrow{PQ} = \overrightarrow{QP}$. Der Nullvektor ist

$$\vec{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

mit $\vec{0} = \overrightarrow{PP}$ für alle Punkte P. Führt man zwei Parallelverschiebungen, erst $\vec{a} = \overrightarrow{PQ}$, dann $\vec{b} = \overrightarrow{QR}$, hintereinander aus, so ergibt sich wieder eine Parallelverschiebung, nämlich $\vec{c} = \overrightarrow{PR}$. Wir nennen $\vec{c}$ die <u>Summe</u> von $\vec{a}$ und $\vec{b}$ und schreiben $\vec{c} = \vec{a} + \vec{b}$.



Sind $\vec{a}$ und $\vec{b}$ durch ihre Komponenten gegeben, so kann die Summe durch Addition der Komponenten erhalten werden:

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \Rightarrow \vec{a} + \vec{b} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \end{pmatrix}.$$

Offenbar gelten für beliebige Vektoren $\vec{a}, \vec{b}, \vec{c}$ die folgenden Rechenregeln:

$$\begin{array}{rcll} \vec{a} + \vec{0} & = & \vec{a}, \quad \vec{a} + (-\vec{a}) = \vec{0} & \\ \vec{a} + \vec{b} & = & \vec{b} + \vec{a} \qquad \text{(Kommutativgesetz)} & \qquad (VA) \\ \vec{a} + (\vec{b} + \vec{c}) & = & (\vec{a} + \vec{b}) + \vec{c} \qquad \text{(Assoziativgesetz).} & \end{array}$$

Die skalaren Vielfachen eines Vektors

Zu einer reellen Zahl $\lambda \geq 0$ und einem Vektor $\vec{a}$ bezeichne $\lambda\vec{a}$ denjenigen Vektor, der dieselbe Richtung wie $\vec{a}$ besitzt, aber die λ-fache Länge. Im Fall $\lambda < 0$ setzt man $\lambda\vec{a} := -(|\lambda|\vec{a})$.
Sonderfälle dieser Definition sind $0\vec{a} = \vec{0}$ und $\lambda\vec{0} = \vec{0}$ für jede Zahl $\lambda \in \mathbb{R}$ und jeden Vektor $\vec{a}$.
Für diese Multiplikation von Vektoren mit Zahlen (Skalarmultiplikation) gelten mit mit $\lambda, \mu \in \mathbb{R}$ und Vektoren $\vec{a}, \vec{b}$ folgende Rechenregeln:

$$\begin{array}{rcll} \lambda(\mu\vec{a}) & = & (\lambda\mu)\vec{a} & \\ \lambda(\vec{a} + \vec{b}) & = & \lambda\vec{a} + \lambda\vec{b} & \qquad (VM) \\ (\lambda + \mu)\vec{a} & = & \lambda\vec{a} + \mu\vec{a}. & \end{array}$$

Geraden

Ein Punkt X liegt genau dann auf der Geraden g durch A in Richtung $\vec{c}$ (für $\vec{c} \neq \vec{0}$), wenn $\overrightarrow{AX}$ zu $\vec{c}$ parallel ist, falls es eine also Zahl $t \in \mathbb{R}$ mit $\overrightarrow{AX} = t\vec{c}$ gibt. Man sagt, dass g die <u>Punkt-Richtungsgleichung</u>

$$\overrightarrow{AX} = t\vec{c} \qquad (*)$$

mit $t \in \mathbb{R}$ besitzt. Die in $(*)$ auftretende Variable t nennt man einen Parameter. Zu jedem Parameter $t = t_0$ gehört genau ein Punkt X_0 auf der Geraden g mit $\overrightarrow{AX_0} = t_0\vec{c}$, und umgekehrt. Wegen $\overrightarrow{AX} = \overrightarrow{PX} - \overrightarrow{PA}$ lässt sich g bezüglich eines beliebigen Punktes P durch

$$\overrightarrow{PX} = \overrightarrow{PA} + t\vec{c}$$

mit $t \in \mathbb{R}$ darstellen. Ist $A = (a_1, a_2)$, $B = (b_1, b_2)$ und $\vec{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$, so ergibt ein Komponentenvergleich für die Geradenpunkte $X = (x, y)$ die zwei Gleichungen

$$\begin{cases} x = a_1 + tc_1 \\ y = a_2 + tc_2 \end{cases} \quad \text{(Punkt-Richtungs-Gleichung)}$$

$$\begin{cases} x = a_1 + t \cdot (b_1 - a_1) \\ y = a_2 + t \cdot (b_2 - a_2) \end{cases} \quad \text{(Zwei- Punkte- Gleichung)}$$

Löst man die zwei unteren Gleichungen nach t auf, und setzt die Ausdrücke gleich, so erhält man mit $A = (a_1, a_2)$ und $B = (b_1, b_2)$ eine parameterfreie Darstellung.

Koordinatengleichung der Geraden durch A und B

Es ist

- $\dfrac{x - a_1}{b_1 - a_1} = \dfrac{y - a_2}{b_2 - a_2}$, falls $a_i \neq b_i$ für $i = 1, 2$
- $x = a_1$, falls $a_1 = b_1$
- $y = a_2$, falls $a_2 = b_2$.

Aus der (parameterfreien) Zwei-Punkte-Form für g findet man über

$$t = \frac{x - a_1}{b_1 - a_1} \quad \text{bzw.} \quad t = \frac{y - a_2}{b_2 - a_2}$$

zur Parameterform zurück.

Beispiel 1.2.1. Die durch die Parametergleichung

$$g\colon \begin{cases} x = 3 - 2t \\ y = 4 + 5t \end{cases}$$

bestimmte Gerade in der Ebene hat die Koordinatengleichung

$$\frac{3 - x}{2} = \frac{y - 4}{5}.$$

Beispiel 1.2.2. Man finde die Parameterdarstellung der durch die Gleichung $2x + 3y = 5$ gegebenen Geraden. Die Rechnung ist

$$\begin{aligned} 2x + 3y &= 5 \\ 2x - 5 &= -3y \\ \frac{x - \frac{5}{2}}{\frac{1}{2}} &= \frac{y}{-\frac{1}{3}} \\ x - \tfrac{5}{2} &= \tfrac{1}{2}t \\ y &= -\tfrac{1}{3}t \quad \Rightarrow \quad \begin{cases} x = \frac{5}{2} + \frac{1}{2}t \\ y = -\frac{1}{3}t \end{cases} \end{aligned}$$

Schnittpunkt zweier Geraden

Die Bestimmung des Schnittpunkts zweier Geraden führt auf die Lösung eines linearen Gleichungssystems, und zwar in jedem Fall: ob die Gerade nun durch die Punkt-Richtungs-Gleichung oder durch die Zwei-Punkte-Gleichung gegeben ist. Auch lineare Gleichungssysteme stehen im Zentrum der Linearen Algebra. Wir werden sie in Abschnitt 1.4 systematisch behandeln.

Beispiel 1.2.3. Seien zwei Geraden durch Punkt-Richtungs-Gleichungen gegeben:

$$g_1\colon \begin{cases} x = 3 + 5t \\ y = 2 - t \end{cases} \quad \text{und} \quad g_2\colon \begin{cases} x = 4 - 2u \\ y = 1 + 3u \end{cases}$$

Es ist wichtig, zwei verschiedene Variablen für die Parameter zu benützen. Gleichsetzen liefert

$$\begin{aligned} 3 + 5t &= 4 - 2u \\ 2 - t &= 1 + 3u \end{aligned}$$

oder

$$\begin{aligned} 5t + 2u &= 1 \\ -t - 3u &= -1. \end{aligned}$$

Addition des 5-fachen der 2. Zeile zur 1. Zeile ergibt

$$\begin{aligned} -13u &= -4 \\ -t - 3u &= -1 \\ u &= \frac{4}{13}. \end{aligned}$$

Einsetzen in das System von g_2 ergibt $x = \frac{44}{13}$ und $y = \frac{25}{13}$, also erhalten wir den Schnittpunkt $\left(\frac{44}{13}, \frac{25}{13}\right)$.

Sind die beiden Geraden durch Koordinatengleichungen gegeben, so erhält man das Gleichungssystem unmittelbar:

Beispiel 1.2.4. Seien $g_1 :\ x + 3y = 5$ und $g_2 :\ 3x - 2y = 7$ gegeben, dann ist

$$\begin{cases} x + 3y = 5 \\ 3x - 2y = 7 \end{cases} \Rightarrow \begin{cases} x + 3y = 5 \\ -11y = -8 \end{cases}$$

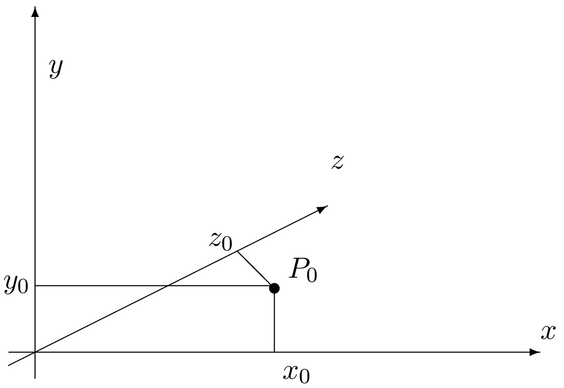
nach Subtraktion des 3-fachen der ersten Zeile von der zweiten Zeile. Dies führt zu $y = \frac{8}{11}$ und $x = \frac{31}{11}$, also zu dem Schnittpunkt $P_0 = \left(\frac{31}{11}, \frac{8}{11}\right)$.

Die Gleichungssysteme haben keine bzw. unendlich viele Lösungen, falls es sich um parallele bzw. identische Geraden handelt.

Der Raum

Es sei nun $\mathbb{R}^3 = \{(x, y, z)\colon x, y, z \in \mathbb{R}\}$.
Ähnlich wie die Ebene mit dem $\mathbb{R}^2$ kann der Raum nun mit dem $\mathbb{R}^3$ identifiziert werden. Ein kartesisches Koordinatensystem im Raum besteht aus dem Nullpunkt 0 und drei sich in 0 schneidenden Zahlengeraden gleicher Längeneinheit. Man bezeichnet sie als x, y und z-Achse derart, dass diese drei Achsen ein Rechtssystem bilden, das heißt, die Drehung der positiven x-Achse um $90°$ in die positive

y-Achse, zusammen mit einer Verschiebung in Richtung der positiven z-Achse, muss eine Rechtsschraube darstellen. Diese drei durch je zwei Achsen bestimmten Ebenen heißen Koordinatenebenen, bzw. (x, y)-Ebene, (y, z)-Ebene und (z, x)-Ebene. Die Koordinaten x_0, y_0, z_0 eines Punktes P_0 gewinnt man aus den Schnittpunkten der entsprechenden Achsen mit dem zu den Koordinatenebenen parallelen Ebenen durch P_0. Man schreibt $P_0 = (x_0, y_0, z_0)$.



Völlig analog zum Fall der Ebene werden nun auch im Raum Vektoren als Parallelverschiebungen des Raumes definiert. Auch die Pfeildarstellung sowie die Operationen der Addition und der Skalarmultiplikation verlaufen völlig analog zum Fall der Ebene. Der einzige Unterschied liegt in der Tatsache, dass Vektoren im Raum drei Komponenten besitzen. Wie in der Ebene besitzt eine Gerade im Raum eine Parameterdarstellung

$$\overrightarrow{OX} = \overrightarrow{OA} + t\vec{x}$$

mit $t \in \mathbb{R}$. Falls $A = (a_1, a_2, a_3)$ und

$$\vec{c} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

ist, so ergibt ein Komponentenvergleich die drei Gleichungen

$$\left\{ \begin{array}{l} x = a_1 + tc_1 \\ y = a_2 + tc_2 \\ z = a_3 + tc_3 \end{array} \right. \qquad \text{(Punkt- Richtungs- Gleichung)}$$

Für eine Gerade durch die zwei verschiedenen Punkte $A = (a_1, a_2, a_3)$ und $B = (b_1, b_2, b_3)$ erhält man die Gleichung

$$\left\{ \begin{array}{l} x = a_1 + t \cdot (b_1 - a_1) \\ y = a_2 + t \cdot (b_2 - a_2) \\ z = a_3 + t \cdot (b_3 - a_3) \end{array} \right. \qquad \text{(Zwei- Punkte- Gleichung)}$$

Die parameterfreien Koordinatengleichungen der Geraden durch A und B sind

- $\dfrac{x - a_1}{b_1 - a_1} = \dfrac{y - a_2}{b_2 - a_2} = \dfrac{z - a_3}{b_3 - a_3}$, falls $a_i \neq b_i$ für $i = 1, 2, 3$
- $\dfrac{x - a_1}{b_1 - a_1} = \dfrac{y - a_2}{b_2 - a_2}$ und $z = a_3$, falls $a_i \neq b_i$ für $i = 1, 2$ und $a_3 = b_3$
- $x = a_1$ und $y = a_2$, falls $a_1 = b_1$, $a_2 = b_2$ und $a_3 \neq b_3$.

Neben den Geraden sind die Ebenen wichtige Teilmengen des Raumes.
Wir betrachten nun die Ebene E mit dem "Aufpunkt" A und den beiden (von $\vec{0}$ verschiedenen und nicht parallelen) Richtungsvektoren $\vec{u}$ und $\vec{v}$. Ein Raumpunkt X liegt genau dann auf E, wenn sich

der Vektor $\overrightarrow{AX}$ in der Form $\overrightarrow{AX} = t\vec{u} + s\vec{v}$ mit Zahlen $t, s \in \mathbb{R}$ darstellen lässt, das heißt, man hat (mit den zwei Parametern $t, s \in \mathbb{R}$) die Parameterdarstellung von E:

$$\overrightarrow{AX} = t\vec{u} + s\vec{v}. \tag{1}$$

Wird ein Kartesisches Koordinatensystem festgelegt, so dass $A = (a_1, a_2, a_3)$ und

$$\vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \quad \text{bzw.} \quad \vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

ist, so die Parameterdarstellung (1) äquivalent zu den drei Koordinatengleichungen

$$\begin{cases} x = a_1 + tu_1 + sv_1 \\ y = a_2 + tu_2 + sv_2 \\ z = a_3 + tu_3 + sv_3. \end{cases} \tag{2}$$

Werden $\vec{u}$ und $\vec{v}$ durch die verschiedenen Punkte $A = (a_1, a_2, a_3)$, $B = (b_1, b_2, b_3)$ und $C = (c_1, c_2, c_3)$ mit $\vec{u} = \overrightarrow{AB}$, $\vec{v} = \overrightarrow{BC}$, also $u_i = b_i - a_i$ und $v_i = c_i - a_i$, bestimmt, dann geht (2) in die Drei-Punkte-Gleichung für die Ebene

$$\begin{cases} x = a_1 + t \cdot (b_1 - a_1) + s \cdot (c_1 - a_1) \\ y = a_2 + t \cdot (b_2 - a_2) + s \cdot (c_2 - a_2) \\ z = a_3 + t \cdot (b_3 - a_3) + s \cdot (c_3 - a_3) \end{cases}$$

über.

1.3 Der $\mathbb{R}^n$

Unsere geometrische Anschauung ist auf drei Dimensionen beschränkt, da unsere physikalische Welt dreidimensional ist. Jedoch gibt es keine Gründe, in der Mathematik, die von der Natur unabhängig ist, nicht auch Räume mit mehr Dimensionen zu betrachten.

Definition 1.3.1. Es sei n eine natürliche Zahl. Unter dem $\mathbb{R}^n$ verstehen wir die Menge aller n- tupel reeller Zahlen:

$$\mathbb{R}^n = \{(x_1, \ldots, x_n) \colon x_i \in \mathbb{R} \text{ für } i = 1, \ldots, n\}.$$

Diese n- tupel heißen auch Punkte des $\mathbb{R}^n$. Sie können wiederum mit Vektoren identifiziert werden. Diese werden als Spaltenvektoren geschrieben:

$$\vec{v} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Die Zahlen $x_1, \ldots, x_n \in \mathbb{R}$ heißen die Komponenten von $\vec{v}$.

Addition und Skalarmultiplikation werden komponentenweise definiert. Es gelten die Rechenregeln (VA) und (VM). Auch die Definition der Geraden kann übertragen werden. Die Definitionen von Abschnitt 1.2 ergeben sich offenbar als die Spezielfälle $n = 2$ (Ebene) und $n = 3$ (Raum).

1.4 Lineare Gleichungssysteme

Definition 1.4.1. Es seien $\alpha_{ij}, \beta_i \in \mathbb{R}$ für $i = 1, \ldots, m$ und $j = 1, \ldots, n$. Man bezeichnet

$$\begin{array}{ccccccccc} \alpha_{11}x_1 & + & \alpha_{12}x_2 & + & \cdots & + & \alpha_{1n}x_n & = & \beta_1 \\ \alpha_{21}x_1 & + & \alpha_{22}x_2 & + & \cdots & + & \alpha_{2n}x_n & = & \beta_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ \alpha_{m1}x_1 & + & \alpha_{m2}x_2 & + & \cdots & + & \alpha_{mn}x_n & = & \beta_m \end{array} \tag{$*$}$$

als ein Lineares Gleichungssystem (LGS) mit m Gleichungen und n Unbestimmten $x_1, \ldots, x_n$ und Koeffizienten $\alpha_{ij} \in \mathbb{R}$. Eine einfachere Beschreibung erhält man, wenn man die Koeffizienten in einer Matrix, einem rechteckigen Schema, sammelt:

$$\mathcal{A} := \begin{pmatrix} \alpha_{11} & \ldots & \alpha_{1n} \\ \vdots & \ldots & \vdots \\ \alpha_{m1} & \ldots & \alpha_{mn} \end{pmatrix}$$

Man nennt $\mathcal{A}$ auch die Koeffizientenmatrix des LGS. Sie besitzt m Zeilen und n Spalten. Der Koeffizient α_{ij} steht in der i- ten Zeile und der j- ten Spalte. Man scheibt auch

$$\mathcal{A} = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$$

oder einfach $\mathcal{A} = (a_{ij})$, wenn m und n klar sind. Dann heißt $\mathcal{A}$ eine (m, n)- Matrix oder eine Matrix vom Typ (m, n).
Die Menge aller (m, n)- Matrizen mit reellen Koeffizienten bezeichnen wir mit $\mathbb{R}^{(m,n)}$. Die Elemente der rechten Seite können zu einem Vektor des $\mathbb{R}^m$ zusammengefasst werden:

$$\vec{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix}.$$

Er heißt rechte Seite des Systems. Fügt man zu $\mathcal{A}$ als $(n+1)$-te Spalte $\vec{\beta}$ hinzu, so erhält man die erweiterte Koeffizientenmatrix $(\mathcal{A}|\vec{\beta}) \in \mathbb{R}^{(m,n+1)}$ des LGS $(*)$.
Man kann die Unbestimmten $x_1, \ldots, x_n$ zu einem Spaltenvektor des $\mathbb{R}^n$

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

zusammenfassen.
Das Produkt $\mathcal{A}\vec{x}$ der Matrix $\mathcal{A}$ mit dem Vektor $\vec{x}$ wird dann als die linke Seite von $(*)$ definiert, womit sich $(*)$ kurz als

$$\mathcal{A}\vec{x} = \vec{\beta} \tag{$**$}$$

schreiben lässt.
Wir werden später das Produkt von Matrizen definieren und dabei feststellen, dass sich dabei dieses Produkt $\mathcal{A}\vec{x}$ als Spezialfall herausstellt.
Jedes $\vec{x} \in \mathbb{R}^n$, für das $(*)$ bzw. $(**)$ gilt, heißt eine Lösung des LGS. Gibt es ein solches $\vec{x}$, so heißt $(*)$ lösbar, ansonsten unlösbar. Die Menge aller Lösungen heißt Lösungsmenge des Systems. Ist $\vec{\beta} = \vec{0}$, so heißt das LGS homogen, ansonsten inhomogen. Man nennt dann $\mathcal{A}\vec{x} = \vec{0}$ das zu $(**)$ gehörende homogene LGS. Ein homogenes LGS besitzt immer die triviale Lösung $\vec{x} = \vec{0}$.

In diesem Abschnitt begnügen wir uns vorerst mit der Beschreibung eines Lösungsverfahrens, des Gaußschen Algorithmus.
Der Gaußsche Algorithmus besteht aus einer Reihe von elementaren Umformungen des LGS, die es in eine einfache "Standardform" bringen. Diese Umformungen können direkt am System $(*)$ oder aber auch an der Koeffizientenmatrix vorgenommen werden.

Bevor wir das Verfahren allgemein beschreiben, beginnen wir mit einem Beispiel:

Beispiel 1.4.1. Finde die Lösungsmenge des LGS

$$\begin{array}{ccccccc} w & -2x & +5y & +3z & = & 3 & (I) \\ 3w & -5x & +18y & +13z & = & 16 & (II) \\ 2w & -2x & +17y & +16z & = & 21 & (III) \\ w & -x & +10y & +11z & = & 12 & (IV) \end{array}$$

mit der Koeffizientenmatrix

$$\left(\begin{array}{cccc|c} 1 & -2 & 5 & 3 & 3 \\ 3 & -5 & 18 & 13 & 16 \\ 2 & -2 & 17 & 16 & 21 \\ 1 & -1 & 10 & 11 & 12 \end{array}\right).$$

1. Schritt:
Wir elimieren die Unbestimmte w aus den Gleichungen (II)- (IV) durch Addition passender Vielfache der Gleichung (I). Wir notieren die Umformungen an der Koeffizientenmatrix:

$$\begin{array}{cccccc} w & -2x & +5y & +3z & = & 3 \\ & x & +3y & +4z & = & 7 \\ & 2x & +7y & +10z & = & 15 \\ & x & +5y & +8z & = & 9 \end{array} \quad \Leftrightarrow \quad \left(\begin{array}{cccc|c} 1 & -2 & 5 & 3 & 3 \\ 0 & 1 & 3 & 4 & 7 \\ 0 & 2 & 7 & 10 & 15 \\ 0 & 1 & 5 & 8 & 9 \end{array}\right). \quad \begin{array}{c} \\ (II) - 3 \cdot (I) \\ (III) - 2 \cdot (I) \\ (IV) - 1 \cdot (I) \end{array}$$

Wir benutzen also das Pivotelement $\alpha_{11} = 1$ um zu erreichen, dass unter diesem Element in der ersten Spalte nur Nullen stehen.

2. Schritt:
Wir elimieren die Unbestimmte x aus den Gleichungen (I), (III) und (IV) durch Addition passender Vielfache der Gleichung (II).

$$\begin{array}{cccccc} w & & +11y & +11z & = & 17 \\ & x & +3y & +4z & = & 7 \\ & & y & +2z & = & 1 \\ & & 2y & +4z & = & 2 \end{array} \quad \Leftrightarrow \quad \left(\begin{array}{cccc|c} 1 & 0 & 11 & 11 & 17 \\ 0 & 1 & 3 & 4 & 7 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 2 & 4 & 2 \end{array}\right). \quad \begin{array}{c} (I) + 2 \cdot (II) \\ \\ (III) - 2 \cdot (II) \\ (IV) - 1 \cdot (II) \end{array}$$

3. Schritt:
Wir elimieren die Unbestimmte y aus den Gleichungen (I), (II) und (IV).

$$\begin{array}{cccccc} w & & & -11z & = & 6 \\ & x & & -2z & = & 4 \\ & & y & +2z & = & 1 \\ & & & 0 & = & 0 \end{array} \quad \Leftrightarrow \quad \left(\begin{array}{cccc|c} 1 & 0 & 0 & -11 & 6 \\ 0 & 1 & 0 & -2 & 4 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right). \quad \begin{array}{c} (I) - 11 \cdot (III) \\ (II) - 3 \cdot (III) \\ \\ (IV) - 2 \cdot (III) \end{array}$$

Das LGS ist lösbar und besitzt die Lösungsmenge $\mathcal{L} = \{(w, x, y, z) = (6+11z, 4+2z, 1-2z, z) \colon z \in \mathbb{R}\}$.

Beispiel 1.4.2. Wir ändern in Beispiel 1.4.1 die Gleichung (IV) ab und betrachten

$$\begin{array}{rrrrcrc} w & -2x & +5y & +3z & = & 3 & (I) \\ 3w & -5x & +18y & +13z & = & 16 & (II) \\ 2w & -2x & +17y & +16z & = & 21 & (III) \\ w & -x & +10y & +11z & = & 13 & (IV) \end{array}$$

Dieselben Umformungen wie in Beispiel 1.4.1 führen zu

$$\begin{array}{rrrrcr} w & & & -11z & = & 6 \\ & x & & -2z & = & 4 \\ & & y & +2z & = & 1 \\ & & & 0 & = & 1 \end{array}$$

Die letzte Gleichung ist falsch. Dies zeigt, dass das LGS unlösbar ist.

Es gibt noch eine andere Version des Gaußschen Algorithmus, bei der nur die Koeffizienten unterhalb der Diagonalen eliminiert werden.
Wir betrachten wieder Beispiel 1.4.1:

$$\begin{array}{rrrrcr} w & -2x & +5y & +3z & = & 3 \\ & x & +3y & +4z & = & 7 \\ & 2x & +7y & +10z & = & 15 \\ & x & +5y & +8z & = & 9 \end{array} \quad \Leftrightarrow \quad \left(\begin{array}{rrrr|r} 1 & -2 & 5 & 3 & 3 \\ 0 & 1 & 3 & 4 & 7 \\ 0 & 2 & 7 & 10 & 15 \\ 0 & 1 & 5 & 8 & 9 \end{array}\right).$$

$$\begin{array}{rrrrcr} w & -2x & +5y & +3z & = & 3 \\ & x & +3y & +4z & = & 7 \\ & & y & +2z & = & 1 \\ & & 2y & +4z & = & 2 \end{array} \quad \Leftrightarrow \quad \left(\begin{array}{rrrr|r} 1 & -2 & 5 & 3 & 3 \\ 0 & 1 & 3 & 4 & 7 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 2 & 4 & 2 \end{array}\right).$$

$$\begin{array}{rrrrcr} w & -2x & +5y & +3z & = & 3 \\ & x & +3y & +4z & = & 7 \\ & & y & +2z & = & 1 \\ & & & 0 & = & 0 \end{array} \quad \Leftrightarrow \quad \left(\begin{array}{rrrr|r} 1 & -2 & 5 & 3 & 3 \\ 0 & 1 & 3 & 4 & 7 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Während die Umformung des Systems hier weniger Arbeit macht, ist die Gewinnung der Lösung aus der "Standardform" aufwändiger. Hier kann die Lösung gefunden werden, indem man die Gleichungen von unten nach oben löst.
In manchen Fällen treten zusätzliche Komplikationen auf. Es kann passieren, dass zum Beispiel nach dem ersten Schritt in der zweiten Zeile und zweiten Spalte ebenfalls eine Null steht. Steht in der zweiten Spalte in der i- ten Zeile ein von null verschiedenes Element, so kann eine Pivotisierung durch Vertauschung der zweiten Zeile mit der i- ten Zeile erreicht werden.
Stehen nach dem ersten Schritt in der zweiten Spalte unterhalb der ersten Zeile nur Nullen, so muss die zweite Spalte mit einer der Spalten $3, \ldots, n$ vertauscht werden.

Beispiel 1.4.3. Es liegt das LGS

$$\begin{array}{rrrrcr} w & +2x & +y & -z & = & 1 \\ 2w & +4x & +3y & +5z & = & -1 \\ 3w & +6x & +2y & +z & = & -2 \end{array}$$

mit der Koeffizientenmatrix

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 1 \\ 2 & 4 & 3 & 5 & -1 \\ 3 & 6 & 2 & 1 & -2 \end{array}\right)$$

vor. Der 1. Schritt ergibt:

$$\begin{array}{cccccc} w & +2x & +y & -z & = & 1 \\ & & y & +7z & = & -3 \\ & & -y & +4z & = & -5 \end{array} \quad \Leftrightarrow \quad \left(\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 1 \\ 0 & 0 & 1 & 7 & -3 \\ 0 & 0 & -1 & 4 & -5 \end{array}\right).$$

Man nimmt eine Pivotisierung vor, indem man die Variablen x und y, das entspricht der zweiten und dritten Spalte, vertauscht. Arbeitet man nur mit der Koeffizientenmatrix, so muss diese Vertauschung notiert werden.

$$\begin{array}{cccccc} w & +y & +2x & -z & = & 1 \\ & y & & +7z & = & -3 \\ & -y & & +4z & = & -5 \end{array} \quad \Leftrightarrow \quad \left(\begin{array}{cccc|c} w & y & x & z & \\ \hline 1 & 1 & 2 & -1 & 1 \\ 0 & 1 & 0 & 7 & -3 \\ 0 & -1 & 0 & 4 & -5 \end{array}\right).$$

Daraus resultiert dann

$$\begin{array}{cccccc} w & +y & +2x & -z & = & 1 \\ & y & & +7z & = & -3 \\ & & & +11z & = & -8 \end{array} \quad \Leftrightarrow \quad \left(\begin{array}{cccc|c} w & y & x & z & \\ \hline 1 & 1 & 2 & -1 & 1 \\ 0 & 1 & 0 & 7 & -3 \\ 0 & 0 & 0 & 11 & -8 \end{array}\right).$$

Die Lösungsmenge ergibt sich als

$$\mathcal{L} = \left\{(w,x,y,z) = \left(w, -\frac{10}{11} - \frac{1}{2}w, \frac{23}{11}, -\frac{8}{11}\right)\right\} = \left\{(w,x,y,z) = \left(-\frac{20}{11} - 2x, x, \frac{23}{11}, -\frac{8}{11}\right)\right\}.$$

Wir schließen mit einer allgemeinen Beschreibung:
Gegeben sei ein LGS

$$\mathcal{A}\vec{x} = \vec{\beta}, \quad \mathcal{A} \in \mathbb{R}^{(m,n)} \quad \text{mit} \quad (\mathcal{A}|\vec{\beta}) = \left(\begin{array}{ccc|c} \alpha_{11} & \cdots & \alpha_{1n} & \beta_1 \\ \vdots & & \vdots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} & \beta_m \end{array}\right). \tag{$*$}$$

1. Schritt:
Falls in der ersten Spalte mindestens ein Element $\alpha_{k1} \neq 0$ vorkommt, bringe man $(\mathcal{A}|\vec{\beta})$ durch elementare Zeilenumformungen in die Gestalt

$$(\mathcal{A}^{(1)} \,|\, \vec{\beta}^{(1)}) = \left(\begin{array}{cccc|c} 1 & \alpha_{12}^{(1)} & \cdots & \alpha_{1n}^{(1)} & \beta_1^{(1)} \\ 0 & \alpha_{22}^{(1)} & \cdots & \alpha_{2n}^{(1)} & \beta_2^{(1)} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & \alpha_{m2}^{(1)} & \cdots & \alpha_{mn}^{(1)} & \beta_m^{(1)} \end{array}\right). \tag{1}$$

Das LGS $\mathcal{A}^{(1)}\vec{x} = \vec{\beta}^{(1)}$ hat dann dieselbe Lösungsmenge wie $\mathcal{A}\vec{x} = \vec{b}$, denn jede elementare Zeilenumformung lässt sich ja durch passende elementare Zeilenumformungen wieder rückgängig machen. Enthält die erste Spalte von $\mathcal{A}$ nur Nullen, so vertausche man sie zuvor mit einer Spalte von $\mathcal{A}$ (die

Spalte $\vec{\beta}$ kommt nicht in Betracht), welche mindestens ein von null verschiedenes Element enthält. Diese Umformung muss registriert werden; sie bedeutet eine Umnummerierung der Unbekannten. Die so erhaltene Matrix bringe man dann durch elementare Zeilenumformungen auf die Gestalt (1). Ist $\mathcal{A} = 0^{(m,n)}$, so ist man von vorneherein fertig.

2. Schritt:
Falls in der Matrix $\mathcal{A}^{(1)}$ in der zweiten Spalte unter den $\alpha_{i2}^{(1)}$ für $j = 2, \ldots, m$ mindestens ein von null verschiedenes Element vorkommt, bringe man $(\mathcal{A}^{(1)}|\vec{\beta}^{(1)})$ durch elementare Zeilenumformungen in die Gestalt

$$(\mathcal{A}^{(2)} \mid \vec{\beta}^{(2)}) = \left(\begin{array}{ccccc|c} 1 & 0 & \alpha_{13}^{(2)} & \cdots & \alpha_{1n}^{(2)} & \beta_1^{(2)} \\ 0 & 1 & \alpha_{23}^{(2)} & \cdots & \alpha_{2n}^{(2)} & \beta_2^{(2)} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \alpha_{m3}^{(2)} & \cdots & \alpha_{mn}^{(2)} & \beta_m^{(2)} \end{array}\right). \tag{2}$$

Sind $\alpha_{22}^{(1)} = \ldots = \alpha_{m2}^{(1)} = 0$, dann vertausche man die zweite Spalte von $\mathcal{A}^{(1)}$ mit einer Spalte der Nummer j_0 mit $3 \leq j_0 \leq n$, welche mindestens ein $\alpha_{i_0,j_0}^{(1)} \neq 0$ enthält ($i_0 \geq 2$). Danach bringe man die Matrix durch elementare Zeilenumformungen in die Form (2). Verschwindet jedoch die gesamte Teilmatrix $(\alpha_{ij}^{(1)})$ für $2 \leq i \leq m$ und $2 \leq j \leq n$, so ist man fertig.

So fortfahrend, gelangt man zu einer Matrix der Gestalt

$$(\mathcal{A}^{(p)}|\vec{b}^{(p)}) = \left(\begin{array}{cccccccc|c} 1 & 0 & 0 & \cdots & 0 & \alpha_{1,p+1}^{(p)} & \cdots & \alpha_{1,n}^{(p)} & \beta_1^{(p)} \\ 0 & 1 & 0 & \cdots & 0 & \alpha_{2,p+1}^{(p)} & \cdots & \alpha_{2,n}^{(p)} & \beta_2^{(p)} \\ 0 & 0 & 1 & \cdots & 0 & \alpha_{3,p+1}^{(p)} & \cdots & \alpha_{3,n}^{(p)} & \beta_3^{(p)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \alpha_{p,p+1}^{(p)} & \cdots & \alpha_{p,n}^{(p)} & \beta_p^{(p)} \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \beta_{p+1}^{(p)} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \beta_m^{(p)} \end{array}\right). \tag{p}$$

Die Lösungsmenge des LGS $\mathcal{A}^{(p)}\vec{x} = \vec{\beta}$ stimmt nun (bis auf eventuelle Umnummerierung der Unbekannten) mit der von $\mathcal{A}\vec{x} = \vec{\beta}$ überein. An den letzten $m - p$ Zeilen kann die Lösbarkeit des Systems abgelesen werden:

- Fall 1:
  Ist mindestens eines der $\beta_{p+1}^{(p)}, \ldots, \beta_m^{(p)}$ nicht null, so ist das LGS nicht lösbar
- Fall 2:
  Ist dagegen $\beta_{p+1}^{(p)} = \cdots = \beta_m^{(p)} = 0$, so ist das LGS lösbar.

1.5 Verknüpfungen und Gruppen

Definition 1.5.1. Es sei $X \neq \emptyset$ eine Menge. Eine Abbildung $\circ\colon X \times X \to X$ heißt eine Verknüpfung auf X.

Bemerkung 1.5.1. Jedem Paar $(x, y) \in X \times X$ wird also als Ergebnis der Verknüpfung das Element $x \circ y \in X$ zugeordnet.

Beispiel 1.5.1. Anstelle von "$\circ$" kann die Verknüpfung jeden Namen erhalten. Auf der Menge $\mathbb{R}$ der reellen Zahlen haben wir die Verknüpfungen "$+$" der Addition und "$\cdot$" der Multiplikation. So ordnet etwa die Addition dem Paar $(5,3)$ als Ergebnis der Verknüpfung die Summe 8 $(= 5+3)$ zu.

Bemerkung 1.5.2. Im Falle einer endlichen Menge $X = \{x_1, x_2, \ldots, x_n\}$ kann man $\circ$ durch eine Verknüpfungstafel darstellen:

$$\begin{array}{c|cccccc}
\circ & x_1 & x_2 & \ldots & x_k & \ldots & x_n \\
\hline
x_1 & & & & \vdots & & \\
x_2 & & & & \vdots & & \\
\vdots & & & & \vdots & & \\
x_l & \cdots & \cdots & \cdots & x_k \circ x_l & & \\
\vdots & & & & & & \\
x_n & & & & & &
\end{array}$$

Beispiel 1.5.2. Die Menge $X = \{-1, 0, 1\}$ besitzt mit der Multiplikation die Verknüpfungstafel

$\cdot$	-1	0	1
-1	1	0	-1
0	0	0	0
1	-1	0	1

Definition 1.5.2. Es sei $\circ$ eine innere Verknüpfung auf einer Menge $G \neq \emptyset$. Dann heißt $(G, \circ)$ eine Gruppe, wenn die folgenden Axiome gelten:

(G1) $(a \circ b) \circ c = a \circ (b \circ c)$ für alle $a, b, c \in G$ (Assoziativgesetz).

(G2) Es gibt mindestens ein neutrales Element $e \in G$ mit $e \circ a = a \circ e = a$ für alle $a \in G$.

(G3) Ist ein neutrales Element $e \in G$ gegeben, so gibt es zu jedem $a \in G$ ein inverses Element $a' \in G$ mit $a \circ a' = a' \circ a = e$.

Gilt zusätzlich das Axiom

(G4) $a \circ b = b \circ a$ für alle $a, b \in G$ (Kommutativgesetz),

so heißt G eine abelsche (oder auch kommutative) Gruppe.

Beispiel 1.5.3. Es sei $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ die Menge der ganzen Zahlen mit der Verknüpfung der Addition. Das Assoziativgesetz (G1) ist hier offenbar erfüllt: $(a+b)+c = a+(b+c)$. Die Zahl 0 ist das einzige neutrale Element: $0+a = a+0 = a$ für alle $a \in \mathbb{Z}$. Zu jeder Zahl $a \in \mathbb{Z}$ existiert ein eindeutiges Inverses, nämlich $-a$ mit $(-a)+a = 0$. Außerdem ist (G4) erfüllt: $a+b = b+a$ für alle $a, b \in \mathbb{Z}$. Somit ist $(\mathbb{Z}, +)$ eine abelsche Gruppe.

Beispiel 1.5.4. Ebenso sind $(\mathbb{Q}, +)$ und $(\mathbb{R}, +)$, wobei $\mathbb{Q}$ die Menge der rationalen Zahlen und $\mathbb{R}$ die Menge der reellen Zahlen darstellen, abelsche Gruppen bzgl. der Addition. Für $\mathbb{N} = \{1, 2, 3, \ldots\}$, der Menge der natürlichen Zahlen, ist hingegen $(\mathbb{N}, +)$ keine Gruppe, da kein neutrales Element existiert.

Beispiel 1.5.5. Es sei $n \in \mathbb{N}$. Die Rechenregeln (VA) der Vektoraddition (Abschnitt 1.2) zeigen, dass $(\mathbb{R}^n, +)$ eine abelsche Gruppe bildet. Das neutrale Element ist der Nullvektor $\vec{0}$. Zu jedem $\vec{a} \in \mathbb{R}^n$ existiert ein eindeutiges Inverses, nämlich $-\vec{a}$.

Beispiel 1.5.6. Die Menge $\mathbb{Z}$ der ganzen Zahlen bildet keine Gruppe bzgl. der Multiplikation. Es gibt zwar ein neutrales Element, nämlich die Zahl 1, aber es gibt für die Zahlen $a \neq \pm 1$ in $\mathbb{Z}$ keine Inversen.

Beispiel 1.5.7. Die Mengen $\mathbb{Q}\backslash\{0\}$ und $\mathbb{R}\backslash\{0\}$ bilden abelsche Gruppen bzgl. der Multiplikation. In beiden Fällen ist die Zahl 1 das einzige neutrale Element. Die Zahl $a \neq 0$ hat $\frac{1}{a} = a^{-1}$ als Inverses.

Beispiel 1.5.8. (Eine nicht-abelsche Gruppe)
Es sei S_3 die Menge der Permutationen von $\{1,2,3\}$, d.h. die Menge der bijektiven Abbildungen von $\{1,2,3\}$ auf sich selbst. Die Verknüpfung auf S_3 ist die Komposition (Hintereinanderausführung) der Permutationen. Jede Permutation τ werde in der Form

$$\tau = \begin{pmatrix} 1 & 2 & 3 \\ \tau(1) & \tau(2) & \tau(3) \end{pmatrix}$$

notiert. Dann bildet S_3 eine Gruppe mit neutralem Element

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

Jedes $\tau \in S_3$ besitzt als Inverses seine Umkehrabbildung τ^{-1}. Die Gruppe S_3 ist nicht abelsch: sei z.B.

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \quad \text{und} \quad \tau = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

Dann ist

$$\tau \circ \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \quad \text{aber} \quad \sigma \circ \tau = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

Wir zeigen nun unseren ersten

Satz 1.5.1. *In einer beliebigen Gruppe $(G, \cdot)$ gelten folgende Eigenschaften:*

i) Es gibt genau ein neutrales Element.

ii) Jedes $a \in G$ hat genau ein Inverses.

Beweis. Zu (i): Beweis durch Widerspruch:
Angenommen, es gibt verschiedene neutrale Elemente $e_1, e_2 \in G$. Da e_1 neutral ist gilt $e_1 \cdot a = a$ für alle $a \in G$, wir können also $a = e_2$ einsetzen und erhalten $e_1 \cdot e_2 = e_2$. Da auch e_2 neutral ist gilt $a \cdot e_2 = a$ für alle $a \in G$, einsetzen von $a = e_1$ ergibt $e_1 \cdot e_2 = e_1$. Zusammensetzen der beiden Gleichungen ergibt $e_1 = e_1 \cdot e_2 = e_2$, ein Widerspruch zur Annahme, dass $e_1 \neq e_2$ ist.

Zu (ii): Es sei $a \in G$ beliebig und $a_1, a_2 \in G$ zwei Inverse, also

$$a \cdot a_1 = a_1 \cdot a = e \quad \text{und} \quad a \cdot a_2 = a_2 \cdot a = e$$

mit dem nach (i) eindeutig bestimmten neutralen Element e von G. Wir multiplizieren die zweite Gleichung von links mit a_1 und erhalten

$$a_1 \cdot (a \cdot a_2) = a_1 \cdot e.$$

Anwendung des Axioms (G1) ergibt die Gleichungen

$$(a_1 \cdot a) \cdot a_2 = a_1 \cdot e.$$

Anwendung von (G3) auf der linken Seite ergibt

$$e \cdot a_2 = a_1 \cdot e$$

und nach (G2) folgt $a_2 = a_1$, also waren die Inversen gleich. □

Bemerkung 1.5.3. Für eine multiplikativ geschriebene Gruppe nennt man das neutrale Element auch Einselement und schreibt 1 statt e. Der Multiplikationspunkt wird oft weggelassen, man schreibt also ab statt $a \cdot b$. Für das Inverse a' schreiben wir auch a^{-1}.

Wir zeigen nun ein paar Rechenregeln für Gruppen:

Satz 1.5.2. *In jeder Gruppe $(G, \cdot)$ gilt:*

i) Zu beliebigen $a, b \in G$ gibt es eindeutig bestimmte $x, y \in G$ mit $ax = b$ und $ya = b$ (nämlich $x = a^{-1}b$ und $y = ba^{-1}$).

ii) Es gilt $(a^{-1})^{-1} = a$ für alle $a \in G$.

Beweis. Zu (i): Wir müssen zeigen, dass es eine Lösung der Gleichungen $ax = b$ bzw. $ya = b$ gibt und dass diese eindeutig bestimmt ist.

Existenz:
Einsetzen von $x = a^{-1}b$ und $y = ba^{-1}$ ergibt

$$ax = a(a^{-1}b) \underset{\text{(G1)}}{=} (aa^{-1})b \underset{\text{(G3)}}{=} eb \underset{\text{(G2)}}{=} b \quad \text{und} \quad ya = (ba^{-1})a \underset{\text{(G1)}}{=} b(a^{-1}a) \underset{\text{(G3)}}{=} be \underset{\text{(G2)}}{=} b.$$

Eindeutigkeit:
Sind $x_1, x_2 \in G$ zwei Lösungen von $ax = b$, so gilt $ax_1 = b = ax_2$. Nach Multiplikation von links mit a^{-1} erhalten wir $a^{-1}ax_1 = a^{-1}ax_2$, daraus folgt mit (G3) und (G2) $x_1 = x_2$, die Lösungen waren also gleich. Die Rechnung für die Lösungen von $ya = b$ verläuft analog.

Zu (ii): Nach (G3) ist $\tilde{a} = (a^{-1})^{-1}$ ein Element aus G mit der Eigenschaft $\tilde{a} \cdot a^{-1} = a^{-1} \cdot \tilde{a} = e$. Auch a erfüllt diese Eigenschaft wegen (G2). Nach Satz 1.5.1(ii) ist $a = \tilde{a} = (a^{-1})^{-1}$. □

Satz 1.5.3. *Für endlich viele Elemente $a_1, a_2, \ldots, a_n \in G$ gilt $(a_1 \cdot a_2 \cdots a_n)^{-1} = a_n^{-1} \cdot a_{n-1}^{-1} \cdots a_1^{-1}$.*

Beweis. Diese Aussage beweisen wir durch vollständige Induktion nach n:
Der Induktionsanfang ist $n = 1$: für nur ein einziges Element ist die Aussage $a_1^{-1} = a_1^{-1}$ richtig.
Die Induktionsannahme ist, dass für ein beliebiges $n \geq 1$ die Aussage $(a_1 \cdots a_n)^{-1} = a_n^{-1} \cdots a_1^{-1}$ gilt.
Der Induktionsschritt besteht darin, dass wir die Aussage für $n + 1$ zeigen, indem wir die Induktionsannahme für n verwenden. Die Aussage für n lautet, dass die rechte Seite des Satzes das Axiom (G3) erfüllt, also gilt

$$(a_1 \cdots a_n) \cdot (a_n^{-1} \cdots a_1^{-1}) = e.$$

Nach (G2) dürfen wir e in der Mitte einfügen ohne den Wert der linken Seite zu ändern:

$$(a_1 \cdots a_n) \cdot e \cdot (a_n^{-1} \cdots a_1^{-1}) = e.$$

Nach (G3) können wir e durch $a_{n+1}a_{n+1}^{-1}$ ersetzen, und erhalten

$$(a_1 \cdots a_n) \cdot (a_{n+1}a_{n+1}^{-1}) \cdot (a_n^{-1} \cdots a_1^{-1}) = e.$$

Wegen (G1) dürfen wir die Klammen umsetzen zu

$$(a_1 \cdots a_n \cdot a_{n+1}) \cdot (a_{n+1}^{-1} \cdot a_n^{-1} \cdots a_1^{-1}) = e.$$

Damit ist die rechte Klammer das Inverse der linken Klammer nach (G3). Das ist die gewünschte Aussage für $n+1$. Damit haben wir die Induktion abgeschlossen, und der Satz gilt für alle $n \in \mathbb{N}$. □

Insbesondere ist die aus der Schule bekannte Regel $(a_1 \cdots a_n)^{-1} = a_1^{-1} \cdots a_n^{-1}$ nur in abelschen Gruppen richtig!

Bemerkung 1.5.4. Für eine additiv geschriebene Gruppe $(G, +)$ ändern sich die Bezeichnungen und die Rechenregeln folgendermaßen:

i) Man schreibt "0" für das neutrale Element und nennt es das Nullelement von $(G, +)$. Es gilt $a + 0 = 0 + a = a$ für alle $a \in G$.

ii) Man schreibt $-a$ für das Inverse statt a' und $a - b$ als Abkürzung für $a + (-b)$. Es gilt somit $a - a = -a + a = a + (-a) = 0$ für alle $a \in G$.

iii) Die Gleichungen $a + x = b$ und $y + a = b$ sind eindeutig lösbar mit $x = -a + b$ und $y = b - a$.

iv) Es gilt $-(-a) = a$ und

$$-(a_1 + a_2 + \cdots + a_n) = -a_n - a_{n-1} - \cdots - a_1 = (-a_n) + (-a_{n-1}) + \cdots + (-a_1).$$

Üblicherweise benützt man die additive Schreibweise nur für abelsche Gruppen.

Eine zentrale Frage der Algebra befasst sich mit Unterstrukturen einer gegebenen Struktur. Ist die gegebene Struktur eine Gruppe, so handelt es sich um eine Untergruppe.

Definition 1.5.3. Es sei $(G, \circ)$ eine Gruppe und $U \subset G$ eine Teilmenge von G. Dann heißt $(U, \circ)$ Untergruppe von G, falls es ebenfalls eine Gruppe ist.

Beispiel 1.5.9. Es sei $(G, \circ) = (\mathbb{Z}, +)$ und $U = \{-1, 0, 1\}$. Somit ist $(U, +)$ keine Untergruppe von $(\mathbb{Z}, +)$, da keine Abgeschlossenheit vorliegt. Die Summe $1 + 1 = 2$ ist zwar durch die auf G gegebene Addition definiert, liegt aber nicht in U.

Beispiel 1.5.10. Es sei $(G, \circ) = (\mathbb{R}\backslash\{0\}, \cdot)$ und $H = \{-1, 1\}$. Hier ist $(H, \cdot)$ eine Untergruppe von $(G, \cdot)$. Die Multiplikation ist abgeschlossen. Weiter enthält H das neutrale Element 1 und ist auch abgeschlossen bzgl. der Inversenbildung. Wir haben die Verknüpfungstafel

$\cdot$	1	-1
1	1	-1
-1	-1	1

Das Assoziativgesetz gilt auf H, da es auf der gesamten Menge G gilt, in der H enthalten ist.

Zur Nachprüfung der Untergruppeneigenschaft müssen nicht sämtliche Eigenschaften der Gruppenverknüpfung nachgewiesen werden. Es genügt vielmehr die Überprüfung des folgenden Kriteriums:

Satz 1.5.4. *(Untergruppenkriterium)*
Es sei $(G, \circ)$ eine Gruppe und $U \subset G$. Dann ist $(U, \circ)$ genau dann Untergruppe von $(G, \circ)$, wenn

i) Die Menge U ist nichtleer, also $U \neq \emptyset$.

ii) Für alle $a, b \in U$ gilt $ab^{-1} \in U$.

Beweis. "$\Leftarrow$":
Es ist klar, dass (i) und (ii) für die Untergruppeneigenschaft notwendig sind.

"⇒":
Das Assoziativgesetz gilt, da es in der Menge G gilt, in der U enthalten ist. Wegen $U \neq \emptyset$ gibt es ein $a \in U$. Nach (ii) ist dann $e = aa^{-1} \in U$. Also hat U ein neutrales Element.
Es sei $b \in U$. Wegen $e, b \in U$ folgt nach (ii) nun

$$b^{-1} = eb^{-1}. \qquad (*)$$

Damit ist U bezüglich der Inversenbildung abgeschlossen.
Es seien $a, b \in U$. Wegen $a, b^{-1} \in U$ folgt nach (ii) und $(*)$

$$ab = a\left(b^{-1}\right)^{-1} \in U.$$

Dies zeigt die Abgeschlossenheit der Verknüpfung. □

1.6 Ebene und Raum als reelle Vektorräume, Unterraum

Wie schon in den Vorbemerkungen ausgeführt, hat sich die Lineare Algebra, in deren Mittelpunkt der Begriff des Vektorraums (VR) steht, aus der Elementargeormetrie entwickelt. Wir wollen hier zunächst die Definition des reellen Vektorraums geben und dann ausführen, dass Geraden, Ebenen und Raum- falls sie den Ursprung enthalten- unter diesen Begriff fallen. Den Vektorraumaxiomen liegen die Regeln (VA) und (VM) von Abschnitt 1.2 zugrunde. Nach Definition 1.5.2 können die Regeln (VA) kurz im Gruppenbegriff zusammengefasst werden.

Definition 1.6.1. Es sei V eine Menge mit einer Verknüpfung $\oplus$ sowie $\circ \colon \mathbb{R} \times V \to V$ eine Abbildung. Das Tripel $(V, \oplus, \circ)$ heißt ein reeller Vektorraum (oder Vektorraum über dem Körper der reellen Zahlen), wenn die folgenden Axiome für alle $\vec{u}, \vec{v} \in V$ und $\lambda, \mu \in \mathbb{R}$ gelten:

(V1) $(V, \oplus)$ bildet eine abelsche Gruppe.

(V2) Es gilt $(\lambda + \mu) \circ \vec{u} = (\lambda \circ \vec{u}) \oplus (\mu \circ \vec{u})$ sowie $\lambda \circ (\vec{u} \oplus \vec{v}) = (\lambda \circ \vec{u}) \oplus (\lambda \circ \vec{v})$. (Distributivgesetze)

(V3) Es ist $\lambda \circ (\mu \circ \vec{u}) = (\lambda \cdot \mu) \circ \vec{u}$. (Assoziativgesetz)

(V4) Schließlich ist $1 \circ \vec{u} = \vec{u}$. (Unitaritätsgesetz)

Die Elemente von V heißen Vektoren, die von $\mathbb{R}$ Skalare. Die Verknüpfung $\circ$ nennt man dann die äußere Multiplikation von Vektoren mit Skalaren; die Operationen $\oplus$ und $\circ$ nennt man auch lineare Operationen.
Statt $(V, \oplus, \circ)$ nennt man meist kurz V einen Vektorraum, wenn aus dem Zusammenhang Klarheit über $\oplus$ und $\circ$ besteht.

Bemerkung 1.6.1. In der Vorlesung bezeichnen wir Vektoren mit einem Pfeil, also $\vec{v}$, und schreiben $\vec{0}$ für das Nullelement der abelschen Gruppe $(V, \oplus)$, genannt Nullvektor. In der Literatur gebräuchlich sind auch die altdeutschen Buchstaben $\mathfrak{u}$, Unterstriche $\underline{u}$ oder Fettdruck $\mathbf{u}$, um Vektoren von Skalaren zu unterscheiden. Sobald wir mit dem Sachverhalt vertraut sind, schreiben wir statt $\oplus$ und $\circ$ einfach $+$ und $\cdot$.

Wir stellen einige einfache Rechenregeln zusammen:

Satz 1.6.1. *In einem reellen Vektorraum $(V, \oplus, \circ)$ gilt für $\vec{v}, \vec{w} \in V$ und $\lambda, \mu \in \mathbb{R}$:*

i) $\lambda \circ \vec{v} = \vec{0} \Leftrightarrow \lambda = 0$ *oder* $\vec{v} = \vec{0}$

ii) $(-\lambda) \circ \vec{v} = \lambda \circ (\ominus \vec{v}) = \ominus(\lambda \circ \vec{v})$, *wobei* $\ominus$ *die Inversion in* $(V, \oplus)$ *ist.*

iii) $(\lambda - \mu) \circ \vec{v} = \lambda \circ \vec{v} \ominus \mu \circ \vec{v}$ *sowie* $\lambda \circ (\vec{v} \ominus \vec{w}) = \lambda \circ \vec{v} \ominus \lambda \circ \vec{w}$

Beweis. i) Zur Richtung "$\Leftarrow$":

Fall $\lambda = 0$:

$$0 \circ \vec{v} = (0+0) \circ \vec{v} \underset{(V2)}{=} 0 \circ \vec{v} \oplus 0 \circ \vec{v}.$$

Es ist auch $0 \circ \vec{v} \oplus \vec{0} = 0 \circ \vec{v}$. Also hat die Gleichung $0 \circ \vec{v} \oplus \vec{x} = 0 \circ \vec{v}$ die Lösungen $\vec{x} = 0 \circ \vec{v}$ und $\vec{x} = \vec{0}$. Nach Satz 1.5.2 ist die Lösung dieser Gleichung aber eindeutig bestimmt. Damit folgt $0 \circ \vec{v} = \vec{0}$.

Andererseits gilt für $\vec{v} = \vec{0}$:

$$\lambda \circ \vec{0} = \lambda \circ (\vec{0} + \vec{0}) \underset{(V2)}{=} \lambda \circ \vec{0} \oplus \lambda \circ \vec{0}.$$

Wiederum folgt nach Satz 1.5.2 nun $\lambda \circ \vec{0} = \vec{0}$.

Zur Richtung "$\Rightarrow$":

Es sei $\lambda \circ \vec{v} = \vec{0}$. Ist $\lambda = 0$, so sind wir fertig. Ist $\lambda \neq 0$, so existiert $\lambda^{-1} \in \mathbb{R}$, und es folgt

$$\vec{v} \underset{(V4)}{=} 1 \circ \vec{v} = (\lambda^{-1} \cdot \lambda) \circ \vec{v} \underset{(V3)}{=} \lambda^{-1} \circ (\lambda \circ \vec{v}) = \lambda^{-1} \circ \vec{0} = \vec{0}.$$

ii) Es gilt

$$\lambda \circ \vec{v} \oplus (-\lambda) \circ \vec{v} \underset{(V2)}{=} (\lambda - \lambda) \circ \vec{v} = 0 \circ \vec{v} \underset{(i)}{=} \vec{0},$$

also $(-\lambda) \circ \vec{v} = \ominus(\lambda \circ \vec{v})$. Ebenso folgt aus

$$\lambda \circ \vec{v} \oplus \lambda \circ (\ominus \vec{v}) \underset{(V2)}{=} \lambda \circ (\vec{v} \ominus \vec{v}) = \lambda \circ \vec{0} \underset{(i)}{=} \vec{0},$$

die Gleichung $\lambda \circ (\ominus \vec{v}) = \ominus(\lambda \circ \vec{v})$.

iii) Dies folgt sofort aus (V2) und (ii).

□

Beispiel 1.6.1. Die in den früheren Abschnitten diskutierte Mengen von Vektoren: $\mathbb{R}^2$ (Ebene), $\mathbb{R}^3$ (Raum) und allgemein der $\mathbb{R}^n$ mit $n \in \mathbb{N}$ erfüllen nun zusammen mit den auf ihnen definierten Rechenoperationen Vektoraddition und Skalarmultiplikation die Axiome (V1)- (V4) von Definition 1.5.2. Wir haben also die reellen Vektorräume $(\mathbb{R}^2, +, \cdot)$, $(\mathbb{R}^3, +, \cdot)$ und allgemein den $(\mathbb{R}^n, +, \cdot)$.

Wir betrachten nun ein Beispiel, das von der Elementargeometrie weit entfernt zu sein scheint:

Beispiel 1.6.2. Es sei $n \in \mathbb{N}$.
Unter einem Polynom vom Grad $n \in \mathbb{N}$ versteht man die Funktion $p \colon \mathbb{R} \to \mathbb{R}$, $x \to p(x)$, wobei $p(x) = a_0 + a_1 x + \ldots + a_n x^n$ mit $a_i \in \mathbb{R}$ fest und $a_n \neq 0$ ist.
Es sei nun P_n die Menge aller Polynome vom Grad $\leq n$. Polynome können addiert werden: Ist $p(x) = a_0 + a_1 x + \ldots + a_n x^n$ und $q(x) = b_0 + b_1 x + \ldots + b_n x^n$, so ist

$$\begin{aligned} (p+q)(x) &= (a_0 + a_1 x + \ldots + a_n x^n) + (b_0 + b_1 x + \ldots + b_n x^n) \\ &= (a_0 + b_0) + (a_1 + b_1)x + \ldots + (a_n + b_n)x^n \end{aligned}$$

wieder ein Polynom. Ebenso können Polynome mit Skalaren multipliziert werden:

$$(\lambda \cdot p)(x) = (\lambda a_0) + (\lambda a_1)x + \ldots + (\lambda a_n)x^n.$$

Man sieht leicht, dass mit dieser Addition und Skalarmultiplikation die Axiome (V1)- (V4) erfüllt sind, wenn man als Nullvektor das Nullpolynom mit $0 \colon x \to 0 + 0 \cdot x + \ldots 0 \cdot x^n = 0$ nimmt.
Man hat also den reellen Vektorraum $(P_n, +, \cdot)$.

An diesem Beispiel werden die Vorteile klar, die eine Verallgemeinerung einer Theorie haben kann. Sie umfasst eine größere Menge an Beispielen als die speziellere Theorie. Ideen, die in dieser- in unserem Beispiel der Elementargeometrie- entstanden sind, können auf auf die übrigen Beispiele übertragen werden.
Wir kehren zur Elementargeometrie (und zum $\mathbb{R}^n$ zurück): So wie man in Gruppen nach Untergruppen suchen kann, so kann man in Vektorräumen nach Unterräumen suchen.

Definition 1.6.2. Es sei $(V, \oplus, \circ)$ ein Vektorraum und $U \subset V$ eine Teilmenge von V. Dann heißt $(U, \oplus, \circ)$ ein Unterraum von V, wenn es ebenfalls ein Vektorraum ist.

Wiederum ist die Abgeschlossenheit von Bedeutung, dieses Mal nicht nur die der Addition, sondern auch der Skalarmultiplikation.

Beispiel 1.6.3. Es sei $V = \mathbb{R}^2$ und $U = \{(x, y) \colon x^2 + y^2 = 1\}$.
Wir werden später sehen, dass U der Einheitskreis, der Kreis mit dem Ursprung als Mittelpunkt und Radius 1. Wie man leicht sieht, ist die Addition nicht abgeschlossen. Es ist $\vec{v}_1 = (1, 0) \in U$ und $\vec{v}_2 = (0, 1) \in U$, aber $\vec{v}_1 + \vec{v}_2 = (1, 1) \notin U$.
Also ist U kein Unterraum von V.

Beispiel 1.6.4. Es sei $V = \mathbb{R}^2$ und $U = \{(m, m) \colon m \in \mathbb{Z}\}$.
Die Menge U ist abgeschlossen bezüglich der Addition, jedoch nicht abgeschlossen bezüglich der Skalarmultiplikation. Zum Beispiel ist $(1, 1) \in U$, aber für $\lambda = \frac{1}{3}$ ist $\lambda \cdot (1, 1) = \left(\frac{1}{3}, \frac{1}{3}\right) \notin U$.
Damit ist U kein Unterraum von V.

Wir geben nun in Analogie zum Untergruppenkriterium aus Satz 1.5.4 ein Unterraumkriterium, aus dem die Unterraumeigenschaft einer Teilmenge U eines Vektorraums folgt.

Satz 1.6.2. *Es sei V ein reeller Vektorraum und $U \subset V$ eine nichtleere Teilmenge von V. Dann ist U genau dann ein Unterraum von V, wenn*

$$\vec{u}, \vec{v} \in U \Rightarrow \lambda\vec{u} + \mu\vec{v} \in U \ \forall \lambda, \mu \in \mathbb{R} \tag{UR}$$

gilt.

Beweis. "$\Leftarrow$":
Es gelte (UR). Zu zeigen ist, dass U ein Unterraum von V ist. Die Rechenregeln (Assoziativ- Kommutativ- und Distributivgesetze sowie das Unitaritätsgesetz) gelten auf U, da sie auf der Obermenge V gelten. Es bleibt zu zeigen, dass $(U, +)$ eine Untergruppe von $(V, +)$ ist, sowie die Abgeschlossenheit bzgl. der Skalarmultiplikation.
Es seien $\vec{u}, \vec{v} \in U$. In (UR) wählen wir $\lambda = 1$ und $\mu = -1$. Also ist damit $\lambda\vec{u} + \mu\vec{v} = \vec{u} - \vec{v} \in U$. Nach Satz 1.5.4 ist $(U, +)$ eine Untergruppe von $(V, +)$. Wählen wir $\lambda \in \mathbb{R}$ und $\mu = 0$, so folgt $\lambda\vec{u} \in U$. Also ist U auch abgeschlossen bzgl. der Skalarmultiplikation.

"$\Rightarrow$":
Es sei nun U ein Unterraum von V. Zu zeigen ist, dass (UR) gilt.
Es seien also $\vec{u}, \vec{v} \in U$ und $\lambda, \mu \in \mathbb{R}$. Wegen der Abgeschlossenheit bzgl. der Skalarmultiplikation sind $\lambda\vec{u} \in U$ und $\mu\vec{v} \in U$. Wegen der Untergruppeneigenschaft von $(U, +)$ ist $\lambda\vec{u} + \mu\vec{v} \in U$, also gilt (UR). □

Bemerkung 1.6.2. Der Ausdruck $\lambda\vec{u}+\mu\vec{v}$ in (UR) ist ein einfaches Beispiel einer Linearkombination. Dieser Begriff ist zentral in der Linearen Algebra.

Definition 1.6.3. Es sei V ein reeller Vektorraum und $\vec{v}_1, \ldots, \vec{v}_n \in V$. Eine Linearkombination von $\vec{v}_1, \ldots, \vec{v}_n$ ist eine Summe der Form $\lambda_1\vec{v}_1 + \ldots + \lambda_n\vec{v}_n$ mit $\lambda_j \in \mathbb{R}$. Die λ_j heißen die Koeffizienten der Linearkombination.

Bemerkung 1.6.3. Linearkombinationen sind uns schon früher begegnet:

i) In der Punkt- Richtungs- Gleichung der Geraden $\vec{PX} = \vec{PA} + t\vec{c}$ aus Abschnitt 1.2 ist der Summand $t\vec{c}$ eine Linearkombination des Richtungsvektors $\vec{c}$.

ii) In der Beschreibung der Ebene in der Form $\vec{AX} = t\vec{u} + s\vec{v}$ ist die rechte Seite eine Linearkombination der Richtungsvektoren $\vec{u}$ und $\vec{v}$.

Wir werden nun die allgemeine Regel formulieren, die hinter diese Beobachtungen steckt. Wir beginnen mit dem einfachen

Satz 1.6.3. *Der Durchschnitt beliebig vieler Unterräume eines Vektorraums ist wieder ein Unterraum.*

Beweis. Es sei V ein reeller Vektorraum und U_j mit $j \in J$ seien Unterräume von V. Es sei

$$U = \bigcap_{j \in J} U_j.$$

Es seien $\vec{u}, \vec{v} \in U$. Dann gilt auch $\vec{u}\vec{v} \in U_j$ für alle $j \in J$. Da U_j Unterräume sind, gilt nach dem Unterraumkriterium (Satz 1.6.2) auch $\lambda\vec{u} + \mu\vec{v} \in U_j$ für alle $\lambda, \mu \in \mathbb{R}$. Daraus folgt

$$\lambda\vec{u} + \mu\vec{v} \in \bigcap_{j \in J} U_j = U$$

für alle $\lambda, \mu \in \mathbb{R}$. Damit erfüllt auch U das Unterraumkriterium, weswegen U ein Unterraum von V ist. $\square$

Definition 1.6.4. Es sei V ein reeller Vektorraum und $M \subset V$.
Unter dem von M erzeugten oder aufgespannten Unterraum von V (auch Erzeugnis von M oder lineare Hülle von M mit der Schreibweise $\langle M \rangle$ versteht man den Durchschnitt aller Unterräume von V, die M enthalten. Im Falle einer endlichen Menge $M = \{\vec{v}_1, \ldots, \vec{v}_n\}$ schreibt man auch $\langle \vec{v}_1, \ldots, \vec{v}_n \rangle$ statt $\langle \{\vec{v}_1, \ldots, \vec{v}_n\} \rangle$.

Beispiel 1.6.5. Das Erzeugnis der leeren Menge $\emptyset$ ist der Nullraum $\{\vec{0}\}$.

Definition 1.6.5. Es sei V ein reeller Vektorraum und U ein Unterraum von V. Eine Teilmenge $M \subset V$ heißt Erzeugendensystem von U, wenn $U = \langle M \rangle$ ist.

Definition 1.6.6. Es sei V ein reeller Vektorraum und $M \subset V$. Dann setzen wir

$$L(M) = \{\lambda_1\vec{v}_1 + \ldots + \lambda_n\vec{v}_n \colon n \in \mathbb{N},\ \vec{v}_1, \ldots, \vec{v}_n \in M,\ \lambda_1, \ldots, \lambda_n \in \mathbb{R}\}$$

als die Menge aller Linearkombinationen von Elementen von M.

Bemerkung 1.6.4. Die leere Summe, die null Summanden enthält, gilt als Nullvektor $\vec{0}$. Somit ist $L(\emptyset) = \{\vec{0}\}$.

Satz 1.6.4. *Es sei V ein reeller Vektorraum und $M \subset V$. Dann ist $L(M)$ ein Unterraum von V.*

Beweis. Wir wenden das Unterraumkriterium (Satz 1.6.2) an. Es seien $\vec{u}, \vec{v} \in L(M)$. Dann gibt es $\kappa_1, \ldots, \kappa_l, \lambda_1, \ldots, \lambda_m \in \mathbb{R}$ und $\vec{u}_1, \ldots, \vec{u}_l, \vec{w}_1, \ldots, \vec{w}_m \in M$ mit

$$\vec{u} = \kappa_1 \vec{u}_1 + \ldots + \kappa_l \vec{u}_l \quad \text{und} \quad \vec{v} = \lambda_1 \vec{v}_1 + \ldots + \lambda_m \vec{v}_m.$$

Durch mögliches Hinzufügen von Summanden $0 \cdot \vec{u}_i$ bzw. $0 \cdot \vec{v}_j$ kann erreicht werden, dass in den Darstellungen von $\vec{u}$ und $\vec{v}$ dieselben Vektoren und dieselbe Anzahl von Summanden auftreten. Es ist $\vec{u} = \mu_1 \vec{w}_1 + \ldots + \mu_n \vec{w}_n$ und $\vec{v} = \nu_1 \vec{w}_1 + \ldots + \nu_n \vec{w}_n$ mit $\mu_i, \nu_i \in \mathbb{R}$ und $\vec{w}_j \in M$. Dann ist mit $\alpha, \beta \in \mathbb{R}$ auch

$$\alpha \vec{u} + \beta \vec{v} = (\alpha \mu_1 + \beta \nu_1) \vec{w}_1 + \ldots + (\alpha \mu_n + \beta \nu_n) \vec{w}_n \in L(M),$$

womit $L(M)$ das Unterraumkriterium erfüllt. □

Satz 1.6.5. *Es sei V ein reeller Vektorraum und $M \subset V$. Dann ist $\langle M \rangle = L(M)$, d.h. $\langle M \rangle$ ist die Menge aller Linearkombinationen von Elementen aus M.*

Beweis. Wir zeigen zunächst: Ist U ein Unterraum von V mit $M \subset U$, so gilt

$$L(M) \subset U. \tag{1}$$

Wir führen den Beweis durch vollständige Induktion nach der Anzahl n der Summanden:
$n = 0$:
Für den Nullvektor gilt $\vec{0} \in U$ nach dem Unterraumkriterium (Satz 1.6.2).
$n \to n + 1$:
Es seien $\vec{v}_1, \ldots, \vec{v}_{n+1} \in M$ und $\lambda_1, \ldots, \lambda_{n+1} \in \mathbb{R}$.
Nach der Induktionshypothese ist $\vec{u} := \lambda_1 \vec{v}_1 + \ldots + \lambda_n \vec{v}_n \in U$. Das Unterraumkriterium mit $\vec{v} = \vec{v}_{n+1}$, $\lambda = 1$ und $\mu = \lambda_{n+1}$ ergibt

$$\lambda \vec{u} + \mu \vec{v} = \lambda_1 \vec{v}_1 + \ldots + \lambda_n \vec{v}_n + \lambda_{n+1} \vec{v}_{n+1} \in U,$$

womit (1) gezeigt ist. Wegen $M \subset \langle M \rangle$ folgt aus (1)

$$L(M) \subset \langle M \rangle . \tag{2}$$

Andererseits gilt für jeden Unterraum U von V mit $M \subset U$

$$\langle M \rangle \subset U. \tag{3}$$

Da dies nach Satz 1.6.4 ein Unterraum von V ist, können wir (3) mit $U = L(M)$ anwenden und erhalten

$$\langle M \rangle \subset L(M). \tag{4}$$

Aus (2) und (4) folgt die Behauptung. □

Wir betrachten nun einige Spezialfälle:

Beispiel 1.6.6. Es sei $n \in \mathbb{N}$ und $\vec{v}_1 \in \mathbb{R}^n \backslash \{\vec{0}\}$. Nach Satz 1.6.5 ist $\langle \vec{v}_1 \rangle = \{t \vec{v}_1 \colon t \in \mathbb{R}\}$ die Gerade durch $\vec{0}$ mit Richtungsvektor $\vec{v}$. Es sei U ein Unterraum von $\langle \vec{v} \rangle$. Wir unterscheiden zwei Fälle:
Fall 1: $U = \{\vec{0}\}$.
Fall 2: $U \neq \{\vec{0}\}$. Dann gibt es $s \neq 0$ mit $s\vec{v}_1 \in U$, weswegen U auch alle Vielfachen von $s\vec{v}_1$ enthält. Für $t \in \mathbb{R}$ ist $t\vec{v}_1 = ts^{-1}(s\vec{v}_1) \in U$, d. h. $U = \langle \vec{v}_1 \rangle$.
Damit enthält $\langle \vec{v}_1 \rangle$ nur zwei Unterräume: den Nullraum $\{\vec{0}\}$ und den Gesamtraum $\langle \vec{v}_1 \rangle$.

Beispiel 1.6.7. Es sei $n \in \mathbb{N}$ und $\vec{v}_1, \vec{v}_1 \in \mathbb{R}^n$, wobei $\vec{v}_1$ und $\vec{v}_2$ nicht parallel sind. Nach Satz 1.6.5 ist $E = \langle \vec{v}_1, \vec{v}_2 \rangle = \{t_1\vec{v}_1 + t_2\vec{v}_2 \colon t_1, t_2 \in \mathbb{R}\}$ die Ebene durch $\vec{0}$ mit den Richtungsvektoren $\vec{v}_1$ und $\vec{v}_2$. Wir untersuchen die Unterräume von E. Es sei U ein Unterraum von E.
Fall 1: $U = \{\vec{0}\}$.
Fall 2: $U \neq \{\vec{0}\}$. Dann gibt es $\vec{w}_1 \in U - \{\vec{0}\}$.
Unterfall 1: Für alle $\vec{w} \in U$ gilt $\vec{w} = s_1\vec{w}_1$ mit $s_1 \in \mathbb{R}$.
Dann ist $U = \langle \vec{w}_1 \rangle$, die Gerade durch den Ursprung mit Richtungsvektor $\vec{w}_1$.
Unterfall 2: Es existiert ein $\vec{w}_2 \in U$ mit $\vec{w}_2 \notin \langle \vec{w}_1 \rangle$. Wir behaupten $E = \langle \vec{w}_1, \vec{w}_2 \rangle$.
Wegen $\vec{v}_1, \vec{v}_2 \in E$ gibt es $\lambda_{11}, \lambda_{12}, \lambda_{21}, \lambda_{22} \in \mathbb{R}$ mit

$$\vec{w}_1 = \lambda_{11}\vec{v}_1 + \lambda_{21}\vec{v}_2 \quad \text{bzw.} \quad \vec{w}_2 = \lambda_{12}\vec{v}_1 + \lambda_{22}\vec{v}_2. \tag{1}$$

Es sei

$$\vec{v} = t_1\vec{v}_1 + t_2\vec{v}_2 \in E. \tag{2}$$

Wir wollen nun die Existenz von x_1 und x_2 mit $x_1\vec{w}_1 + x_2\vec{w}_2 = \vec{v}$ zeigen. Einsetzen von (1) und (2) führt auf das LGS

$$\begin{aligned} \lambda_{11}x_1 + \lambda_{12}x_2 &= t_1 \\ \lambda_{21}x_1 + \lambda_{22}x_2 &= t_2 \end{aligned}$$

oder

$$\begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}.$$

Wegen $\vec{w}_2 \notin \langle \vec{w}_1 \rangle$ ist $(\lambda_{11}, \lambda_{12}) \neq (0,0)$ und $(\lambda_{21}, \lambda_{22}) \neq (0,0)$. Durch den Gaußschen Algorithmus kann das LGS (nach möglicher Umnummerierung der Unbestimmten x_1, x_2) in folgende Form gebracht werden:

$$\begin{pmatrix} 1 & \lambda_{12}^{(1)} \\ 0 & \lambda_{22}^{(1)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}.$$

Annahme: $\lambda_{22}^{(1)} = 0$
Dann gibt es $u \in \mathbb{R}$ mit $\lambda_{21} = u\lambda_{11}$ und $\lambda_{22} = u\lambda_{12}$. Die Koeffizientenmatrix hat dann die Form

$$\begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{pmatrix} = \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ u\lambda_{11} & u\lambda_{12} \end{pmatrix}.$$

Man sieht, dass dann auch die Spalten der Matrix Vielfache voneinander sind, im Widerspruch zu $\vec{w}_2 \notin \langle \vec{w}_1 \rangle$. Somit ist $\lambda_{22}^{(1)} \neq 0$, d.h. das LGS ist für beliebige t_1, t_2 lösbar. Es ist $U = E$.

Wir erhalten:
Die Ebene E hat die Unterräume $\{\vec{0}\}$, $\{t\vec{v}_1 \colon t \in \mathbb{R}\}$, also Geraden durch den Ursprung, und E selbst.

Geraden und Ebenen, die nicht den Nullpunkt enthalten, sind keine Unterräume. Sie fallen unter den allgemeinen Begriff der linearen Mannigfaltigkeit.

Definition 1.6.7. Es sei $(G, +)$ eine abelsche Gruppe und $S, T \subset G$. Unter der Summe $S + T$ der Mengen S und T versteht man

$$S + T = \{s + t \colon s \in S, t \in T\}.$$

Besteht S nur aus einem einzigen Element, also $S = \{s\}$, so schreibt man für $S + T$ statt $\{s\} + T$ auch $s + T$.

Beispiel 1.6.8. Es sei $S = \{1, 3, 4\}$ und $T = \{-2, 5, 8, 9\}$. Dann ist

$$\begin{aligned} S + T &= \{1 + (-2), 1 + 5, 1 + 8, 1 + 9, 3 + (-2), 3 + 5, 3 + 8, 3 + 9, 4 + (-2), 4 + 5, 4 + 8, 4 + 9\} \\ &= \{-1, 1, 2, 6, 8, 9, 10, 11, 12, 13\}. \end{aligned}$$

Definition 1.6.8. Es sei V ein reeller Vektorraum. Unter einer linearen Mannigfaltigkeit M (von V) versteht man eine Menge der Gestalt $M = \vec{v}_0 + U$ mit $\vec{v}_0 \in V$ und einem Unterraum U von V.

Beispiel 1.6.9. Geraden und Ebenen des $\mathbb{R}^n$ sind lineare Mannigfaltigkeiten. Eine Gerade g ist eine Menge der Form $g = \{\vec{v}_0 + U_1\}$ mit $U_1 = \{t\vec{c}\colon t \in \mathbb{R}\}$, dem vom Richtungsvektor $\vec{c}$ aufgespannten Unterraum. Dann geht g aus U_1 durch Parallelverschiebung um $\vec{v}_0$ hervor.
Eine Ebene E ist eine Menge der Form $E = \{\vec{v}_0 + U_2\}$ mit $U_2 = \{s\vec{u} + t\vec{v}\colon s, t \in \mathbb{R}\}$, dem von den Richtungsvektoren $\vec{u}$ und $\vec{v}$ aufgespannten Unterraum. Dann geht analog E aus U_2 durch Parallelverschiebung um $\vec{v}_0$ hervor.

1.7 Lineare Unabhängigkeit, Basis, Dimension

Wir betrachten die Definition der Geraden als Punktrichtungsgleichung $g(\vec{v}_1) = \{\vec{0A} + t_1\vec{v}_1 : t_1 \in \mathbb{R}\}$ mit

$$\vec{v}_1 \neq \vec{0} \tag{1}$$

und der Ebene $E = E(\vec{v}_1, \vec{v}_2) = \{\vec{0A} + t_1\vec{v}_1 + t_2\vec{v}_2 : t_1, t_2 \in \mathbb{R}\}$, wobei

$$\vec{v}_1 \text{ und } \vec{v}_2 \text{ nicht parallel sind,} \tag{2}$$

von Abschnitt 1.2.

Zur Beschreibung der Punkte einer Geraden $g(\vec{v}_1)$ ist also ein Parameter t_1 notwendig, in der Umgangssprache bezichnen wir die Gerade als eindimensional. Zur Beschreibung der Punkte einer Ebene $E(\vec{v}_1, \vec{v}_2)$ sind zwei Parameter t_1 und t_2 notwendig, die Ebene wird zweidimensional genannt.

Was passiert nun, wenn die Bedingungen (1) und (2) nicht erfüllt sind?

1. Ist $\vec{v}_1 = \vec{0}$, so gilt wegen $t\vec{v}_1 = \vec{0}$ für alle $t \in \mathbb{R}$ folglich $g(\vec{v}_1) = \{\vec{OA}\}$, d.h. die Menge g ist keine Gerade, also ein eindimensionales Objekt, dessen Punkte durch einen Parameter beschrieben werden, sondern nur noch ein Punkt, ein nulldimensionales Objekt, zu dessen Beschreibung kein Parameter mehr nötig ist. Das Erzeugendensystem $\{\vec{v}_1\}$ kann zur leeren Menge verkleinert werden.

2. Es seien $\vec{v}_1$ und $\vec{v}_2$ parallel. Wir nehmen an, dass $\vec{v}_1 \neq \vec{0}$ und $\vec{v}_2 \neq 0$ gilt. Dann gibt es $\lambda \in \mathbb{R}\backslash\{0\}$ mit $\vec{v}_2 = \lambda\vec{v}_1$. Es ist dann

$$E(\vec{v}_1, \vec{v}_2) = \{\vec{0A} + t_1\vec{v}_1 + t_2\vec{v}_2 : t_1, t_2 \in \mathbb{R}\} = \{\vec{0A} + (t_1 + t_2\lambda)\vec{v}_1\} = \{\vec{0A} + t_3\vec{v}_1 : t_3 \in \mathbb{R}\}$$

keine Ebene, ein zweidimensionales Objekt, dessen Punkte durch zwei Parameter beschrieben werden, sondern nur noch eine Gerade, ein eindimensionales Objekt, dessen Punkte durch einen Parameter beschrieben werden. Das Erzeugendensystem $\{\vec{v}_1, \vec{v}_2\}$ kann zu $\{\vec{v}_1\}$ verkleinert werden.

Wir suchen nun nach einer geeigneten Formulierung für diese Ausnahmesituation:
Im Fall (1) ist $\lambda_1\vec{v}_1 = \vec{0}$ für alle $\lambda_1 \in \mathbb{R}$.
Im Fall (2) kann ein Vektor als Vielfaches (eine spezielle Linearkombination) der anderen ausgedrückt werden:

$$\vec{v}_2 = \lambda_1\vec{v}_1 \quad \text{oder auch} \quad \lambda_1\vec{v}_1 + (-1)\vec{v}_2 = \vec{0}.$$

Beide Fälle können so zusammengefasst werden:
Es gibt $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, wobei nicht alle $\lambda_i = 0$ sind, so dass $\lambda_1\vec{v}_1 + \ldots + \lambda_n\vec{v}_n = \vec{0}$ gilt.

Dies gibt Anlass zu folgender

Definition 1.7.1. Es sei V ein reeller Vektorraum. Dann heißen $\vec{v}_1, \ldots \vec{v}_n$ linear abhängig (l.a.), wenn es $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ gibt, so dass $\lambda_i \neq 0$ für mindestens ein i mit $1 \leq i \leq n$ mit $\lambda_1\vec{v}_1 + \ldots + \lambda_n\vec{v}_n = \vec{0}$ gilt. Im gegenteiligen Fall heißen $\vec{v}_1, \ldots \vec{v}_n$ linear unabhängig (l.u.).

Es ergibt sich folgende Formulierung der linearen Unabhängigkeit:
Die Vektoren $\vec{v}_1, \ldots, \vec{v}_n$ sind genau dann linear unabhängig, wenn aus $\lambda_1\vec{v}_1 + \ldots + \lambda_n\vec{v}_n = \vec{0}$ auch $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$ folgt, d.h. die einzige Linearkombination von $\vec{v}_1, \ldots, \vec{v}_n$, die den Nullvektor darstellt, ist diejenige, in welcher sämtliche Koeffizienten verschwinden.

Unsere obigen Beobachtungen können nun verallgemeinert werden: Ein linear abhängiges Erzeugendensystem kann stets verkleinert werden.

Satz 1.7.1. *Es sei V ein reeller Vektorraum und $U \subset V$ ein Unterraum. Ist $U = \langle \vec{v}_1, \ldots, \vec{v}_n \rangle$ und sind $\vec{v}_1, \ldots, \vec{v}_n$ linear abhängig, so gibt es eine echte Teilfolge $\vec{v}_{i_1}, \ldots, \vec{v}_{i_m}$, so dass $U = \langle \vec{v}_{i_1}, \ldots, \vec{v}_{i_m} \rangle$ gilt.*

Beweis. Nach Definition 1.6.4 und Satz 1.6.5 ist $U = \{t_1\vec{v}_1 + \ldots + t_n\vec{v}_n \colon t_i \in \mathbb{R}\}$. Nach der Defintion der linearen Abhängigkeit (Definition 1.7.1) gibt es $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, die nicht alle verschwinden, so dass $\lambda_1\vec{v}_1 + \ldots + \lambda_n\vec{v}_n = \vec{0}$ gilt. O.B.d.A. sei $\lambda_n \neq 0$. Dann ist

$$\vec{v}_n = -\lambda_n^{-1}\lambda_1\vec{v}_1 - \ldots - \lambda_n^{-1}\lambda_{n-1}\vec{v}_{n-1}.$$

Aus jeder Linearkombination von $\vec{v}_1, \ldots \vec{v}_n$ kann dann $\vec{v}_n$ eliminiert werden:

$$t_1\vec{v}_1 + \ldots + t_{n-1}\vec{v}_{n-1} + t_n\vec{v}_n = (t_1 - t_n\lambda_n^{-1}\lambda_1)\vec{v}_1 + \ldots + (t_{n-1} - t_n\lambda_n^{-1}\lambda_{n-1})\vec{v}_n.$$

Es ist also $U = \langle \vec{v}_1, \ldots, \vec{v}_{n-1} \rangle$. □

Wenn wir nun versuchen, eine Definition für den der Umgangssprache entstammenden Begriff der Dimension zu geben, so scheint die Anzahl der Vektoren eines Erzeugendensystems, das nicht mehr verkleinert werden kann, d.h. eines linear unabhängigen Erzeugendensystems, die entscheidende Rolle zu spielen.

Definition 1.7.2. (Basis)
Es sei V ein reeller Vektorraum. Eine Basis von V ist ein linear unabhängiges Erzeugendensystem von V. Ist $\{\vec{v}_1, \ldots, \vec{v}_n\}$ eine Basis von V, so schreibt man auch $V = \langle \langle \vec{v}_1, \ldots, \vec{v}_n \rangle \rangle$

Man könnte nun erwägen, die Dimension eines Vektorraum als die Anzahl der Vektoren einer Basis zu definieren. Da jedoch zuerst noch die Existenz einer Basis gezeigt werden muss, empfiehlt sich eine andere Definition.

Definition 1.7.3. (Dimension)
Es sei V ein reeller Vektorraum. Unter der Dimension von V, $\dim V$, versteht man die Maximalzahl linear unabhängiger Vektoren in V.

Bemerkung 1.7.1. Wir werden später sehen, dass für einen Vektorraum mit $\dim V = n$ jede Basis n Elemente hat.

Beispiel 1.7.1. Der $\mathbb{R}^n$ hat die Standardbasis $\{\vec{e}_1, \ldots, \vec{e}_n\}$ mit

$$\vec{e}_1 = (1, 0, \ldots, 0)^T, \vec{e}_2 = (0, 1, 0, \ldots, 0)^T, \ldots, \vec{e}_n = (0, \ldots, 0, 1)^T.$$

Beweis. Es sei $\vec{v} = (\lambda_1, \ldots, \lambda_n)^T \in \mathbb{R}^n$. Dann ist $\vec{v} = \lambda_1 \vec{e}_1 + \ldots \lambda_n \vec{e}_n$. Aus $\lambda_1 \vec{e}_1 + \ldots \lambda_n \vec{e}_n = \vec{0}$ folgt $\lambda_1 = \ldots = \lambda_n = 0$, weswegen $\vec{e}_1, \ldots, \vec{e}_n$ linear unabhängig sind. □

Die lineare Abhängigkeit bzw. Unabhängigkeit eines Erzeugendensystems hat eine weitere wichtige Konsequenz.

Satz 1.7.2. *Es sei V ein reeller Vektorraum und $V = \langle \vec{v}_1, \ldots \vec{v}_n \rangle$.*

i) Sind $\vec{v}_1, \ldots, \vec{v}_n$ linear unabhängig, so gibt es zu jedem $\vec{v} \in V$ genau eine Darstellung der Form $\vec{v} = \lambda_1 \vec{v}_1 + \ldots + \lambda_n \vec{v}_n$.

ii) Sind $\vec{v}_1, \ldots, \vec{v}_n$ linear abhängig, so gibt es zu jedem $\vec{v} \in V$ unendlich viele Darstellungen der Form $\vec{v} = \lambda_1 \vec{v}_1 + \ldots + \lambda_n \vec{v}_n$.

Beweis. i) Da $\{\vec{v}_1, \ldots, \vec{v}_n\}$ ein Erzeugendensystem von V ist, gibt es zu jedem $\vec{v} \in V$ mindestens eine Darstellung

$$\vec{v} = \lambda_1 \vec{v}_1 + \ldots + \lambda_n \vec{v}_n. \tag{1}$$

Es sei nun

$$\vec{v} = \mu_1 \vec{v}_1 + \ldots + \mu_n \vec{v}_n \tag{2}$$

eine beliebige Darstellung. Subtraktion von (1) und (2) ergibt

$$\vec{0} = (\lambda_1 - \mu_1)\vec{v}_1 + \ldots (\lambda_n - \mu_n)\vec{v}_n.$$

Wegen der linearen Unabhängigkeit der $\vec{v}_1, \ldots \vec{v}_n$ folgt $\lambda_1 - \mu_1 = \ldots = \lambda_n - \mu_n = 0$, also $\lambda_i = \mu_i$, woraus sich die Eindeutigkeit ergibt.

ii) Es sei $\vec{v} \in V$. Wiederum gibt es mindestens eine Darstellung

$$\vec{v} = \lambda_1 \vec{v}_1 + \ldots + \lambda_n \vec{v}_n. \tag{1}$$

Wegen der linearen Abhängigkeit der $\vec{v}_i$ gibt es $\nu_1, \ldots \nu_n \in \mathbb{R}$, die nicht alle verschwinden, mit

$$\nu_1 \vec{v}_1 + \ldots + \nu_n \vec{v}_n = \vec{0}. \tag{2}$$

Es sei $\nu_j \neq 0$. Ein beliebiges Vielfaches von (2) kann nun zu (1) addiert werden. Für $t \in \mathbb{R}$ erhält man

$$\vec{v} = (\lambda_1 + t\nu_1)\vec{v}_1 + \ldots (\lambda_j + t\nu_j)\vec{v}_j + \ldots (\lambda_n + t\nu_n)\vec{v}_n.$$

Wegen $\nu_j \neq 0$ nimmt $\lambda_j + t\nu_j$ unendlich viele verschiedene Werte an, wenn t alle reellen Zahlen durchläuft.

□

Kapitel 2

Vektorräume

2.1 Ringe und Körper

Wir führen in diesem Kapitel den Begriff des Vektorraums ein, der sich als Verallgemeinerung des im vorigen Kapitel eingeführten Begriffs des reellen Vektorraums erweisen wird. Während bei reellen Vektorräumen die bei der Skalarmultiplikation verwendeten Skalare dem Körper $\mathbb{R}$ der reellen Zahlen angehören, entstammen hier die Skalare einem beliebigen Körper K. Man spricht dann auch von einem Vektorraum (VR) über (dem Körper) K. Wir wollen daher den Begriff des Körpers und zunächst den allgemeinen Begriff des Rings einführen.

Definition 2.1.1. Ein Ring ist ein Tripel $(R, +, \cdot)$, bestehend aus einer nichtleeren Menge R, einer als Addition bezeichneten Verknüpfung $+$ und einer als Multiplikation bezeichneten Verknüpfung $\cdot$, so dass folgende Ringaxiome erfüllt sind:

Es seien $a, b, c \in R$.

(R1) $(R, +)$ ist eine abelsche Gruppe.

(R2) Für alle $a, b, c \in R$ gilt $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (Assoziativgesetz für $\cdot$).

(R3) Es gelten die Distributivgesetze:

$$a \cdot (b + c) \;=\; (a \cdot b) + (a \cdot c) \;\text{ und }\; (b + c) \cdot a \;=\; (b \cdot a) + (c \cdot a)\,.$$

Gilt zusätzlich das Kommutativgesetz der Multiplikation:

(R4) Für alle $a, b \in R$ ist $a \cdot b = b \cdot a$,

so spricht man von einem kommutativen Ring. Das Nullelement der Gruppe $(R, +)$ wird mit 0 bezeichnet. Gibt es ein neutrales Element der Multiplikation, so heißt dieses das Einselement und wird mit 1 bezeichnet. Man spricht dann von einem Ring mit Eins.

Satz 2.1.1. *In einem Ring R gilt für alle $a, b \in \mathbb{R}$:*

i) $a \cdot 0 = 0 \cdot a = 0$

ii) $a \cdot (-b) = (-a) \cdot b = -(a \cdot b)$

iii) $(-a) \cdot (-b) = a \cdot b$.

Beweis. i) Es gilt

$$a \cdot 0 = a \cdot (0+0) \underset{(R3)}{=} a \cdot 0 + a \cdot 0.$$

Subtraktion von $a \cdot 0$ auf beiden Seiten ergibt $0 = a \cdot 0$.

ii) Es gilt

$$ab + a(-b) \underset{(R3)}{=} a \cdot (b + (-b)) = a \cdot 0 \underset{(i)}{=} 0.$$

Also gilt $a(-b) = -ab$. Der Beweis für $(-a)b = -ab$ ist analog.

iii) Es gilt

$$(-a)(-b) \underset{(ii)}{=} -(-a)b \underset{(ii)}{=} -(-ab) = ab.$$

□

Bemerkung 2.1.1. Es gibt Beispiele von nichtkommutativen Ringen, d.h. Ringe, in denen das Kommutativgesetz der Multiplikation nicht gilt. Die Addition ist immer kommutativ, auch in nichtkommutativen Ringen.

Definition 2.1.2. Ein Ring $(K, +, \cdot)$ heißt Körper, wenn $(K - \{0\}, \cdot)$ eine abelsche Gruppe ist. Ist $a \neq 0$, so wird das zu a inverse Element bzgl. der Multiplikation mit a^{-1} bezeichnet.

Beispiel 2.1.1. Die ganzen Zahlen bilden bzgl. der Addition und Multiplikation einen Ring, den Ring $(\mathbb{Z}, +, \cdot)$ Dieser Ring ist jedoch kein Körper, da $(\mathbb{Z} - \{0\}, \cdot)$ keine Gruppe ist. Die einzigen Elemente von $\mathbb{Z}$, die Inverse bzgl. der Multiplikation besitzen, sind 1 und -1.

Beispiel 2.1.2. Die rationalen Zahlen und die reellen Zahlen bilden bzgl. der Addition und Multiplikation einen Körper, die Körper $(\mathbb{Q}, +, \cdot)$ bzw. $(\mathbb{R}, +, \cdot)$. Im nächsten Abschnitt werden wir den Körper $\mathbb{C}$ der komplexen Zahlen kennenlernen.

In der Mathematik und ihren Anwendungen spielen auch noch völlig anders geartete Körper eine Rolle. Der einfachste Körper, den es überhaupt gibt, hat nur zwei Elemente: 0 und 1.

Beispiel 2.1.3. Es sei $\mathbb{F} = \{0, 1\}$. Die Verknüpfungen $+$ der Addition und $\cdot$ der Multiplikation sind durch folgende Tafeln definiert:

$$\begin{array}{c|cc} + & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array} \qquad \text{bzw.} \qquad \begin{array}{c|cc} \cdot & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1. \end{array}$$

2.2 Der Körper der komplexen Zahlen

Definition 2.2.1. Die Menge $\mathbb{C}$ der komplexen Zahlen ist die Menge aller Paare reeller Zahlen:

$$\mathbb{C} = \{(x, y) \colon x, y \in \mathbb{R}\}.$$

Addition und Multiplikation sind wie folgt definiert:

i) $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ für $(x_1, y_1), (x_2, y_2) \in \mathbb{C}$.

ii) $(x_1, y_1) \cdot (x_2, y_2) = (x_1 x_2 - y_1 y_2, x_1 y_2 + y_1 x_2)$.

Diese Regeln werden durch folgende Definition übersichtlich:

Definition 2.2.2. Wir setzen $i := (0,1)$.

Aus Definition 2.2.1 (ii) ergibt sich dann die Regel $i^2 = (-1,0)$. Die Menge $\mathbb{R}$ der reellen Zahlen kann nun durch eine leichte Änderung der Definition 2.2.1 zu einer Teilmenge der komplexen Zahlen gemacht werden.

Definition 2.2.3. Es sei $\tilde{\mathbb{C}} = (\mathbb{C} - \{(x,0) \colon x \in \mathbb{R}\}) \cup \mathbb{R}$.
Wir "werfen also die Elemente $(x,0)$ hinaus" und ersetzen sie durch die reellen Zahlen x. Die Regel $i^2 = (-1,0)$ wird zu $i^2 = -1$, und (x,y) kann als $(x,y) = x+iy$ geschrieben werden. Die Rechenregeln lassen sich wie folgt sehr leicht merken:
Es gelten die üblichen Regeln (Assoziativ-, Kommutativ- und Distributivgesetze), und es ist $i^2 = -1$.

Im folgenden schreiben wir wieder $\mathbb{C}$ statt $\tilde{\mathbb{C}}$.

Beispiel 2.2.1. Es gilt

$$(3+5i)\cdot(7-2i) = 3\cdot 7 - 3\cdot 2i + 7\cdot 5i - (5i)\cdot 2i = 21 - 10i^2 + (7\cdot 5 - 3\cdot 2)i \underset{i^2=-1}{=} 31 + 29i.$$

Satz 2.2.1. *Die Struktur* $(\mathbb{C}, +, \cdot)$ *ist ein Körper mit der Null 0 und der Eins 1.*
Für $x + iy \neq 0$ *haben wir*

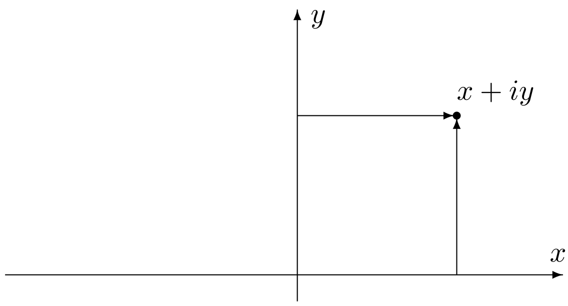
$$(x+iy)^{-1} = \frac{x-iy}{x^2+y^2}.$$

Beweis. Die Körperaxiome folgen durch Nachrechnen aus Definition 2.2.1.
Zum Nachweis der Inversen verwendet man

$$(x+iy)^{-1} = \frac{1}{x+iy}\cdot\frac{x-iy}{x-iy} = \frac{x}{x^2+y^2} - i\frac{y}{x^2+y^2},$$

falls $x + iy \neq 0$. □

So wie der Körper $\mathbb{R}$ der reellen Zahlen durch die Zahlengerade veranschaulicht werden kann, kann der Körper $\mathbb{C}$ der komplexen Zahlen durch die Zahlenebene veranschaulicht werden. Die Menge der Punkte, die der Teilmenge $\mathbb{R}$ entsprechen, wird auch als reelle Achse bezeichnet, die Menge der Punkte, die der Teilmenge $\{iy \colon y \in \mathbb{R}\}$ entsprechen, als imaginäre Achse.



Definition 2.2.4. Es sei $z = x + iy \in \mathbb{C}$ mit $x, y \in \mathbb{R}$. Wir definieren den Realteil ($\Re(z)$) und den Imaginärteil ($\Im(z)$) von z durch $\Re(z) = x$ und $\Im(z) = y$.
Der Betrag von z wird durch

$$|z| = \sqrt{x^2+y^2}$$

definiert.

Bemerkung 2.2.1. Die geometrische Bedeutung des Betrages $|z|$ ist die Entfernung des Punktes z in der komplexen Zahlenebene vom Ursprung.

2.3 Der allgemeine Begriff des Vektorraums

Wir verallgemeinern nun die Definition des reellen Vektorraums von Definition 1.6.1, indem wir an die Stelle des Körpers $\mathbb{R}$ der reellen Zahlen einen beliebigen Körper K treten lassen.

Definition 2.3.1. Es sei V eine Menge mit einer Verknüpfung $\oplus$, weiter $(K, +, \cdot)$ ein Körper mit Null 0 und Eins 1 sowie $\circ\colon K \times V \to V$ eine Abbildung.
Das Tripel $(V, K, \circ)$ heißt ein Vektorraum über (dem Körper) K, wenn die folgenden Axiome für alle $\vec{u}, \vec{v} \in V$ und $\lambda, \mu \in K$ gelten:

(V1) $(V, \oplus)$ bildet eine abelsche Gruppe.

(V2) Es gilt $(\lambda + \mu) \circ \vec{u} = (\lambda \circ \vec{u}) \oplus (\mu \circ \vec{u})$ sowie $\lambda \circ (\vec{u} \oplus \vec{v}) = (\lambda \circ \vec{u}) \oplus (\lambda \circ \vec{v})$. (Distributivgesetze)

(V3) Es ist $\lambda \circ (\mu \circ \vec{u}) = (\lambda \cdot \mu) \circ \vec{u}$. (Assoziativgesetz)

(V4) Schließlich ist $1 \circ \vec{u} = \vec{u}$. (Unitaritätsgesetz)

Die Elemente von V heißen Vektoren, die von K Skalare. Wie in Abschnitt 1.6 werden wir später statt $\oplus$ und $\circ$ einfach $+$ und $\cdot$ schreiben

In den Beweisen sämtlicher Sätze von Abschnitt 1.6 wurde von der Menge $\mathbb{R}$ nur die Körpereigenschaft benutzt. Diese Sätze lassen sich daher sofort auf Vektorräume über beliebigen Körpern verallgemeinern. Ihre Beweise erhält man aus den Beweisen der Sätze in Abschnitt 1.6 einfach dadurch, indem man den Körper $\mathbb{R}$ überall durch den allgemeinen Körper K ersetzt.

Als Verallgemeinerung von Satz 1.6.1 erhalten wir

Satz 2.3.1. *In einem Vektorraum $(V, \oplus, \circ)$ über dem Körper K gilt für $\vec{v}, \vec{w} \in V$ und $\lambda, \mu \in K$:*

i) $\lambda \circ \vec{v} = \vec{0} \Leftrightarrow \lambda = 0$ *oder* $\vec{v} = \vec{0}$

ii) $(-\lambda) \circ \vec{v} = \lambda \circ (\ominus \vec{v}) = \ominus(\lambda \circ \vec{v})$, *wobei* $\ominus$ *die Inversion in* $(V, \oplus)$ *ist.*

iii) $(\lambda - \mu) \circ \vec{v} = \lambda \circ \vec{v} \ominus \mu \circ \vec{v}$ *sowie* $\lambda \circ (\vec{v} \ominus \vec{w}) = \lambda \circ \vec{v} \ominus \lambda \circ \vec{w}$

Als Verallgemeinerung von Definition 1.6.2 geben wir

Definition 2.3.2. Es sei $(V, \oplus, \circ)$ ein Vektorraum über K und $U \subset V$ eine Teilmenge von V. Dann heißt $(U, \oplus, \circ)$ ein Unterraum von V, wenn es ebenfalls ein Vektorraum ist.

Als Verallgemeinerung von Satz 1.6.2 ergibt sich

Satz 2.3.2. *(Unterraumkriterium)*
Es sei V ein Vektorraum über K und $U \subset V$ eine nichtleere Teilmenge von V. Dann ist U genau dann ein Unterraum von V, wenn

$$\vec{u}, \vec{v} \in U \Rightarrow \lambda\vec{u} + \mu\vec{v} \in U$$

für alle $\lambda, \mu \in K$ gilt.

Definition 2.3.3. Es sei V ein Vektorraum über K und $\vec{v}_1, \ldots, \vec{v}_n \in V$. Eine Linearkombination von $\vec{v}_1, \ldots, \vec{v}_n$ ist eine Summe der Form $\lambda_1\vec{v}_1 + \ldots + \lambda_n\vec{v}_n$ mit $\lambda_j \in K$. Die λ_j heißen die Koeffizienten der Linearkombination.

Definition 2.3.4. Es sei V ein Vektorraum über K und $M \subset V$.

i) Unter dem von M erzeugten oder aufgespannten Unterraum von V (auch Erzeugnis von M oder lineare Hülle von M mit der Schreibweise $\langle M \rangle$ versteht man den Durchschnitt aller Unterräume von V, die M enthalten.

ii) Unter $L(M)$ versteht man

$$L(M) = \{\lambda_1 \vec{v}_1 + \ldots + \lambda_n \vec{v}_n \colon n \in \mathbb{N},\ \vec{v}_1, \ldots, \vec{v}_n \in M,\ \lambda_1, \ldots, \lambda_n \in K\}$$

als die Menge aller Linearkombinationen von Elementen von M.

Als Verallgemeinerung von Satz 1.6.5 erhalten wir

Satz 2.3.3. *Es sei V ein Vektorraum über K und $M \subset V$ mit $M \neq \emptyset$. Dann ist $\langle M \rangle$ ein Unterraum von V und $\langle M \rangle = L(M)$.*

Wir schließen mit einigen Beispielen:

Beispiel 2.3.1. Es sei $(K, +, \cdot)$ ein Körper, $n \in \mathbb{N}$ sowie $K^n = \{\vec{x} = (x_1, \ldots, x_n) \colon x_1, \ldots, x_n \in K\}$. Durch die komponentenweise definierte Addition und die Skalarmultiplikation wird K^n zu einem Vektorraum $V = (K^n, K, \circ)$. Für $\vec{x} = (x_1, \ldots, x_n), \vec{y} = (y_1, \ldots, y_n) \in K^n$ und $\lambda \in K$ definiert man

$$\begin{aligned} \vec{x} + \vec{y} &= (x_1 + y_1, \ldots, x_n + y_n) \\ \lambda \vec{x} &= (\lambda x_1, \ldots, \lambda x_n). \end{aligned}$$

Die Vektorraumeigenschaften werden nachgeprüft wie im Spezialfall $K = \mathbb{R}$ (Beispiel 1.6.1).

Definition 2.3.5. Der $(K^n, K, \circ)$ oder kurz K^n heißt der n- dimensionale Standardraum über dem Körper K.

Für Anwendungen von besonderer Bedeutung ist der Fall, dass $\mathbb{F}_q$ ein endlicher Körper mit q Elementen ist.

Satz 2.3.4. *Es sei $\mathbb{F}_q$ ein endlicher Körper mit q Elementen und $n \in \mathbb{N}$. Dann hat der Standardraum $\mathbb{F}_q^n$ genau q^n Elemente.*

Beweis. Für jede der n Komponenten gibt es q Möglichkeiten, womit sich insgesamt q^n ergeben. □

Beispiel 2.3.2. Es sei $\mathbb{F}_q$ ein endlicher Körper mit q Elementen und $n \in \mathbb{N}$.
Ein linearer Code (der Länge n mit Alphabet $\mathbb{F}_q$) ist ein Unterraum von $\mathbb{F}_q^n$.
Die Elemente von C werden Codeworte genannt.

Für Anwendungen sind folgende Begriffe besonders wichtig:

1. Hammingsbstand
2. Gewicht
3. Minimalgewicht
4. Minimalabstand

Diese Begriffe haben folgende Bedeutung:
Unter dem Hammingabstand $d(\vec{x}, \vec{y})$ zweier Elemente $\vec{x} = (x_1, \ldots, x_n), \vec{y} = (y_1, \ldots, y_n) \in \mathbb{F}_q^n$ versteht man die Anzahl der Stellen, an denen sich $\vec{x}$ und $\vec{x}$ unterscheiden, also

$$d(\vec{x}, \vec{y}) = |\{j \in \{1, \ldots, n\} \colon x_j \neq y_j\}.$$

Unter dem Gewicht w von $\vec{x}$ versteht man die Anzahl der von null verschiedenen Stellen.
Beispiel: $q = 2$ und $n = 7$, $\vec{x} = (1,0,0,1,0,1,0)$, $\vec{y} = (1,0,1,1,0,0,0)$. Dann ist $w(\vec{x}) = w(\vec{y}) = 3$.
Unter dem Minimalgewicht w_0 des Codes C versteht man den kleinsten Wert, den die Gewichtsfunktion $w(\vec{x})$ für $\vec{x} \in C - \{\vec{0}\}$ annimmt.
Unter dem Minimalabstand d_0 von C versteht man den kleinsten Wert, den der Hammingabstand $d(\vec{x}, \vec{y})$ für zwei verschiedene Codeworte $\vec{x}, \vec{y} \in C$ annehmen kann.

Man zeigt leicht:

$$d_0 = w_0, \tag{$*$}$$

der Minimalabstand ist gleich dem Minimalgewicht.
Beweis:
Es sei $\vec{x}_0 \in C$ und $w(\vec{x}_0) = w_0$. Wegen $\vec{0} \in C$ (Unterraumeigenschaft von C) ist

$$d_0 \leq d(\vec{x}_0, \vec{0}) = w(\vec{x}_0) = w_0. \tag{1}$$

Es seien $\vec{x}_1, \vec{y}_1 \in C$ mit $d(\vec{x}_1, \vec{y}_1) = d_0$. Weiter ist $\vec{x}_1 - \vec{y}_1 \in C$ und

$$w_0 \leq w(\vec{x}_1 - \vec{y}_1) \leq d(\vec{x}_1, \vec{y}_1) = d_0. \tag{2}$$

Aus (1) und (2) folgt $(*)$.

Beispiel 2.3.3. Der erste lineare Code wurde 1947 von Richard Hamming (1915- 1998) aufgestellt: der $(7,4)$- Hamming- Code.
Es ist $C \subset \mathbb{F}_2^7$ und definiert als die Lösungsmenge des folgenden LGS über $\mathbb{F}_2$:

$$\begin{array}{rcl}
x_2 + x_3 + x_4 + x_5 & = & 0 \\
x_1 + x_3 + x_4 + x_6 & = & 0 \\
x_1 + x_2 + x_4 + x_7 & = & 0
\end{array}$$

oder $A \cdot \vec{x} = \vec{0}$ mit

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

und dem Spaltenvektor

$$\begin{pmatrix} x_1 \\ \vdots \\ x_7 \end{pmatrix}$$

Das LGS befindet sich schon in der im Gaußschen Algroithmus angestrebten Endform, die in Abschnitt 1.4 beschrieben wurde. Allerdings stehen hier die Variablen, nach denen aufgelöst wird, rechts. Offenbar können die Werte für x_1, x_2, x_3, x_4 beliebig vorgegeben werden (Informationsbits), während die Werte von x_5, x_6, x_7 dadaurch bestimmt sind (Prüfbits).
Von den $2^7 = 128$ Zeichenfolgen aus $\mathbb{F}_2^7$ sind somit $2^4 = 16$ Codeworte.

Eines der Probleme der Codierungstheorie ist, Fehler bei der Übertragung zu erkennen. Fehler können nur dann nicht erkannt werden, wenn das beabsichtigte Codewort durch Fehler in ein anderes Codewort umgewandelt wird. So kann etwa das Codewort $(0,0,0,0,0,0,0)$ durch drei Fehler in das Codewort $(1,0,0,0,0,1,1)$ übergehen.
Ein Codewort kann nur dann in ein anderes Codewort übergehen, wenn die Anzahl der unbekannten Fehler mindestens gleich dem Minimalabstand d_0 von C ist. Wir behaupten, dass für den $(7,4)$-Hamming- Code der Minimalabstand $d_0 = 3$ ist. Damit können Fehler bei der Übertragung erkannt werden, wenn ihre Anzahl höchstens zwei ist.
Nach Beispiel 2.3.2 ist d_0 gleich dem Minimalgewicht w_0.
Es sei $\vec{e}_j$ die j- te Spalte in $\mathcal{A}$.

1. Es gibt kein $\vec{x} \in \mathcal{C}$ mit $w(\vec{x}) = 1$.
   Aus $\vec{x} = (x_1, \ldots, x_j, \ldots, x_7)$ mit $x_j = 1$ und $x_k = 0$ für $k \neq j$ folgt $\vec{x} = \vec{e}_j \neq \vec{0}$, also $\vec{x} \notin \mathcal{C}$.
2. Es gibt kein $\vec{x} \in \mathcal{C}$ mit $w(\vec{x}) = 2$.
   Es sei $\vec{x} = (x_1, \ldots, x_7)$ mit $x_i = x_j = 1$ und $x_k = 0$ für $k \notin \{i, j\}$. Dann ist $\mathcal{A}\vec{x} = \vec{e}_i + \vec{e}_j \neq \vec{0}$, da $-\vec{e}_j = \vec{e}_i$ ist, aber andererseits alle Spalten verschieden sind.
3. Es sei $\vec{x}_0 = (1,0,0,0,0,1,1)$. Also ist $w(\vec{x}_0) = 3$, und damit ist $w_0 = d_0 = 3$.

2.4 Lineare Abhängigkeit, Basis, Dimension

Wir verallgemeinern hier Definitionen und Sätze von Abschnitt 1.7 auf Vektorräume über beliebigen Körpern und klären auch noch Fragen, die in Abschnitt 1.7 offen geblieben waren:

- die Existenz einer Basis und
- der Zusammenhang mit der Dimension des Vektorraums.

Es sei V stets ein Vektorraum über dem Körper K.

Definition 2.4.1. Vektoren $\vec{v}_1, \ldots, \vec{v}_k \in V$ heißen linear abhängig (kurz l.a.), wenn es $\lambda_1, \ldots, \lambda_k \in K$ gibt, so dass

$$\lambda_1\vec{v}_1 + \ldots + \lambda_k\vec{v}_k = \vec{0} \quad \text{und} \quad (\lambda_1, \ldots, \lambda_k) \neq (0, \ldots, 0)$$

gilt. Andernfalls, d.h. wenn für $\lambda_j \in K$ stets

$$\lambda_1\vec{v}_1 + \ldots + \lambda_k\vec{v}_k = \vec{0} \Rightarrow \lambda_1 = \lambda_2 = \ldots = \lambda_k = 0$$

gilt, heißen $\vec{v}_1, \ldots, \vec{v}_k$ linear unabhängig (kurz l.u.).

Satz 2.4.1. *Es gilt:*

i) Der Nullvektor $\vec{0}$ ist stets l.a., ein einzelner Vektor $\vec{v} \neq \vec{0}$ ist stets l.u.

ii) Mit $\vec{v}_1, \ldots, \vec{v}_k$ sind auch $\vec{v}_1, \ldots, \vec{v}_k, \ldots, \vec{v}_l$ mit $l \geq k$ l.a.

iii) Sind $\vec{v}_1, \ldots, \vec{v}_k$ l.u., so auch $\vec{v}_1, \ldots, \vec{v}_m$ für $1 \leq m \leq k$.

iv) Ist $\vec{v}$ eine Linearkombination von $\vec{v}_1, \ldots, \vec{v}_k$, so sind $\vec{v}_1, \ldots, \vec{v}_k, \vec{v}$ l.a.

v) Sind $k \geq 2$ Vektoren l.a., so ist wenigstens einer von ihnen eine Linearkombination der anderen.

vi) Sind $\vec{v}_1, \ldots, \vec{v}_k$ l.u., aber $\vec{v}_1, \ldots, \vec{v}_k, \vec{v}$ l.a., so ist $\vec{v}$ eine Linearkombination der $\vec{v}_1, \ldots, \vec{v}_k$.

Beweis. i) Es gilt $1 \cdot \vec{0} = \vec{0}$, also ist $\vec{0}$ l.a., andererseits folgt aus $\lambda\vec{v} = \vec{0}$ mit $\vec{v} \neq \vec{0}$ dann $\lambda = 0$ wegen Satz 2.3.1(i).

ii) Aus $\lambda_1\vec{v}_1 + \ldots + \lambda_k\vec{v}_k = \vec{0}$ und $(\lambda_1, \ldots, \lambda_k) \neq (0, \ldots, 0)$ folgt auch

$$\lambda_1\vec{v}_1 + \ldots + \lambda_k\vec{v}_k + \lambda_{k+1}\vec{v}_{k+1} + \ldots + \lambda_l\vec{v}_l = \vec{0}$$

und $(\lambda_1, \ldots, \lambda_l) \neq (0, \ldots, 0)$, wenn man $\lambda_{k+1} = \ldots = \lambda_l = 0$ einsetzt.

iii) Wären $\vec{v}_1, \ldots, \vec{v}_m$ l.a., so nach (ii) auch $\vec{v}_1, \ldots, \vec{v}_k$.

iv) Ist $\vec{v} = \lambda_1\vec{v}_1 + \ldots + \lambda_k\vec{v}_k$, so folgt $\lambda_1\vec{v}_1 + \ldots + \lambda_k\vec{v}_k + (-1)\vec{v} = \vec{0}$.

v) Es sei $\lambda_1\vec{v}_1 + \ldots + \lambda_k\vec{v}_k = \vec{0}$ und $(\lambda_1, \ldots, \lambda_k) \neq (0, \ldots, 0)$. Nach eventueller Umnummerierung können wir annehmen, dass $\lambda_k \neq 0$ ist. Dann folgt

$$\vec{v}_k = (-\lambda_k^{-1})\lambda_1\vec{v}_1 + \ldots + (-\lambda_k)^{-1}\lambda_{k-1}\vec{v}_{k-1}.$$

vi) Es gelte $\lambda_1\vec{v}_1 + \ldots + \lambda_k\vec{v}_k + \lambda\vec{v} = \vec{0}$ mit $(\lambda_1, \ldots, \lambda_k, \lambda) \neq (0, \ldots, 0)$. Dann muss $\lambda \neq 0$ gelten, da wegen der linearen Unabhängigkeit von $\vec{v}_1, \ldots, \vec{v}_k$ die Gleichung $\lambda_1\vec{v}_1 + \ldots + \lambda\vec{v}_k = \vec{0}$ nur für $\lambda_1 = \ldots = \lambda_k = 0$ möglich ist. Dann ist aber $\vec{v} = (-\lambda^{-1})\lambda_1\vec{v}_1 + \ldots + (-\lambda^{-1})\lambda_k\vec{v}_k$.

$\square$

Der Begriff der linearen (Un-)Abhängigkeit lässt sich auf beliebige (auch unendliche) Mengen von Vektoren erweitern:

Definition 2.4.2. Eine Menge $M \subset V$ heißt l.u., wenn je endlich viele verschiedene (!) Vektoren aus M l.u. sind, andernfalls l.a.

Bemerkung 2.4.1. Wir stellen einige Spezialfälle zusammen:

1. Die leere Menge ist l.u.
2. Ist $\vec{0} \in M$, so ist M l.a.
3. Jede Teilmenge einer l.u. Menge ist l.u
4. Jede in V liegende Obermenge einer l.a. Menge ist l.a.
5. Achtung: Im Falle $\vec{a}_1 = \vec{a}_2 \neq \vec{0}$ sind $\vec{a}_1$ und $\vec{a}_2$ l.a., aber $\{\vec{a}_1, \vec{a}_2\}$ ist l.u. (weil es tatsächlich die Menge $\{\vec{a}_1\}$ ist).

Definition 2.4.3. Ist $M \subset V$ l.u., so schreiben wir $\langle M \rangle = \langle\langle M \rangle\rangle$, und nennen M eine Basis des Erzeugnisses $\langle M \rangle$, kurz:

$$V = \langle\langle M \rangle\rangle \Leftrightarrow M \text{ ist l.u. und } \langle M \rangle = V.$$

Speziell ist eine Basis eines Vektorraums V also ein linear unabhängiges Erzeugendensystem von V. Im Falle $M = \{\vec{v}_1, \ldots, \vec{v}_k\}$ sagen wir auch: "Die Vektoren $\vec{v}_1, \ldots, \vec{v}_k$ bilden eine Basis von V", und schreiben wieder kurz $\langle\langle \vec{v}_1, \ldots, \vec{v}_k \rangle\rangle$ statt $\langle\langle \{\vec{v}_1, \ldots, \vec{v}_k\} \rangle\rangle$.

Beispiel 2.4.1. Es ist $\{\vec{0}\} = \langle\langle \emptyset \rangle\rangle$, und $\emptyset$ ist die einzige Basis von $\{\vec{0}\}$.

Beispiel 2.4.2. Der K^n hat die Standardbasis $\{\vec{e}_1, \ldots, \vec{e}_n\}$ mit den Einheitsvektoren $\vec{e}_1 = (1, 0, \ldots, 0)$, $\vec{e}_2 = (0, 1, 0, \ldots, 0) \ldots, \vec{e}_n = (0, \ldots, 0, 1)$.
Der Beweis erfolgt wie im Spezialfall $K = \mathbb{R}$ in Beispiel 1.7.1.

Beispiel 2.4.3. Der Körper $\mathbb{C}$ der komplexen Zahlen, aufgefaßt als Vektorraum über $K = \mathbb{R}$, hat die Basis $\{1, i\}$.

Man kann zeigen, dass jeder Vektorraum eine Basis besitzt. Wir werden uns jedoch bei der Diskussion der Basis auf relativ einfache Fälle, sogenannte <u>endlichdimensionale</u> Vektorräume beschränken.

Definition 2.4.4. Gibt es eine maximale Zahl n von l.u. Vektoren in V, so heißt n die <u>Dimension</u> von V, geschrieben $\dim V$:

$$\dim V = \max\{|M|\colon M \subset V \text{ l.u.}\} \in \mathbb{N}_0.$$

Gibt es kein solches n (existiert also zu jedem $k \in \mathbb{N}$ eine l.u.-Teilmenge $M \subset V$ mit $|M| = k$), so heißt V <u>unendlichdimensional</u>, und wir schreiben $\dim V = \infty$.

Beispiel 2.4.4. Es ist $\dim\{\vec{0}\} = 0$.

Beispiel 2.4.5. Es sei $V = K$ ein Körper aufgefasst als Vektorraum über sich selbst. Die Menge $\{1\}$ ist l.u., und sind $\lambda_1, \lambda_2 \in K$ mit $\lambda_1 \neq \lambda_2$, so gilt $\lambda_2\lambda_1 + (-\lambda_1)\lambda_2 = 0$, d.h. die Vektoren λ_1, λ_2 sind stets l.a., also $\dim K = 1$.

Beispiel 2.4.6. Es sei $\mathcal{F} = \{f\colon \mathbb{R} \to \mathbb{R}\}$ die Menge aller Funktionen von $\mathbb{R}$ nach $\mathbb{R}$.
Für $f, g \in \mathcal{F}$ und $\lambda \in \mathbb{R}$ definieren wir die Summe $f + g\colon \mathbb{R} \to \mathbb{R}$, $x \to f(x) + g(x)$ und die Skalarmultiplikation $\lambda f\colon \mathbb{R} \to \mathbb{R}$, $x \to \lambda f(x)$.
Man sieht leicht, dass $\mathcal{F}$ dadurch zu einem reellen Vektorraum wird. Wir betrachten die Menge $M := \{f_\nu\colon \nu \in \mathbb{Z}\}$ mit

$$f_\nu(x) := \begin{cases} 1, & \text{falls} \quad \nu \leq x < \nu + 1 \\ 0 & \text{sonst.} \end{cases}$$

Dann ist M l.u.
Beweis:
Sei $\lambda_1 f_{\nu_1} + \ldots + \lambda_k f_{\nu_k} = 0$, wobei 0 die Nullfunktion darstellt, mit paarweise verschiedenen $\nu_1, \ldots, \nu_k$. Für x mit $\nu_i \leq x < \nu_i + 1$ erhalten wir

$$0 = (\lambda_1 f_{\nu_1} + \ldots + \lambda_k f_{\nu_k})(x) = \lambda_i f_{\nu_i}(x) = \lambda_i,$$

also $\lambda_1 = \ldots = \lambda_k = 0$.
Daraus folgt, dass der Vektorraum $\mathcal{F} = \{f\colon \mathbb{R} \to \mathbb{R}\}$ die Dimension $\dim \mathcal{F} = \infty$ besitzt.

Satz 2.4.2. *Es sei $\dim V = n < \infty$. Dann besitzt V eine Basis, genauer bildet jede l.u. Teilmenge mit n Vektoren eine Basis von V.*

Beweis. Seien $\vec{v}_1, \ldots, \vec{v}_n \in V$ l.u. und $\vec{v} \in V$ beliebig. Nach Definition der Dimension sind $\vec{v}_1, \ldots, \vec{v}_n, \vec{v}$ l.a. Nach Satz 2.4.1 (vi) ist $\vec{v}$ eine Linearkombination von $\vec{v}_1, \ldots, \vec{v}_n$, also $\vec{v} \in \langle\, \vec{v}_1, \ldots, \vec{v}_n \,\rangle$. Da $\vec{v} \in V$ beliebig war, folgt $V = \langle\, \vec{v}_1, \ldots, \vec{v}_n \,\rangle = \langle\langle\, \vec{v}_1, \ldots, \vec{v}_n \,\rangle\rangle$. □

Wir werden in kürze zeigen, dass es keine Basis von V mit weniger als $\dim V$ Elementen gibt.

Satz 2.4.3. *Für $\vec{v}_1, \dots \vec{v}_n \in V$ sind äquivalent:*

i) $V = \langle\langle\, \vec{v}_1, \dots, \vec{v}_n \,\rangle\rangle$.

ii) Jedes $\vec{v}$ besitzt eine Darstellung

$$\vec{v} = \lambda_1 \vec{v}_1 + \ldots + \lambda_n \vec{v}_n \qquad (*)$$

mit eindeutig bestimmten Koeffizienten $(\lambda_1, \dots, \lambda_n) \in K^n$, oder: die durch $(*)$ vermittelte Abbildung

$$K^n, \to V, \quad (\lambda_1, \dots, \lambda_n) \to \lambda_1 \vec{v}_1 + \ldots + \lambda_n \vec{v}_n$$

ist bijektiv.

Beweis. "$\Rightarrow$:"
Da $\{\vec{v}_1, \dots, \vec{v}_n\}$ ein Erzeugendensystem von V bildet, hat jedes $\vec{v} \in V$ mindestens eine Darstellung der Form $(*)$. Es seien $\vec{v} = \lambda_1 \vec{v}_1 + \ldots + \lambda_n \vec{v}_n$ und $\vec{v} = \mu_1 \vec{v}_1 + \ldots + \mu_n \vec{v}_n$ zwei Darstellungen des gleichen Vektors. Dann folgt aus der linearen Unabhängigkeit der $\vec{v}_j$

$$(\lambda_1 - \mu_1)\vec{v}_1 + \ldots + (\lambda_n - \mu_n)\vec{v}_n = \vec{0} \Rightarrow \lambda_j - \mu_j = 0$$

für $j = 1, \dots, n$, also $\lambda_j = \mu_j$, d.h. die Darstellung $(*)$ ist eindeutig.
"$\Leftarrow$:"
Eine (also die einzige) Darstellung $(*)$ von $\vec{0}$ ist $\vec{0} = 0 \cdot \vec{v}_1 + \ldots + 0 \cdot \vec{v}_n$. Also sind $\vec{v}_1, \dots, \vec{v}_n$ l.u. □

Satz 2.4.4. *Es sei $V = \langle\langle\, \vec{v}_1, \dots, \vec{v}_n \,\rangle\rangle$ und $\vec{w} \in V$ mit $\vec{w} \neq \vec{0}$. In der Darstellung*

$$\vec{w} = \lambda_1 \vec{v}_1 + \ldots + \lambda_n \vec{v}_n \qquad (*)$$

mit $(\lambda_1, \dots, \lambda_n) \in K^n$ sei j ein Index mit $\lambda_j \neq 0$. Dann ist auch

$$V = \langle\langle\, \vec{v}_1, \dots, \vec{v}_{j-1}, \vec{w}, \vec{v}_{j+1}, \dots, \vec{v}_n \,\rangle\rangle \ .$$

Beweis. Es sei O.B.d.A. $j = 1$ (sonst nummerieren wir die $\vec{v}_j$ um), und wir müssen $V = \langle\langle\, \vec{w}, \vec{v}_2, \dots, \vec{v}_n \,\rangle\rangle$ zeigen.

i) $V = \langle\, \vec{w}, \vec{v}_2, \dots, \vec{v}_n \,\rangle$:
Auflösen von $(*)$ nach $\vec{v}_1$ ergibt

$$\vec{v}_1 = \mu_1 \vec{w} + \mu_2 \vec{v}_2 + \ldots + \mu_n \vec{v}_n \qquad (**)$$

mit $\mu_i = -\lambda_i \lambda_1^{-1}$ für $i = 2, \dots, n$.
Sei nun $\vec{v} \in V$ beliebig, etwa $\vec{v} = \alpha_1 \vec{v}_1 + \ldots + \alpha_n \vec{v}_n$ mit $\alpha_1, \dots, \alpha_n \in K$. Einsetzen von $(**)$ ergibt

$$\begin{aligned} \vec{v} &= \alpha_1(\mu_1 \vec{w} + \mu_2 \vec{v}_2 + \ldots + \mu_n \vec{v}_n) + \alpha_2 \vec{v}_2 + \ldots + \alpha_n \vec{v}_n \\ &= \alpha_1 \mu_1 \vec{w} + (\alpha_1 \mu_2 + \alpha_2)\vec{v}_2 + \ldots + (\alpha_1 \mu_n + \alpha_n)\vec{v}_n \in \langle\, \vec{w}, \vec{v}_2, \dots, \vec{v}_n \,\rangle \,. \end{aligned}$$

ii) $\{\vec{w}, \vec{v}_2, \dots, \vec{v}_n\}$ ist l.u.:
Angenommen wir haben

$$\mu_1 \vec{w} + \mu_2 \vec{v}_2 + \ldots + \mu_n \vec{v}_n = \vec{0}.$$

Dann folgt mit $(*)$:

$$\begin{aligned}\vec{0} &= \mu_1(\lambda_1\vec{v}_1 + \ldots + \lambda_n\vec{v}_n) + \mu_2\vec{v}_2 + \ldots + \mu_n\vec{v}_n \\ &= (\mu_1\lambda_1)\vec{v}_1 + (\mu_1\lambda_2 + \mu_2)\vec{v}_2 + \ldots + (\mu_1\lambda_n + \mu_n)\vec{v}_n\end{aligned}$$

und damit

$$\mu_1\lambda_1 = \mu_1\lambda_2 + \mu_2 = \ldots = \mu_1\lambda_n + \mu_n = 0,$$

da $\vec{v}_1, \ldots, \vec{v}_n$ l.u. sind, also ist $\mu_1 = 0$ wegen $\lambda_1 \neq 0$. Daraus folgt dann $\mu_2 = \ldots = \mu_n = 0$, also ist $\{\vec{w}, \vec{v}_2, \ldots, \vec{v}_n\}$ l.u.

□

Wir verallgemeinern Satz 2.4.4 zum

Satz 2.4.5. *(Austauschsatz von Steinitz, Ergänzungssatz)*
Es sei V ein Vektorraum und $V = \langle\langle\, \vec{v}_1, \ldots, \vec{v}_n \,\rangle\rangle$ und $\vec{w}_1, \ldots, \vec{w}_k \in V$ l.u., womit $k \leq n$ ist. Ferner lässt sich aus $\{\vec{v}_1, \ldots, \vec{v}_n\}$ eine Teilmenge $\{\vec{v}_{k+1}, \ldots, \vec{v}_n\}$ so auswählen, dass

$$V = \langle\langle\, \vec{w}_1, \ldots, \vec{w}_k, \vec{v}_{k+1}, \ldots, \vec{v}_n \,\rangle\rangle$$

ist.
Mit anderen Worten: Man kann k geeignet gewählte Vektoren der Basis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ gegen die $\vec{w}_i$ austauschen, so dass man wieder eine Basis von V erhält.
Insbesondere lässt sich jede l.u. Teilmenge eines endlichdimensionalen Vektorraums zu einer Basis von V ergänzen.

Beweis. Wir führen eine vollständige Induktion nach k für festes n.
Induktionsanfang $k = 1$:
Es ist $n \geq 1$. Wegen $\vec{w}_1 \neq \vec{0}$ ist in der Basisdarstellung $\vec{w}_1 = \lambda_1\vec{v}_1 + \ldots + \lambda_n\vec{v}_n$ mindestens ein $\lambda_j \neq 0$ enthalten. Nach Satz 2.4.4 kann man $\vec{v}_j$ gegen $\vec{w}_1$ austauschen, so dass $\{\vec{v}_1, \ldots, \vec{v}_{j-1}, \vec{w}_1, \vec{v}_{j+1}, \ldots, \vec{v}_n\}$ wieder eine Basis von V ist.
Induktionsschritt $k \to k+1$:
Die Behauptung des Satzes gelte für ein $k \in \mathbb{N}$, und es seien $\vec{w}_1, \ldots, \vec{w}_k, \vec{w}_{k+1} \in V$ linear unabhängig. Nach Induktionsannahme (angewandt auf $\vec{w}_1, \ldots, \vec{w}_k$) ist $k \leq n$, und es gibt eine Teilmenge $\{\vec{v}_{k+1}, \ldots, \vec{v}_n\} \subseteq \{\vec{v}_1, \ldots, \vec{v}_n\}$ derart, dass

$$V = \langle\langle\, \vec{w}_1, \ldots, \vec{w}_k, \vec{v}_{k+1}, \ldots, \vec{v}_n \,\rangle\rangle\,. \qquad (*)$$

Wäre $k = n$, so wäre schon $V = \langle\langle\, \vec{w}_1, \ldots, \vec{w}_k \,\rangle\rangle$, also $\vec{w}_{k+1}$ eine Linearkombination von $\vec{w}_1, \ldots, \vec{w}_k$. Da $\vec{w}_1, \ldots, \vec{w}_k, \vec{w}_{k+1}$ aber l.u. sein sollen, muss folglich $k < n$, d.h. $k + 1 \leq n$ sein. Wegen $(*)$ und $\vec{w}_{k+1} \neq 0$ gilt

$$\vec{w}_{k+1} = \mu_1\vec{w}_1 + \ldots + \mu_k\vec{w}_k + \mu_{k+1}\vec{v}_{k+1} + \cdots + \mu_n\vec{v}_n$$

mit $\mu_1, \ldots, \mu_n \in K$, wobei mindestens ein $\mu_i \neq 0$ ist. Dabei kann nicht $\mu_{k+1} = \ldots = \mu_n = 0$ sein, sonst wären $\vec{w}_1, \ldots, \vec{w}_k, \vec{w}_{k+1}$ l.a., also ist $\mu_i \neq 0$ für mindestens ein $i \in \{k+1, \ldots, n\}$. Austauschen von $\vec{v}_i'$ gegen $\vec{w}_{k+1}$ gemäß Satz 2.4.4 ergibt die Behauptung für $k + 1$. □

Satz 2.4.6. *Es sei $V = \langle\langle\, \vec{v}_1, \ldots, \vec{v}_n \,\rangle\rangle$. Dann ist* $\dim V = n$, *und eine Teilmenge $\mathcal{B} \subseteq V$ ist genau dann eine Basis von V, wenn $\mathcal{B}$ aus n l.u. Vektoren besteht.*

Beweis. Für jede l.u. Menge $M \subseteq V$ gilt nach Satz 2.4.5, dass $|M| \leq n$ ist, also $\dim V \leq n$. Nach Definition der Dimension ist andererseits $n \leq \dim V$. Also folgt insgesamt $\dim V = n$.
Nun sei $\mathcal{B}$ irgendeine Basis von V. Dann folgt zunächst $m := |\mathcal{B}| \leq \dim V = n$ und dann wie oben $\dim V = m$, also $m = n$. Umgekehrt ist nach Satz 2.4.2 auch jede l.u. Menge $\mathcal{B} \subseteq V$ mit $|\mathcal{B}| = \dim \mathcal{B}$ eine Basis von V. □

Beispiel 2.4.7. Für die Standardvektorräume gilt $\dim K^n = n$, denn die Standardbasis $\{\vec{e}_1, \dots, \vec{e}_n\}$ hat n Elemente. Insbesondere ist $\dim \mathbb{R}^n = n$ und $\dim \mathbb{C}^n = n$ (als Vektorraum über $\mathbb{C}$), aber $\mathbb{C}^n$ als Vektorraum über $\mathbb{R}$ besitzt die Dimension $2n$, eine Basis ist $\{\vec{e}_1, \dots, \vec{e}_n, i\vec{e}_1, \dots, i\vec{e}_n\}$.

Satz 2.4.7. *Es sei* $\dim V < \infty$ *und* U *ein Unterraum von* V*. Dann gilt:*

i) $\dim U \leq \dim V$.

ii) $\dim U = \dim V \Leftrightarrow U = V$.

Beweis. i) Dies folgt sofort aus der Definition der Dimension.

ii) Es sei $\dim U = \dim V = n$. Dann besitzt U nach Satz 2.4.2 eine Basis $\{\vec{v}_1, \dots, \vec{v}_n\}$. Diese bildet dann ebenfalls nach Satz 2.4.2 eine Basis von U, also $U = V$.

□

Satz 2.4.8. *(Dimensionssatz für Summenräume)*
Es seien U_1, U_2 *endlichdimensionale Unterräume eines Vektorraums* V*. Dann gilt*

$$\dim U_1 + \dim U_2 = \dim(U_1 + U_2) + \dim(U_1 \cap U_2).$$

Beweis. Da $U_1 \cap U_2$ ein Unterraum von U_1 (und ebenso von U_2) ist, gilt nach Satz 2.4.7 dann $d = \dim(U_1 \cap U_2) < \infty$. Es sei also nach Satz 2.4.2 $\{\vec{a}_1, \dots, \vec{a}_d\}$ eine Basis von $U_1 \cap U_2$ ($= \emptyset$ falls $d = 0$ ist). Wir ergänzen diese Basis nach Satz 2.4.5 zu je einer Basis von U_1 und U_2:

$$\begin{aligned} \mathcal{B}_1 &:= \{\vec{a}_1, \dots, \vec{a}_d, \vec{b}_1, \dots, \vec{b}_r\}, \quad \text{Basis von } U_1, \\ \mathcal{B}_2 &:= \{\vec{a}_1, \dots, \vec{a}_d, \vec{c}_1, \dots, \vec{c}_s\}, \quad \text{Basis von } U_2. \end{aligned}$$

Behauptung: $\mathcal{B} := \mathcal{B}_1 \cup \mathcal{B}_2 = \{\vec{a}_1, \dots, \vec{a}_d, \vec{b}_1, \dots, \vec{b}_r, \vec{c}_1, \dots, \vec{c}_s\}$ ist eine Basis von $U_1 + U_2$. Wir haben zwei Aussagen zu zeigen: $\langle \mathcal{B} \rangle = U_1 + U_2$ und $\mathcal{B}$ l.u.:

i) $\langle \mathcal{B} \rangle = U_1 + U_2$:
Wegen $\mathcal{B} \subseteq U_1 \cup U_2 \subseteq U_1 + U_2$ ($U_1 + U_2$ ist Unterraum) folgt $\langle \mathcal{B} \rangle \subseteq U_1 + U_2$. Andererseits ist

$$U_1 + U_2 = \langle \mathcal{B}_1 \rangle + \langle \mathcal{B}_2 \rangle \subseteq \langle \mathcal{B} \rangle + \langle \mathcal{B} \rangle = \langle \mathcal{B} \rangle$$

da $\mathcal{B}_1, \mathcal{B}_2 \subseteq \mathcal{B}$.

ii) $\mathcal{B}$ ist l.u.:
Es sei

$$\alpha_1 \vec{a}_1 + \ldots + \alpha_d \vec{a}_d + \beta_1 \vec{b}_1 + \ldots \beta_r \vec{b}_r + \gamma_1 \vec{c}_1 + \ldots + \gamma_s \vec{c}_s = \vec{0}$$

mit $\alpha_i, \beta_j, \gamma_k \in K$, und wir müssen nun zeigen, dass alle Koeffizienten verschwinden. Sortieren ergibt

$$\underbrace{\alpha_1 \vec{a}_1 + \ldots + \alpha_d \vec{a}_d + \beta_1 \vec{b}_1 + \ldots + \beta_r \vec{b}_r}_{\in U_1} = -\underbrace{\gamma_1 \vec{c}_1 + \ldots + \gamma_s \vec{c}_s}_{\in U_2} \subseteq U_1 \cap U_2.$$

Dann gibt es Koeffizienten λ_l mit $-(\gamma_1\vec{c}_1 + \ldots + \gamma_s\vec{c}_s) = \lambda_1\vec{a}_1 + \ldots + \lambda_d\vec{a}_d$, also

$$\lambda_1\vec{a}_1 + \ldots + \lambda_d\vec{a}_d + \gamma_1\vec{c}_1 + \ldots + \gamma_s\vec{c}_s = \vec{0}.$$

Da $\mathcal{B}_2$ als Basis l.u. ist folgt $\lambda_1 = \ldots = \lambda_d = \gamma_1 = \ldots = \gamma_s = 0$. Daraus folgt

$$\alpha_1\vec{a}_1 + \ldots + \alpha_d\vec{a}_d + \beta_1\vec{b}_1 + \ldots + \beta_r\vec{b}_r = \vec{0}$$

und somit $\alpha_1 = \ldots = \alpha_d = \beta_1 = \ldots = \beta_r = 0$ da auch $\mathcal{B}_1$ l.u. ist. Also verschwinden sämtliche Koeffizienten $\alpha_i, \beta_j, \gamma_k$, womit $\mathcal{B}$ l.u. ist.

Da wir jetzt Basen für alle beteiligten Räume haben, können wir die Aussage des Satzes durch Zählen der Basisvektoren zeigen:

$$\dim(U_1 + U_2) = |\mathcal{B}| = d + r + s = (d + r) + (d + s) - d = \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2)\,.$$

$\square$

Beispiel 2.4.8. Es sei $V = \mathbb{R}^3$ und $U_1 = \langle\langle\,(1,0,0),(0,1,0)\,\rangle\rangle$ sowie $U_2 = \langle\langle\,(1,-1,0),(0,0,1)\,\rangle\rangle$, also $\dim U_1 = \dim U_2 = 2$. Die Unterräume U_1,U_2 sind Ebenen durch $\vec{0}$, und zwar explizit

$$\begin{aligned} U_1 &= \{(x,y,0)\colon x,y \in \mathbb{R}\} \quad \text{die } xy\text{-Ebene} \\ U_2 &= \{(x,-x,z)\colon x,z \in \mathbb{R}\} \\ U_1 \cap U_2 &= \{(x,-x,0)\colon x \in \mathbb{R}\} = \langle\langle\,(1,-1,0)\,\rangle\rangle \quad \text{die Gerade durch } y = -x,\ z = 0. \end{aligned}$$

Insbesondere ist $\dim(U_1 \cap U_2) = 1$. Mit dem Dimensionssatz folgt: $\dim(U_1 + U_2) = 2 + 2 - 1 = 3$, also $U_1 + U_2 = \mathbb{R}^3$, d.h. alle $\vec{v} \in \mathbb{R}^3$ lassen sich als $\vec{v} = \vec{u}_1 + \vec{u}_2$ mit $\vec{u}_1 \in U_1$ und $\vec{u}_2 \in U_2$ schreiben. Diese Darstellung ist jedoch nicht eindeutig.

Kapitel 3

Lineare Abbildungen und Matrizen

3.1 Lineare Abbildungen

Es seien stets $V, W, V', \ldots$ Vektorräume über dem selben Körper K. Dieser Abschnitt befasst sich mit Abbildungen zwischen Vektorräumen, die mit den linearen Operationen verträglich sind:

Definition 3.1.1. Eine Abbildung $\varphi \colon V \to V'$ heißt linear oder ein (Vektorraum-) Homomorphismus, falls folgende Eigenschaften gelten:

(L1) $\varphi(\vec{v} + \vec{w}) = \varphi(\vec{v}) + \varphi(\vec{w})$ für alle $\vec{v}, \vec{w} \in V$,

(L2) $\varphi(\lambda\vec{v}) = \lambda\varphi(\vec{v})$ für alle $\lambda \in K$ und $\vec{v} \in V$,

bzw. was dazu äquivalent ist:

(L) $\varphi(\lambda\vec{v} + \mu\vec{w}) = \lambda\varphi(\vec{v}) + \mu\varphi(\vec{w})$ für alle $\lambda, \mu \in K$ und $\vec{v}, \vec{w} \in V$.

Die Menge aller linearen Abbildungen $\varphi \colon V \to V'$ wird mit $L(V, V')$ bezeichnet.
Ein $\varphi \in L(V, V')$ heißt ein (Vektorraum-) Isomorphismus, wenn φ bijektiv ist. Existiert ein Isomorphismus $\varphi \colon V \to V'$, so heißt V isomorph zu V', geschrieben $V \cong V'$. Ein $\varphi \in L(V, V)$ (also $V = V'$) heißt Endomorphismus, bzw. im Falle der Bijektivität Automorphismus von V.

Beispiel 3.1.1. Es gibt stets den trivialen Homomorphismus: $\varphi_0(\vec{v}) = \vec{0}' \in V'$ für alle $\vec{v} \in V$.

Beispiel 3.1.2. Der Endomorphismus $\varphi \colon V \to V$, $\varphi(\vec{v}) = \lambda\vec{v}$ für festes $\lambda \in K$ ist trivial für $\lambda = 0$, und ein Automorphismus für $\lambda \neq 0$. Er ist wegen $\varphi(\vec{v}) = \varphi(\vec{w}) \Leftrightarrow \vec{v} = \vec{w}$ injektiv und wegen $\varphi(\lambda^{-1}\vec{v}) = \vec{v}$ surjektiv.

Beispiel 3.1.3. Für $V = \mathbb{R}^2$ ist $\varphi((x, y)) = (\lambda x, \mu y)$ mit $\lambda, \mu \in \mathbb{R} - \{0\}$ die sogenannte Eulerabbildung. Die Ebene wird in x-Richtung um den Faktor λ und in y-Richtung um den Faktor μ gestreckt. So ist φ offenbar ein Automorphismus.

Beispiel 3.1.4. Es sei $V = \mathbb{R}^2$. Die Projektion auf die x-Achse $\varphi \colon V \to V$, $\varphi((x, y)) = (x, 0)$ ist weder injektiv noch surjektiv, also kein Automorphismus.

Beispiel 3.1.5. Für $V = \mathbb{R}^2$ heißt der Automorphismus $\varphi \colon V \to V$, $\varphi((x, y)) = (x + \lambda y, y)$ für festes $\lambda \in \mathbb{R}$ Scherung.

Beispiel 3.1.6. Der Homomorphismus $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ mit

$$\varphi(x_1, \dots, x_n) = \sum_{j=1}^{n} a_j x_j$$

für feste $a_1, \dots, a_n \in \mathbb{R}$ ist eine sogenannte <u>Linearform</u>.

Beispiel 3.1.7. Es sei $\mathcal{F}_0 = \{p\colon \mathbb{R} \to \mathbb{R},\ p \text{ ein Polynom}\}$. Für $p \in \mathcal{F}_0$ mit

$$p(x) = \sum_{\nu=0}^{n} a_\nu x^\nu$$

sei $\varphi(p) = p'$ die Ableitung mit

$$p'(x) = \sum_{\nu=1}^{n} a_\nu \cdot \nu \cdot x^{\nu-1}.$$

Dann ist φ ein Endomorphismus von $\mathcal{F}_0$. Er ist surjektiv, denn für

$$\tilde{p}(x) = \sum_{\nu=0}^{n} \frac{a_\nu}{\nu+1} x^{\nu+1}$$

ist $\varphi(\tilde{p}) = p$.
Er ist nicht injektiv, beispielsweise gilt $\varphi(p_0) = 0$ für jedes konstante Polynom $p_0(x) = a_0$. Somit ist φ ein Beispiel für einen linearen <u>Differentialoperator</u>.

Satz 3.1.1. *Für Vektorräume V, V' und V'' gilt:*

i) Ist $\varphi \in L(V, V')$ und $\psi \in L(V', V'')$, so ist $\psi \circ \varphi \in L(V, V'')$.

ii) Ist $\varphi\colon V \to V'$ ein Isomorphismus, so auch die inverse Abbildung $\varphi^{-1}\colon V' \to V$.

Beweis. i) Seien $\vec{v}, \vec{w} \in V$ und $\lambda, \mu \in K$. Dann gilt

$$(\psi \circ \varphi)(\lambda\vec{v} + \mu\vec{w}) = \psi(\lambda\varphi(\vec{v}) + \mu\varphi(\vec{w})) = \lambda(\psi \circ \varphi)(\vec{v}) + \mu(\psi \circ \varphi)(\vec{w}).$$

ii) Da φ^{-1} offenbar injektiv ist, müssen wir nur (L) zeigen. Es seien $\vec{v}', \vec{w}' \in V'$ und $\lambda, \mu \in K$ mit $\vec{v}' = \varphi(\vec{v})$ und $\vec{w}' = \varphi(\vec{w})$ für $\vec{v}, \vec{w} \in V$. Dann folgt

$$\varphi^{-1}(\lambda\vec{v}' + \mu\vec{w}') = \varphi^{-1}(\lambda\varphi(\vec{v}) + \mu\varphi(\vec{w})) = \varphi^{-1}(\varphi(\lambda\vec{v} + \mu\vec{w})) = \lambda\vec{v} + \mu\vec{w} = \lambda\varphi^{-1}(\vec{v}') + \mu\varphi^{-1}(\vec{w}').$$

□

Beispiel 3.1.8. Die Umkehrabbildung des Endomorphismus $\varphi\colon V \to V$, $\varphi(\vec{v}) = \lambda\vec{v}$ mit $\lambda \in K\backslash\{0\}$ ist $\varphi^{-1}\colon V \to V$, $\varphi^{-1}(\vec{v}) = \lambda^{-1}\vec{v}$.

Satz 3.1.2. *Für $\varphi, \psi \in L(V, V')$ und $\lambda \in L$ seien $\varphi + \psi, \lambda\varphi \in L(V, V')$ werteweise durch*

$$\begin{aligned} (\varphi + \psi)(\vec{v}) &= \varphi(\vec{v}) + \psi(\vec{v}) \quad \textit{und} \\ (\lambda\varphi)(\vec{v}) &= \lambda\varphi(\vec{v}) \end{aligned}$$

für alle $\vec{v} \in V$ definiert. Mit diesen Operationen wird $L(V, V')$ zu einem Vektorraum über K. Der Nullvektor ist der triviale Homomorphismus $\varphi_0(\vec{v}) = \vec{0}' \in V'$ für alle $\vec{v} \in V$.

Beweis. Durch Nachrechnen sieht man leicht die Gültigkeit der Regel (L) für $\varphi + \psi$ sowie $\lambda\varphi$. Auch die Vektorraumaxiome werden leicht nachgeprüft. $\square$

Satz 3.1.3. *Mit der oben definierten Addition sowie der Komposition von Abbildungen "$\circ$" als Multiplikation ist $(L(V,V),+,\circ)$ ein Ring mit Eins, der Endomorphismenring von V. Einselement ist die Identität $id_V: \vec{v} \to \vec{v}$ für alle $\vec{v} \in V$.*

Beweis. Nach Satz 3.1.2 ist $(L(V,V),+)$ eine abelsche Gruppe. Nach Satz 3.1.1 ist $\circ$ eine Verknüpfung auf $L(V,V)$. Das Assoziativgesetz der Multiplikation gilt, da es allgemein für die Komposition von Abbildungen gilt. Die Distributivgesetze gelten ebenfalls, da für $\psi, \varphi, \chi \in L(V,V)$ und alle $\vec{v} \in V$

$$(\varphi \circ (\psi + \chi))(\vec{v}) = \varphi(\psi(\vec{v}) + \chi(\vec{v})) = \varphi(\psi(\vec{v})) + \varphi(\chi(\vec{v})) = (\varphi \circ \psi)(\vec{v}) + (\varphi \circ \chi)(\vec{v}) = (\varphi \circ \psi + \varphi \circ \chi)(\vec{v}).$$

gilt. Ebenso ist

$$((\psi + \chi) \circ \varphi)(\vec{v}) = \psi(\varphi(\vec{v})) + \chi(\varphi(\vec{v})) = (\psi \circ \varphi + \chi \circ \varphi)(\vec{v}).$$

Für das Einselement gilt $id_V \circ \varphi = \varphi \circ id_V = \varphi$ für alle $\varphi \in L(V,V)$. $\square$

Satz 3.1.4. *Die Automorphismen von V bilden bzgl. $\circ$ eine Gruppe, die lineare Gruppe von V, geschrieben* $\mathrm{GL}(V)$.

Beweis. Es ist $\circ$ ist eine Verknüpfung auf $\mathrm{GL}(V)$: mit φ, ψ ist auch $\psi \circ \varphi \colon V \to V$ bijektiv, die Linearität folgt nach Satz 3.1.1. Das Assoziativgesetz folgt nach Satz 3.1.3, das Einselement ist id_V, das Inverse von φ ist die Umkehrabbildung φ^{-1}. $\square$

3.2 Kern und Bild

Satz 3.2.1. *Für $\varphi \in L(V,V')$ gilt:*

i) $\varphi(\vec{0}) = \vec{0}'$.

ii) Sind $\vec{v}_1, \ldots, \vec{v}_n \in V$ l.a., so auch $\varphi(\vec{v}_1), \ldots, \varphi(\vec{v}_n) \in V'$.

iii) Sind $\varphi(\vec{v}_1), \ldots, \varphi(\vec{v}_n) \in V'$ l.u., so auch $\vec{v}_1, \ldots, \vec{v}_n \in V$.

Beweis. i) Es gilt $\varphi(\vec{0}) = \varphi(\vec{0} + \vec{0}) = \varphi(\vec{0}) + \varphi(\vec{0})$, woraus $\varphi(\vec{0}) = \vec{0}'$ folgt.

ii) Aus $\lambda_1\vec{v}_1 + \cdots + \lambda_n\vec{v}_n = \vec{0}$ folgt

$$\lambda_1\varphi(\vec{v}_1) + \ldots + \lambda_n\varphi(\vec{v}_n) = \varphi(\lambda_1\vec{v}_1 + \ldots + \lambda_n\vec{v}_n) = \vec{0}'$$

nach (i).

iii) Dies folgt direkt aus (ii).

$\square$

Satz 3.2.2. *Es sei $\psi \in L(V,V')$. Dann gilt:*

i) Ist $M \subseteq V$, dann gilt $\varphi(\langle M \rangle) = \langle \varphi(M) \rangle$. Insbesondere ist das Bild eines Unterraums $U \subseteq V$ ein Unterraum von V'.

ii) Für einen Unterraum $U' \subseteq V'$ ist $\varphi^{-1}(U') = \{\vec{a} \in V : \varphi(\vec{a}) \in U'\}$ ein Unterraum von V.

Beweis. i) Fall 1: Es sei $M = \emptyset$.
Dann ist $\langle M \rangle = \{\vec{0}\}$ und $\varphi(\langle M \rangle) = \{\vec{0}'\} = \langle \emptyset \rangle = \langle \varphi(M) \rangle$.
Fall 2: $M \neq \emptyset$.
Wir zeigen zunächst $\varphi(\langle M \rangle) \subseteq \langle \varphi(M) \rangle$:
Ist $\vec{v} = \lambda_1 \vec{v}_1 + \ldots + \lambda_n \vec{v}_n \in \langle M \rangle$, dann ist $\varphi(\vec{v}) = \lambda_1 \varphi(\vec{v}_1) + \ldots + \lambda_n \varphi(\vec{v}_n) \in \langle \varphi(M) \rangle$.
Jetzt müssen wir noch $\langle \varphi(M) \rangle \subseteq \varphi(\langle M \rangle)$ zeigen.
Sei dazu $\vec{w} \in \langle \varphi(M) \rangle$, d.h. $\vec{w} = \lambda_1 \vec{w}_1 + \ldots + \lambda_n \vec{w}_n$ mit $\vec{w}_j \in \varphi(M)$. Dann gibt es $\vec{v}_j \in M$ mit $1 \leq j \leq n$ und $\varphi(\vec{v}_j) = \vec{w}_j$. Es gibt $\lambda_j \in K$ mit $\vec{v} = \lambda_1 \vec{v}_1 + \ldots + \lambda_n \vec{v}_n \in \langle M \rangle$ und damit

$$\varphi(\vec{v}) = \lambda_1 \varphi(\vec{v}_1) + \ldots + \lambda_n \varphi(\vec{v}_n) = \lambda_1 \vec{w}_1 + \ldots + \lambda_n \vec{w}_n = \vec{w},$$

also $\vec{w} \in \varphi(\langle M \rangle)$. Damit ist insgesamt $\langle \varphi(M) \rangle = \varphi(\langle M \rangle)$ gezeigt.

ii) Es ist $\varphi^{-1}(U') \neq \emptyset$, da $\varphi(\vec{0}) = \vec{0}' \in U'$ ist, also $\vec{0} \in \varphi^{-1}(U')$. Es seien $\vec{v}, \vec{w} \in \varphi^{-1}(U')$ gegeben, also $\varphi(\vec{v}), \varphi(\vec{w}) \in U'$. Ferner seien $\lambda, \mu \in K$. Dann gilt

$$\varphi(\lambda \vec{v} + \mu \vec{w}) = \lambda \varphi(\vec{v}) + \mu \varphi(\vec{w}) \in U',$$

d.h. $\lambda \vec{v} + \mu \vec{w} \in \varphi^{-1}(U')$.

□

Satz 3.2.3. *Es sei $\varphi \in L(V, V')$ und U ein Unterraum von V. Dann ist $\dim \varphi(U) \leq \dim U$. Insbesondere folgt $\dim V = \dim V'$ aus $V \cong V'$.*

Beweis. Sei O.B.d.A. $\dim U < \infty$. Dann besitzt U nach Satz 2.4.2 eine Basis B, also $U = \langle B \rangle$. Nach Satz 3.2.2 gilt $\varphi(U) = \varphi(\langle B \rangle) = \langle \varphi(B) \rangle$, also ist $\varphi(B)$ ein Erzeugendensystem von $\varphi(U)$. Dann ist $\dim \varphi(U) \leq |\varphi(B)| \leq |B| = \dim U$. □

Definition 3.2.1. Es sei $\varphi \in L(V, V')$ Dann heißt $\mathrm{Bild}(\varphi) = \varphi(V) = \{\varphi(\vec{a}) \colon \vec{a} \in V\}$ das Bild von φ. Die Menge $\mathrm{Kern}(\varphi) = \varphi^{-1}(\{\vec{0}'\}) = \{\vec{v} \in V \colon \varphi(\vec{v}) = \vec{0}'\}$ heißt Kern von φ. Nach Satz 3.2.2 sind dies Unterräume von V' bzw. V. Ferner heißt $\mathrm{rg}(\varphi) = \dim(\mathrm{Bild}(\varphi))$ der Rang und $\mathrm{def}(\varphi) = \dim(\mathrm{Kern}(\varphi))$ der Defekt von φ.

Beispiel 3.2.1. Wir betrachten die Projektion aus Beispiel 3.1.4: $V = \mathbb{R}^2$ und $\varphi((x, y)) = (x, 0)$. Dann ist $\mathrm{Bild}(\varphi) = \{(x, 0) \colon x \in \mathbb{R}\}$, d.h $\mathrm{rg}(\varphi) = 1$.
Andererseits ist $\mathrm{Kern}(\varphi) = \varphi^{-1}(\{(0, 0)\}) = \{\vec{v} \in V \colon \varphi(\vec{v}) = (0, 0)\} = \{(0, y) \colon y \in \mathbb{R}\}$, und somit $\mathrm{def}(\varphi) = 1$.

Satz 3.2.4. *Für $\varphi \in L(V, V')$ sind folgende Aussagen äquivalent:*

i) φ *ist injektiv.*

ii) $\mathrm{Kern}(\varphi) = \{\vec{0}\}$, *d.h.* $\mathrm{def}(\varphi) = 0$.

iii) Sind $\vec{v}_1, \ldots, \vec{v}_n \in V$ *l.u., so auch* $\varphi(\vec{v}_1), \ldots, \varphi(\vec{v}_n)$.

Beweis. (i)⇒(ii):
Ist φ injektiv, so folgt aus $\varphi(\vec{v}) = \vec{0}' = \varphi(\vec{0})$ schon $\vec{v} = \vec{0}$.

(ii)⇒(iii):
Es sei $\mathrm{Kern}(\varphi) = \{\vec{0}\}$, und $\vec{v}_1, \ldots, \vec{v}_n \in V$ l.u. gewählt. Aus $\lambda_1 \varphi(\vec{v}_1) + \ldots + \lambda_n \varphi(\vec{v}_n) = \vec{0}'$ folgt dann $\varphi(\lambda_1 \vec{v}_1 + \ldots + \lambda_n \vec{v}_n) = \vec{0}'$, d.h.

$$\lambda_1 \vec{v}_1 + \ldots + \lambda_n \vec{v}_n \in \mathrm{Kern}(\varphi) = \{\vec{0}'\} \Rightarrow \lambda_1 = \ldots = \lambda_n = 0,$$

da $\vec{v}_1, \ldots, \vec{v}_n$ l.u. sind.

(iii)$\Rightarrow$(i):
Es seien $\vec{v}_1, \vec{v}_2 \in V$ mit $\vec{v}_1 \neq \vec{v}_2$. Dann ist $\vec{v}_1 - \vec{v}_2 \neq \vec{0}$ l.u., also nach (iii) auch $\varphi(\vec{v}_1 - \vec{v}_2) = \varphi(\vec{v}_1) - \varphi(\vec{v}_2)$ l.u., und damit $\varphi(\vec{v}_1) \neq \varphi(\vec{v}_2)$. $\square$

Satz 3.2.5. *(Rangformel)*
Es sei $\varphi \in L(V, V')$ *mit* $\dim V < \infty$. *Dann gilt* $\operatorname{rg}(\varphi) + \operatorname{def}(\varphi) = \dim V$, *oder ausführlich*

$$\dim(\operatorname{Bild}(\varphi)) + \dim(\operatorname{Kern}(\varphi)) = \dim V.$$

Die beiden Extremfälle dieser Gleichheit sind

$$\begin{aligned} \operatorname{def}(\varphi) = 0 \quad &\Leftrightarrow \quad \dim(\operatorname{Bild}(\varphi)) = \dim V \\ \operatorname{rg}(\varphi) = 0 \quad &\Leftrightarrow \quad \operatorname{Kern}(\varphi) = V. \end{aligned}$$

Beweis. Wir wählen eine Basis $\vec{v}_1, \ldots, \vec{v}_r$ von $\operatorname{Kern}(\varphi)$ und ergänzen sie nach dem Austauschsatz von Steinitz (Satz 2.4.5) zu einer Basis $\vec{v}_1, \ldots, \vec{v}_r, \vec{v}_1, \ldots, \vec{v}_s$ von V. Wir zeigen im folgenden, dass

$$\varphi(V) = \langle\langle\, \varphi(\vec{w}_1), \ldots, \varphi(\vec{w}_s) \,\rangle\rangle$$

ist. Daraus folgt dann $\operatorname{rg}(\varphi) + \operatorname{def}(\varphi) = r + s = \dim V$, also die Behauptung.

i) $\varphi(\vec{w}_1), \ldots, \varphi(\vec{w}_s)$ sind l.u.:
Es sei $\vec{0} = \lambda_1 \varphi(\vec{w}_1) + \ldots + \lambda_s \varphi(\vec{w}_s) = \varphi(\lambda_1 \vec{w}_1 + \ldots + \lambda_s \vec{w}_s)$ mit $\lambda_i \in K$. Es folgt

$$\lambda_1 \vec{w}_1 + \ldots + \lambda_s \vec{w}_s \in \operatorname{Kern}(\varphi) \Rightarrow \lambda_1 \vec{w}_1 + \ldots + \lambda_s \vec{w}_s = \mu_1 \vec{v}_1 + \ldots + \mu_r \vec{v}_r$$

mit $\mu_j \in K$. Dann ist $\mu_1 \vec{v}_1 + \ldots + \mu_r \vec{v}_r + (-\lambda_1)\vec{w}_1 + \ldots + (-\lambda_s)\vec{w}_s = \vec{0}$, woraus nun eben $\lambda_1 = \ldots = \lambda_s = \mu_1 = \ldots = \mu_r = 0$ folgt, da $\vec{v}_1, \ldots, \vec{v}_r, \vec{w}_1, \ldots, \vec{w}_s$ l.u. sind.

ii) $\varphi(V) = \langle\, \varphi(\vec{w}_1), \ldots, \varphi(w_s) \,\rangle$:
Es sei dazu $\vec{u}' \in \varphi(V)$, etwa $\vec{u}' = \varphi(\vec{u})$ für ein $\vec{u} \in V$. Dann ist

$$\vec{u} = \alpha \vec{v}_1 + \ldots + \alpha_r \vec{v}_r + \beta_1 \vec{w}_1 + \ldots + \beta_s \vec{w}_s$$

mit $\alpha_i, \beta_j \in K$. Es folgt

$$\begin{aligned} \vec{u}' &= \varphi(\vec{u}) = \alpha_1 \varphi(\vec{v}_1) + \ldots + \alpha_r \varphi(\vec{v}_r) + \beta_1 \varphi(\vec{w}_1) + \ldots + \beta_s \varphi(\vec{w}_s) \\ &= \beta_1 \varphi(\vec{w}_1) + \ldots + \beta_s \varphi(\vec{w}_s) \in \langle\, \varphi(\vec{w}_1), \ldots, \varphi(\vec{w}_s) \,\rangle, \end{aligned}$$

denn es ist $\varphi(\vec{v}_i) = \vec{0}$ für die $\vec{v}_i \in \operatorname{Kern}(\varphi)$.

$\square$

Satz 3.2.6. *Es sei* $\varphi \in L(V, V')$. *Dann gilt:*

i) Ist $\dim V < \infty$, *so ist* φ *genau dann injektiv, wenn* $\operatorname{rg}(\varphi) = \dim V$ *gilt.*

ii) Ist $\dim V = \dim V'$, *so ist* φ *genau dann injektiv, wenn* φ *surjektiv ist.*

Beweis. Es gelten die Äquivalenzen

i) φ injektiv $\underset{\text{Satz 3.2.4}}{\Leftrightarrow} \operatorname{def}(\varphi) = 0 \underset{\text{Satz 3.2.5}}{\Leftrightarrow} \operatorname{rg}(\varphi) = \dim V$.

ii) φ injektiv $\underset{\text{(i)}}{\Leftrightarrow} \dim(\operatorname{Bild}(\varphi)) = \dim V = \dim V' \underset{\text{Satz 2.4.7(ii)}}{\Leftrightarrow} \varphi(V) = V'$.

$\square$

3.3 Lineare Fortsetzung

Wir kommen nun zur Frage der Beschreibung linearer Abbildungen:

Satz 3.3.1. *Es sei $V = \langle\langle \vec{v}_1, \ldots, \vec{v}_n \rangle\rangle$, und es seien $\vec{w}_1, \ldots, \vec{w}_n \in V'$ beliebige Vektoren. Dann gibt es genau eine lineare Abbildung $\varphi \in L(V, V')$ mit $\varphi(\vec{v}_j) = \vec{w}_j$ für $j = 1 \ldots n$.*
Also ist eine lineare Abbildung schon völlig festgelegt, wenn ihre Werte auf einer Basis von V bekannt sind. Andererseits können diese Werte beliebig vorgeschrieben werden.

Beweis. i) Existenz:
Zu $\vec{v} \in V$ existieren eindeutig bestimmte $\lambda_1, \ldots, \lambda_n \in K$ mit $\vec{v} = \lambda_1\vec{v}_1 + \cdots + \lambda_n\vec{v}_n$. Die Abbildung φ wird dann durch $\varphi(\vec{v}) = \lambda_1\vec{w}_1 + \cdots + \lambda_n\vec{w}_n$ definiert. Es bleibt zu zeigen, dass φ linear ist. Dazu seien

$$\vec{v} = \sum_{j=1}^{n} \lambda_j \vec{v}_j \quad \text{und} \quad \vec{w} = \sum_{j=1}^{n} \mu_j \vec{w}_j$$

aus V beliebig und $\alpha, \beta \in K$. Dann ist

$$\begin{aligned} \varphi(\alpha\vec{v} + \beta\vec{w}) &= \varphi\left(\sum_{j=1}^{n}(\alpha\lambda_j + \beta\mu_j)\vec{v}_j\right) = \sum_{j=1}^{n}(\alpha\lambda_j + \beta\mu_j)\vec{w}_j \\ &= \alpha\sum_{j=1}^{n}\lambda_j\vec{w}_j + \beta\sum_{j=1}^{n}\mu_j\vec{w}_j = \alpha\varphi(\vec{v}) + \beta\varphi(\vec{w}). \end{aligned}$$

ii) Eindeutigkeit:
Es sei $\psi \in L(V, V')$ mit $\psi(\vec{v}_j) = \vec{w}_j$ für $j = 1 \ldots n$. Für jedes $\vec{v} = \lambda_1\vec{v}_1 + \ldots + \lambda_n\vec{v}_n$ muss dann

$$\psi(\vec{v}) = \lambda_1\psi(\vec{v}_1) + \ldots + \lambda_n\psi(\vec{v}_n) = \lambda_1\vec{w}_1 + \cdots + \lambda_n\vec{w}_n = \varphi(\vec{v})$$

gelten, also $\psi = \varphi$.

□

Definition 3.3.1. Die durch die Zuordnung $\vec{a}_j \to \vec{b}_j$ nach dem Satz eindeutig festgelegte lineare Abbildung $\varphi \in L(V, V')$ heißt lineare Fortsetzung dieser Zuordnung.

Satz 3.3.2. *Es seien V und V' endlichdimensionale Vektorräume über K. Dann gilt*

$$V \cong V' \Leftrightarrow \dim V = \dim V'.$$

Insbesondere ist also jeder n-dimensionale VR über K isomorph zum Standardraum K^n.

Beweis. "$\Rightarrow$":
Dies folgt aus Satz 3.2.3.
"$\Leftarrow$":
Es sei $\dim V = \dim V' = n \in \mathbb{N}$.
Für $n = 0$ ist die Aussage trivial, andernfalls ist $V = \langle\langle \vec{v}_1, \ldots, \vec{v}_n \rangle\rangle$ und $V' = \langle\langle \vec{w}_1, \ldots, \vec{w}_n \rangle\rangle$. Wir definieren $\varphi \in L(V, V')$ als die lineare Fortsetzung der Zuordnung $\vec{v}_j \to \vec{w}_j$ für $j = 1 \ldots n$ und müssen nur noch die Bijektivität von φ zeigen. Ist $\vec{w} = \lambda_1\vec{w}_1 + \ldots + \lambda_n\vec{w}_n \in V'$ beliebig, so ist $\vec{w} = \varphi(\vec{v})$ für $\vec{v} = \lambda_1\vec{v}_1 + \ldots + \lambda_n\vec{v}_n \in V$ nach Definition von φ. Nach Satz 3.2.6(ii) ist φ auch injektiv, und damit ein Isomorphismus. □

Satz 3.3.3. *Es seien* $\varphi \in L(V, V')$ *und* $\psi \in L(V', V'')$, *also* $\psi \circ \varphi \in L(V, V'')$. *Dann gilt*

$$\mathrm{rg}(\varphi) + \mathrm{rg}(\psi) - \dim V' \leq \mathrm{rg}(\psi \circ \varphi) \leq \min\{\mathrm{rg}(\varphi), \mathrm{rg}(\psi)\}.$$

Beweis. Die rechte Ungleichung folgt aus

$$\mathrm{rg}(\psi \circ \varphi) = \dim(\mathrm{Bild}(\psi \circ \varphi)) = \dim(\psi(\varphi(V))) \leq \begin{cases} \dim \varphi(V) = \mathrm{rg}(\varphi) \\ \dim \psi(V') = \mathrm{rg}(\psi) \end{cases}$$

Für die linke Ungleichung betrachten wir die Abbildung $\psi^* = \psi|_{\varphi(V)}$, die Beschränkung von ψ auf den Unterraum $\varphi(V)$, also $\psi^* \colon \varphi(V) \to V''$. Dann gilt

$$\begin{aligned} \mathrm{rg}(\psi \circ \varphi) &= \dim \psi(\varphi(V)) = \dim \psi^*(V) = \mathrm{rg}(\psi^*) \\ &= \dim \varphi(V) - \mathrm{def}(\psi^*) \geq \dim \varphi(V) - \mathrm{def}(\psi) = \mathrm{rg}(\varphi) - (\dim V' - \mathrm{rg}(\psi)) \end{aligned}$$

nach Satz 3.2.5. □

3.4 Isomorphismen

In der Mathematik, vor allem dem Teilgebiet der Algebra, hat man oft folgende Situation vorliegen: Es sind zwei Mengen G_1 und G_2 gegeben, auf denen durch eine (oder mehrere) Verknüpfungen dieselbe algebraische Struktur (z.B. Gruppe, Ring, Körper oder Vektorraum) gegeben ist. Die Beziehungen zwischen den Elementen in G_1, die durch die Verknüpfungen gegeben sind, sind dieselben, wie die zwischen den Elementen von G_2 (Relationstreue).
Man erhält also die Elemente von G_2 durch Namensänderung aus den Elementen von G_1. In diesem Fall sagt man, dass G_1 und G_2 isomorph sind. Die Abbildung, die die Namensänderung beschreibt, heißt Isomorphismus.
Wir haben schon das Beispiel des Vektorraum- Isomorphismus kennengelernt:
Nach Satz 3.3.2 ist jeder Vektorraum V mit $\dim V = n$ zum K^n isomorph.

Für eine feste Basis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ ist dieser Isomorphismus φ durch

$$\varphi(\lambda_1 \vec{v}_1 + \ldots + \lambda_n \vec{v}_n) = (\lambda_1, \ldots, \lambda_n).$$

gegeben. Durch φ wird also der Vektor $\vec{v} = \lambda_1 \vec{v}_1 + \ldots + \lambda_n \vec{v}_n$ in $(\lambda_1, \ldots, \lambda_n)$ umbenannt.

Es ergibt sich folgendes Bild:

alte Namen: $\underbrace{\vec{v} = \lambda_1 \vec{v}_1 + \ldots + \lambda_n \vec{v}_n,\ \vec{w} = \mu_1 \vec{v}_1 + \ldots + \mu_n \vec{v}_n,\ \vec{v} + \vec{w} = (\lambda_1 + \mu_1)\vec{v}_1 + \ldots + (\lambda_n + \mu_n)\vec{v}_n}$

$$\Downarrow \varphi$$

neue Namen: $\varphi(\vec{v}) = (\lambda_1, \ldots, \lambda_n), \varphi(\vec{w}) = (\mu_1, \ldots, \mu_n), \varphi(\vec{v}+\vec{w}) = (\lambda_1+\mu_1, \ldots, \lambda_n+\mu_n) = \varphi(\vec{v})+\varphi(\vec{w})$

Wir geben nun die Definition des Isomorphismus für die Strukturen Gruppe und Ring.

Definition 3.4.1. Zwei Gruppen $(G, \circ)$ und (G, Δ) heißen isomorph, wenn eine bijektive Abbildung $\varphi \colon G \to G'$ mit $\varphi(a \circ b) = \varphi(a) \Delta \varphi(b)$ für alle $a, b \in G$ existiert.
Dieses φ heißt (Gruppen-) isomorphismus.

Beispiel 3.4.1. Es sei $\mathbb{F}_2 = \{0, 1\}$ der Körper mit zwei Elementen und $(G, +) = (\mathbb{F}_2^2, +)$.
Es ist also $G = \{(0,0), (0,1), (1,0), (1,1)\}$ mit der Verknüpfungstafel

+	(0,0)	(0,1)	(1,0)	(1,1)
(0,0)	(0,0)	(0,1)	(1,0)	(1,1)
(0,1)	(0,1)	(0,0)	(1,1)	(1,0)
(1,0)	(1,0)	(1,1)	(0,0)	(0,1)
(1,1)	(1,1)	(1,0)	(0,1)	(0,0)

$(*)$

Nun sei $(G', \circ)$ die Gruppe bestehend aus den vier Permutationen id, Π_1, Π_2, Π_3 von $\{1, 2, 3, 4\}$ mit der Komposition als Verknüpfung:

$$id = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \quad \Pi_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}, \quad \Pi_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}, \quad \Pi_3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}$$

mit der Verknüpfungstafel

$\circ$	id	Π_1	Π_2	Π_3
id	id	Π_1	Π_2	Π_3
Π_1	Π_1	id	Π_3	Π_2
Π_2	Π_2	Π_3	id	Π_1
Π_3	Π_3	Π_2	Π_1	id

$(**)$

Es sei $\varphi\colon G \to G'$ durch

$$\varphi((0,0)) = id, \quad \varphi((0,1)) = \Pi_1, \quad \varphi((1,0)) = \Pi_2, \quad \varphi((1,1)) = \Pi_3, \quad .$$

gegeben. Führt man die Namensänderung $x \to \varphi(x)$ durch, so geht die Verknüpfungstafel $(*)$ in die Verknüpfungstafel $(**)$ über. Es ist also φ ein Isomorphismus von $(G, +)$ auf $(G', \circ)$. Die beiden Gruppen sind isomorph.

Gruppen- und Ringisomorphismen haben nun Eigenschaften, die den schon besprochenen Vektorraumisomorphismen entsprechen. Wir beschränken uns bei der Diskussion auf Gruppenisomorphismen.

Satz 3.4.1. *Es seien $(G, \circ)$ bzw. (G, Δ) Gruppen mit Einselementen e bzw. e'. Weiter sei $\varphi :\ G \to G'$ ein Gruppenisomorphismus.*
Dann gilt:

i) $\varphi(e) = e'$

ii) $\varphi(a^{-1}) = \varphi(a)^{-1}$ *für alle* $a \in G$

iii) Die Umkehrabbildung $\varphi^{-1}\colon G' \to G$ *ist ebenfalls ein Gruppenisomorphismus.*

Beweis.
i) Für $a \in G$ gilt: $\varphi(a) = \varphi(a \circ e) = \varphi(a)\Delta\varphi(e) \Rightarrow \varphi(e) = e'$

ii) Für $a \in G$ gilt: $e' = \varphi(e) = \varphi(a \circ a^{-1}) = \varphi(a)\Delta\varphi(a^{-1}) \Rightarrow \varphi(a^{-1}) = \varphi(a)^{-1}$

iii) Offenbar ist φ bijektiv. Es bleibt die Relationstreue zu zeigen: es seien $a', b' \in G'$. Dann gibt es $a, b \in G$ mit $\varphi(a) = a'$ und $\varphi(b) = b'$. Es ist

$$\varphi^{-1}(a'\Delta b') = \varphi^{-1}(\varphi(a)\Delta\varphi(b)) = \varphi^{-1}(\varphi(a \circ b)) = a \circ b = \varphi^{-1}(a') \circ \varphi^{-1}(b').$$

□

3.5 Lineare Abbildungen und Matrizen

Es sei K ein Körper.

Definition 3.5.1. Unter einer Matrix vom Typ (m,n) mit $m,n\in\mathbb{N}$ (oder einer $m\times n$-Matrix) über einem Körper K versteht man ein rechteckiges Schema der Gestalt

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

mit $a_{ij}\in K$ für $1\le i\le m$ und $1\le j\le n$. Die Einträge a_{ij} heißen Komponenten oder Koeffizienten der Matrix. Den Vektor $\vec{a}_i=(a_{i1},\dots,a_{in})\in K^n$ für $1\le i\le m$ bezeichnet man als den i-ten Zeilenvektor (oder kurz die i-te Zeile) von A, den Vektor $\vec{b}_j=(a_{1j},\dots,a_{mj})\in K^m$ für $1\le j\le n$ als den j-ten Spaltenvektor (oder kurz die j-te Spalte) von A. Wir schreiben $\vec{b}_j$ oft in Spaltenschreibweise

$$\vec{b}_j=\begin{pmatrix} a_{1j}\\ \vdots\\ a_{mj}\end{pmatrix}.$$

Man schreibt dann auch kurz

$$A=\begin{pmatrix} \vec{a}_1\\ \vdots\\ \vec{a}_m\end{pmatrix}=(\vec{b}_1,\dots,\vec{b}_n)=(a_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}}.$$

Die Menge aller Matrizen vom Typ (m,n) bezeichnet man mit $K^{(m,n)}$ oder $K^{m\times n}$. Die Gerade in A, auf der die Elemente $a_{11},\dots,a_{rr}$ mit $r=\min(m,n)$ stehen, nennt man die Hauptdiagonale von A. Durch Spiegelung an der Hauptdiagonalen erhält man aus A die Matrix A^T, die Transponierte von A. Sie hat die Gestalt

$$A^T=\begin{pmatrix} a_{11} & \cdots & a_{m1}\\ a_{12} & \cdots & a_{m2}\\ \vdots & & \vdots\\ a_{1n} & \cdots & a_{mn}\end{pmatrix}\in K^{(n,m)}.$$

Die Zeilen von A werden also die Spalten von A^T, und die Spalten von A werden die Zeilen von A^T, also

$$A=(a_{kl})_{\substack{1\le k\le m\\ 1\le l\le n}}\Leftrightarrow A^T=(b_{kl})_{\substack{1\le k\le n\\ 1\le l\le m}}$$

mit $b_{kl}=a_{lk}$.

Matrizen desselben Typs über K können komponentenweise addiert und mit Skalaren aus K multipliziert werden:

Definition 3.5.2. Es seien $A,B\in K^{(m,n)}$ mit

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{m1} & \cdots & a_{mn}\end{pmatrix},\quad \text{und}\quad B=\begin{pmatrix} b_{11} & \cdots & b_{1n}\\ \vdots & & \vdots\\ b_{m1} & \cdots & b_{mn}\end{pmatrix}.$$

Dann versteht man unter der Summe von A und B die Matrix

$$A + B = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix}.$$

Ist $\lambda \in K$, so setzt man

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

Man schreibt $(-1) \cdot A = -A$.

Beispiel 3.5.1. Es sei $K = \mathbb{R}$ und

$$A = \begin{pmatrix} 3 & -1 & 5 & 7 \\ 2 & 1 & 4 & 3 \\ 0 & -8 & 2 & 6 \end{pmatrix} \quad \text{sowie} \quad B = \begin{pmatrix} 2 & 0 & -4 & -3 \\ -2 & -1 & 0 & -5 \\ 1 & 8 & 1 & 4 \end{pmatrix}.$$

Dann ist

$$A + B = \begin{pmatrix} 5 & -1 & 1 & 4 \\ 0 & 0 & 4 & -2 \\ 1 & 0 & 3 & 10 \end{pmatrix} \quad \text{und} \quad 3A = \begin{pmatrix} 9 & -3 & 15 & 21 \\ 6 & 3 & 12 & 9 \\ 0 & -24 & 6 & 18 \end{pmatrix}.$$

Definition 3.5.3. Die Matrix vom Typ (m, n), deren sämtliche Komponenten gleich null sind, nennt man die Nullmatrix

$$0 = 0^{(m,n)} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

Man zeigt leicht

Satz 3.5.1. *Der $K^{(m,n)}$ bildet bzgl. der Matrizenaddition und Skalarmultiplikation mit der Nullmatrix als Nullelement einen Vektorraum über K mit Dimension* $\dim K^{(m,n)} = m \cdot n$.

Beweis. Übungen □

Von großer Bedeutung ist auch das Produkt von Matrizen A und B. Im allgemeinen Fall sind hier jedoch A und B von verschiedenem Typ.

Definition 3.5.4. Sind

$$A = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} \in K^{(m,n)} \quad \text{und} \quad B = (b_{kl})_{\substack{1 \le k \le n \\ 1 \le l \le r}} \in K^{(n,r)}$$

so versteht man unter dem Produkt $C = AB$ die Matrix

$$C = (c_{il})_{\substack{1 \le i \le m \\ 1 \le l \le r}} \in K^{(m,r)} \quad \text{mit} \quad c_{il} = \sum_{j=1}^{n} a_{ij} b_{jl}$$

und $1 \le i \le m$ sowie $1 \le l \le r$.

Bemerkung 3.5.1. Das Element in der i-ten Zeile und der l-ten Spalte der Produktmatrix C wird also erhalten, indem die Elemente der i-ten Zeile von A und der l-ten Spalte von B paarweise multipliziert und die Produkte addiert werden:

$$C = \left(\quad \right) \leftarrow A = \left(\text{i-te Zeile} \right), \quad B = \left(\; \overset{\text{l-te Spalte}}{|} \; \right),$$

Damit das Produkt zweier Matrizen A und B definiert ist, muss die Anzahl der Spalten von A mit der Anzahl der Zeilen B übereinstimmen. Die Produktmatrix $C = AB$ hat dann dieselbe Anzahl Zeilen wie A, und die dieselbe Anzahl Spalten wie B.

Beispiel 3.5.2. Es sei $K = \mathbb{R}$ und

$$A = \begin{pmatrix} 2 & 3 & -1 & 5 \\ 0 & -2 & 1 & 1 \\ 3 & 7 & 4 & 2 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & -1 \\ 3 & 0 \\ 2 & 5 \\ 1 & 6 \end{pmatrix}.$$

Dann ist das Produkt

$$A \cdot B = \begin{pmatrix} 14 & 23 \\ -3 & 11 \\ 34 & 29 \end{pmatrix}.$$

Der Grund für die komplizierte Definition der Multiplikation von Matrizen wird klar, wenn wir Matrizen in Zusammenhang mit linearen Abbildungen bringen.

Definition 3.5.5. Es sei $\dim V = n < \infty$ sowie $\dim V' = m < \infty$, $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ eine Basis von V sowie $\mathcal{B}' = \{\vec{b}'_1, \ldots, \vec{b}'_m\}$ eine Basis von V'. Ferner sei $\varphi \in L(V, V')$. Wegen $\varphi(\vec{b}_l) \in V'$ gibt es eindeutig bestimmte $\alpha_{kl} \in K$ mit

$$\varphi(\vec{b}_l) = \alpha_{1l}\vec{b}'_1 + \ldots + \alpha_{ml}\vec{b}'_m = \sum_{k=1}^{m} \alpha_{kl}\vec{b}'_k \tag{$*$}$$

mit $l = 1, \ldots, n$.
Es sei $\mathcal{M}(\varphi; \mathcal{B}', \mathcal{B}) = (\alpha_{kl})$ mit $1 \leq k \leq m$ und $1 \leq l \leq n$. Dabei heißt $\mathcal{M}(\varphi; \mathcal{B}', \mathcal{B})$ die der Abbildung φ bzgl. der Basen $\mathcal{B}$ und $\mathcal{B}'$ zugeordnete Matrix (auch Darstellungsmatrix von φ bzgl. $\mathcal{B}$ und $\mathcal{B}'$).

Ein erster Zusammenhang mit der Matrixmultiplikation ergibt sich durch

Satz 3.5.2. *Mit den obigen Bezeichnungen sei*

$$\vec{v} = \lambda_1\vec{b}_1 + \ldots + \lambda_n\vec{b}_n \in V \quad \textit{und} \quad \varphi(\vec{v}) = \lambda'_1\vec{b}'_1 + \ldots + \lambda'_m\vec{b}'_m \in V'.$$

Dann gilt

$$\begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_m \end{pmatrix} = \mathcal{M}(\varphi; \mathcal{B}', \mathcal{B}) \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}. \tag{$**$}$$

Definition 3.5.6. Wir nennen $\lambda_1, \ldots, \lambda_n$ die Koordinaten von $\vec{v}$ bzgl. $\mathcal{B}$ (entsprechend sind $\lambda'_1, \ldots, \lambda'_m$ die Koordinaten von $\varphi(\vec{v})$ bzgl. $\mathcal{B}'$).

Beweis. (Beweis von Satz 3.5.2)
Es ist

$$\varphi(\vec{v}) = \sum_{l=1}^{n} \lambda_l \varphi(\vec{b}_l) \underset{(*)}{=} \sum_{l=1}^{n} \lambda_l \sum_{k=1}^{m} \alpha_{kl} \vec{b}'_k,$$

also

$$\lambda'_k = \sum_{l=1}^{n} \alpha_{kl} \lambda_l$$

für $k = 1, \dots, m$, d.h.

$$\begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_m \end{pmatrix} = \begin{pmatrix} \alpha_{11}\lambda_1 & + & \cdots & + & \alpha_{1n}\lambda_n \\ & & \vdots & & \\ \alpha_{m1}\lambda_1 & + & \cdots & + & \alpha_{mn}\lambda_n \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

□

Bemerkung 3.5.2. Die Matrix $\mathcal{M}(\varphi; \mathcal{B}', \mathcal{B})$ lässt sich folgendermaßen einfach in Worten beschreiben: Nach $(*)$ besteht die l-te Spalte von $\mathcal{M}(\varphi; \mathcal{B}', \mathcal{B})$ aus den Koordinaten des Bilds $\varphi(\vec{b}_l)$ des l-ten Basisvektors.

Wir betrachten jetzt die Matrizen, die einigen der im letzten Paragraphen als Beispiele aufgeführten linearen Abbildungen zugeordnet sind, sowie ein paar andere Beispiele.

Beispiel 3.5.3. Sei $\dim V = n$ und $\dim V' = m$ sowie $\varphi_0 \colon V \to V'$ der triviale Homomorphismus. Dann ist bzgl. beliebiger Basen $\mathcal{B}$ und $\mathcal{B}'$ von V und V' stets $\mathcal{M}(\varphi_0, \mathcal{B}', \mathcal{B}) = 0^{(m,n)}$ die Nullmatrix.

Beispiel 3.5.4. Es sei $\varphi \colon V \to V$ mit $\varphi(\vec{a}) = \lambda\vec{a}$ für festes $\lambda \in K$ und $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ irgendeine Basis von V. Die l-te Spalte von $\mathcal{M}(\varphi, \mathcal{B}, \mathcal{B})$ besteht dann aus den Koordinaten $\varphi(\vec{b}_l) = \lambda\vec{b}_l$ bzgl. $\mathcal{B}$, also

$$\mathcal{M}(\varphi, \mathcal{B}, \mathcal{B}) = \begin{pmatrix} \lambda & & \\ & \ddots & \\ & & \lambda \end{pmatrix}.$$

Beispiel 3.5.5. Es sei $V = \mathbb{R}^2$ und $\varphi \colon V \to V$ mit $\varphi(x, y) = (\lambda x, \mu y)$ die Eulerabbildung. Es sei $\mathcal{B} = \{\vec{e}_1, \vec{e}_2\}$ mit $\vec{e}_1 = (1, 0)$ und $\vec{e}_2 = (0, 1)$ die Standardbasis. Dann gilt $\varphi(\vec{e}_1) = \lambda\vec{e}_1$ sowie $\varphi(\vec{e}_2) = \mu\vec{e}_2$, also

$$\mathcal{M}(\varphi, \mathcal{B}, \mathcal{B}) = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}.$$

Die Darstellungsmatrix hängt im allgemeinen von der Wahl der Basen ab. Ist etwa $\tilde{\mathcal{B}} = \{\vec{b}_1, \vec{b}_2\}$ mit $\vec{b}_1 = (1, 1)$ und $\vec{b}_2 = (1, -1)$, so ist

$$\begin{aligned} \varphi(\vec{b}_1) &= (\lambda, \mu) = \frac{\lambda + \mu}{2}\vec{b}_1 + \frac{\lambda - \mu}{2}\vec{b}_2 \quad \text{und} \\ \varphi(\vec{b}_2) &= (\lambda, -\mu) = \frac{\lambda - \mu}{2}\vec{b}_1 + \frac{\lambda + \mu}{2}\vec{b}_2 \end{aligned}$$

und damit

$$\mathcal{M}(\varphi, \tilde{\mathcal{B}}, \tilde{\mathcal{B}}) = \begin{pmatrix} \frac{\lambda+\mu}{2} & \frac{\lambda-\mu}{2} \\ \frac{\lambda-\mu}{2} & \frac{\lambda+\mu}{2} \end{pmatrix}.$$

Eine der wichtigen Aufgaben der linearen Algebra ist es, zu einer gegebenen Abbildung φ Basen zu finden, bzgl. denen die Darstellungsmatrix besonders einfach wird.

Beispiel 3.5.6. Sei $V = \mathbb{R}^3$ mit der Standardbasis $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ versehen, und $\varphi \in L(V,V)$ die lineare Abbildung mit bzgl. $\mathcal{B}$ zugeordneter Matrix

$$\mathcal{M}(\varphi, \mathcal{B}, \mathcal{B}) = \begin{pmatrix} -3 & -2 & -4 \\ -3 & 0 & -3 \\ 8 & 4 & 9 \end{pmatrix}.$$

Wir betrachten die neue Basis $\tilde{\mathcal{B}} = \{\vec{b}_1, \vec{b}_2, \vec{b}_3\}$ mit

$$\vec{b}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \vec{b}_2 = \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix} \quad \text{und} \quad \vec{b}_3 = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}.$$

Es ist

$$\mathcal{M}(\varphi, \mathcal{B}, \mathcal{B}) \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \mathcal{M}(\varphi, \mathcal{B}, \mathcal{B}) \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \\ -8 \end{pmatrix}$$

und

$$\mathcal{M}(\varphi, \mathcal{B}, \mathcal{B}) \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} -3 \\ -3 \\ 6 \end{pmatrix},$$

also kurz $\varphi(\vec{b}_1) = \vec{b}_1$, $\varphi(\vec{b}_2) = 2\vec{b}_2$ und $\varphi(\vec{b}_3) = 3\vec{b}_3$. Bezüglich der Basis $\tilde{\mathcal{B}}$ hat φ die einfache Diagonalmatrix

$$\mathcal{M}(\varphi; \tilde{\mathcal{B}}, \tilde{\mathcal{B}}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

als Darstellungsmatrix. Die Abbildung φ streckt daher den $\mathbb{R}^3$ in den Richtungen von $\vec{b}_1$, $\vec{b}_2$ und $\vec{b}_3$ um die Faktoren 1, 2 und 3. Man nennt dann $\vec{b}_1$, $\vec{b}_2$ und $\vec{b}_3$ die Eigenvektoren von φ mit zugehörigen Eigenwerten $1, 2, 3$.

Beispiel 3.5.7. Es sei $\varphi \in L(K^n, K)$ mit

$$\varphi(x_1, \ldots, x_n) = \sum_{j=1}^{n} a_j x_j$$

eine Linearform und $\mathcal{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ bzw. $\mathcal{B}' = \{1\}$ die Standardbasen von K^n bzw. K. Dann ist

$$\varphi(\vec{e}_i) = \varphi(0, \ldots, 0, \underset{\substack{\uparrow \\ i\text{-te Stelle}}}{1}, 0, \ldots, 0) = a_i,$$

also ist die Darstellungsmatrix die Zeile $\mathcal{M}(\varphi, \mathcal{B}', \mathcal{B}) = (a_1, \ldots, a_n)$.

Beispiel 3.5.8. Es sei

$$V = \{p \colon \mathbb{R} \to \mathbb{R},\ p \text{ Polynom vom Grad } \leq n\}$$

mit der Basis $\mathcal{B} = \{1, x, \ldots, x^n\}$ und $\varphi \colon V \to V$ mit $\varphi(p) = p'$ die Ableitung. Dann gilt:

$$\begin{array}{rcccl} \varphi(1) & = & 0 & = & 0 \cdot 1 + 0 \cdot x + \cdots + 0 \cdot x^{n-1} + 0 \cdot x^n \\ \varphi(x) & = & 1 & = & 1 \cdot 1 + 0 \cdot x + \cdots + 0 \cdot x^{n-1} + 0 \cdot x^n \\ \varphi(x^2) & = & 2x & = & 0 \cdot 1 + 2 \cdot x + \cdots + 0 \cdot x^{n-1} + 0 \cdot x^n \\ \vdots & & \vdots & & \qquad\qquad\quad \vdots \\ \varphi(x^n) & = & nx^{n-1} & = & 0 \cdot 1 + 0 \cdot x + \cdots + n \cdot x^{n-1} + 0 \cdot x^n \end{array}$$

und somit

$$\mathcal{M}(\varphi,\mathcal{B},\mathcal{B}) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & n \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{(n+1,n+1)}.$$

Im folgenden soll nun die Analogie zwischen linearen Abbildungen und Matrizen weiter untersucht werden. Dies geschieht mittels des im letzten Abschnitt entwickelten Begriffs des Isomorphismus. Jede der betrachteten Strukturen von linearen Abbildungen- die Gruppe von Satz 3.1.4, der Ring von Satz 3.1.3 und der Vektorraum von Satz 3.1.2- ist isomorph zu einer entsprechenden Struktur von Matrizen.

Satz 3.5.3. *Es sei* $\dim V = n$ *und* $\dim V' = m$. *Dann gilt* $L(V,V') \cong K^{(m,n)}$.
Genauer: Ist $\mathcal{B} = \{\vec{b}_1,\dots,\vec{b}_n\}$ *eine Basis von* V *sowie* $\mathcal{B}' = \{\vec{b}'_1,\dots,\vec{b}'_m\}$ *eine Basis von* V', *so ist durch*

$$\Phi(\varphi) = \mathcal{M}(\varphi,\mathcal{B}',\mathcal{B})$$

für $\varphi \in L(V,V')$ *ein Isomorphismus*

$$\Phi\colon L(V,V') \to K^{(m,n)}$$

gegeben. Insbesondere ist $\dim L(V,V') = \dim K^{(m,n)} = m \cdot n$. *Der inverse Isomorphismus*

$$\Phi^{-1}\colon K^{(m,n)} \to L(V,V')$$

ist durch

$$\Phi^{-1}(A) = \varphi_A, \quad \varphi_A(\lambda_1\vec{b}_1 + \ldots + \lambda_n\vec{b}_n) = \lambda'_1\vec{b}'_1 + \ldots + \lambda'_m\vec{b}'_m, \quad \begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_m \end{pmatrix} = A \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

gegeben.

Beweis. Man rechnet leicht nach, dass die Abbildung Φ linear ist.
Wir zeigen noch die Injektivität und Surjektivität von Φ:
Es sei $\varphi \in \mathrm{Kern}(\Phi)$, d.h. $\mathcal{M}(\varphi,\mathcal{B}',\mathcal{B}) = 0^{(m,n)}$. Dann ist nach Satz 3.5.2 $\varphi(\vec{v}) = \vec{0}'$ für alle $\vec{v} \in V$, d.h. φ ist der Nullvektor von $L(V,V')$, also $\mathrm{Kern}(\Phi) = \{0\}$ und Φ ist injektiv nach Satz 3.2.4. Weiter sei $A \in K^{(m,n)}$, dann werde φ_A wie in diesem Satz definiert: für $\vec{v} = \lambda_1\vec{b}_1 + \ldots + \lambda_n\vec{b}_n \in V$ sei $\varphi_A(\vec{v}) = \lambda'_1\vec{b}'_1 + \ldots + \lambda'_m\vec{b}'_m$ mit

$$\begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_m \end{pmatrix} = A \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Dann ist $\varphi_A \in L(V,V')$ und $\Phi(\varphi_A) = \mathcal{M}(\varphi_A,\mathcal{B}',\mathcal{B}) = A$. Gleichzeitig ergibt sich die angegebene Formel für Φ^{-1}. □

Der folgende Satz beschreibt den grundlegenden Zusammenhang zwischen der Komposition von linearen Abbildungen und der Multiplikation von Matrizen:

Satz 3.5.4. *Es seien* V, V', V'' *endlichdimensionale VR mit Basen* $\mathcal{B},\mathcal{B}',\mathcal{B}''$. *Für* $\varphi \in L(V,V')$ *und* $\psi \in L(V',V'')$ *gilt dann:*

$$\mathcal{M}(\psi\circ\varphi,\mathcal{B}'',\mathcal{B}) = \mathcal{M}(\psi,\mathcal{B}'',\mathcal{B}') \cdot \mathcal{M}(\varphi,\mathcal{B}',\mathcal{B})\,.$$

Beweis. Es sei $\dim V = n$, $\dim V' = m$ und $\dim V'' = p$ sowie $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$, $\mathcal{B}' = \{\vec{b}'_1, \ldots, \vec{b}'_m\}$ und $\mathcal{B}'' = \{\vec{b}''_1, \ldots, \vec{b}''_p\}$. Wir definieren die Matrizen A und B durch

$$\mathcal{M}(\varphi, \mathcal{B}', \mathcal{B}) = A = (\alpha_{kl}) \quad , \quad 1 \leq k \leq m,\ 1 \leq l \leq n, \quad \varphi(\vec{b}_l) = \sum_{k=1}^{m} \alpha_{kl} \vec{b}'_k, \quad (l = 1 \ldots m),$$

$$\mathcal{M}(\psi, \mathcal{B}'', \mathcal{B}') = B = (\beta_{kl}) \quad , \quad 1 \leq k \leq p,\ 1 \leq l \leq m, \quad \psi(\vec{b}'_k) = \sum_{j=1}^{p} \beta_{jk} \vec{b}'_j, \quad (k = 1 \ldots m).$$

Dann ist

$$(\psi \circ \varphi)(\vec{b}_l) \quad = \quad \psi(\varphi(\vec{b}_l)) = \psi\left(\sum_{k=1}^{m} \alpha_{kl} \vec{b}'_k\right) = \sum_{k=1}^{m} \alpha_{kl} \psi(\vec{b}'_k) = \sum_{k=1}^{m} \alpha_{kl} \sum_{j=1}^{p} \beta_{jk} \vec{b}''_j = \sum_{j=1}^{p} \left(\sum_{k=1}^{m} \beta_{jk} \alpha_{kl}\right) \vec{b}''_j,$$

also

$$\mathcal{M}(\psi \circ \varphi, \mathcal{B}'', \mathcal{B}) = \left(\sum_{k=1}^{m} \beta_{jk} \alpha_{kl}\right)_{\substack{1 \leq j \leq p \\ 1 \leq l \leq n}} = B \cdot A.$$

□

Satz 3.5.5. *Für Matrizen A, B, C über K gilt, sobald die Ausdrücke definiert sind (d.h. die Spalten- und Zeilenzahl zueinanderpassen) das Assoziativgesetz:*

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

Beweis. Es sei $A \in K^{(m,n)}$, $B \in K^{(n,r)}$ und $C \in K^{(r,s)}$ für $m, n, r, s \in \mathbb{N}$. Es sei $\mathcal{B}_j$ die Standardbasis von $K^{(j)}$. Nach Satz 3.5.3 gibt es lineare Abbildungen $\varphi\colon K^m \to K^n$, $\psi\colon K^n \to K^r$ und $\sigma\colon K^r \to K^s$ mit $\mathcal{M}(\varphi, \mathcal{B}_n, \mathcal{B}_m) = A$, $\mathcal{M}(\psi, \mathcal{B}_r, \mathcal{B}_n) = B$ und $\mathcal{M}(\sigma, \mathcal{B}_s, \mathcal{B}_r) = C$.
Nach Satz 3.5.4 ist $(A \cdot B) \cdot C = \mathcal{M}((\varphi \circ \psi) \circ \sigma, \mathcal{B}_s, \mathcal{B}_m)$ und $A \cdot (B \cdot C) = \mathcal{M}(\varphi \circ (\psi \circ \sigma), \mathcal{B}_s, \mathcal{B}_m)$. Wegen der offensichtlichen Assoziativität der Komposition von Abbildungen ist $(\varphi \circ \psi) \circ \sigma = \varphi \circ (\psi \circ \sigma)$ und daher auch $(A \cdot B) \cdot C = A \cdot (B \cdot C)$. □

Satz 3.5.6. *Es seien A, B, C Matrizen über K. Es gelten die Distributivgesetze*

$$A \cdot (B + C) = A \cdot B + A \cdot C \text{ und } (B + C) \cdot A = B \cdot A + C \cdot A,$$

sobald die Ausdrücke definiert sind.

Beweis. Durch Nachrechnen. □

Die Sätze 3.5.5 und 3.5.6 beinhalten die Verknüpfungsgesetze, die in einem Ring erfüllt sein müssen. Die Abgeschlossenheit kann erreicht werden, wenn man sich auf Matrizen vom Typ (n, n) mit $n \in \mathbb{N}$, sogenannte quadratische Matrizen, beschränkt.
Wir erhalten einen zu $(L(V, V), +, \circ)$ aus Satz 3.1.3 isomorphen Ring von Matrizen, wobei $\dim V = n$ ist. Dabei hat $(L(V, V), +, \circ)$ die Identität als Einselement. Die zur Identität gehörende Matrix ist die Einheitsmatrix E_n.

Definition 3.5.7. Es sei $n \in \mathbb{N}$. Die Matrix $E_n = (\delta_{ij})$ vom Typ (n, n) mit $1 \leq i, j \leq n$ mit dem Kroneckersymbol

$$\delta_{ij} = \begin{cases} 1, & \text{falls} \quad i = j \\ 0, & \text{falls} \quad i \neq j \end{cases}$$

heißt Einheitsmatrix vom Typ (n, n).

Satz 3.5.7. *Es sei* $\dim V = n$ *und* $\mathcal{B}$ *eine Basis von* V*. Die bijektive Abbildung*

$$\Phi\colon L(V,V) \to K^{(n,n)}, \quad \varphi \to \mathcal{M}(\varphi,\mathcal{B},\mathcal{B})$$

ist ein *Ringisomorphismus, d.h. es gilt*

$$\Phi(\varphi+\psi) = \Phi(\varphi) + \Phi(\psi) \quad \textit{und} \quad \Phi(\varphi\circ\psi) = \Phi(\varphi)\cdot\Phi(\psi).$$

Die Ringe $(L(V,V),+,\circ)$ *bzw.* $(K^{(n,n)},+,\cdot)$ *sind isomorphe Ringe mit Einselementen id bzw. mit* $\Phi(id) = E_n$*.*

Nachdem wir somit Vektorräume bzw. Ringe von Matrizen gefunden haben, die zu den entsprechenden Vektorräumen bzw. Ringen von linearen Abbildungen isomorph sind, suchen wir nun noch die Gruppe der Matrizen, die zu der linearen Gruppe $GL(V)$ der Automorphismen von V aus Satz 3.1.4 isomorph sind.
Dazu müssen die Matrizen charakterisiert werden, die zu Automorphismen gehören. Dies geschieht, indem wir das Konzept des Ranges einer linearen Abbildung auf Matrizen übertragen.

Definition 3.5.8. Die Maximalzahl linear unabhängiger Zeilenvektoren von $A \in K^{(m,n)}$ heißt der Zeilenrang von A.
Die Maximalzahl linear unabhängiger Spaltenvektoren von $A \in K^{(m,n)}$ heißt der Spaltenrang, oder kurz Rang der Matrix A (Schreibweise: $\mathrm{rg}(A)$).

Satz 3.5.8. *Es seien* V, V' *endlichdimensionale VR mit Basen* $\mathcal{B},\mathcal{B}'$ *und* $\varphi \in L(V,V')$*. Dann gilt*

$$\mathrm{rg}(\varphi) = \mathrm{rg}(\mathcal{M}(\varphi,\mathcal{B}',\mathcal{B})).$$

Insbesondere hängt der Rang von $\mathcal{M}(\varphi,\mathcal{B}',\mathcal{B})$ *also nicht von der Wahl der Basen* $\mathcal{B},\mathcal{B}'$ *ab.*

Beweis. Es sei $\mathcal{B} = \{\vec{b}_1,\ldots,\vec{b}_n\}$ und $\mathcal{B}' = \{\vec{b}'_1,\ldots,\vec{b}'_m\}$ sowie $\mathcal{M}(\varphi,\mathcal{B}',\mathcal{B}) = A = (\alpha_{kl})$ mit $1 \le k \le m$ bzw. $1 \le l \le n$. Für $l = 1\ldots n$ ist

$$\varphi(\vec{b}_l) = \sum_{k=1}^{m} \alpha_{kl}\vec{b}'_k,$$

d.h. der l-te Spaltenvektor

$$\vec{a}_l = \begin{pmatrix} \alpha_{1l} \\ \vdots \\ \alpha_{ml} \end{pmatrix}$$

von A, der Koordinatenvektor von $\varphi(\vec{b}_l)$ bzgl. der Basis $\mathcal{B}'$. Wir zeigen für $1 \le l_1 < l_2 < \ldots < l_r \le n$ die Äquivalenz

$$\varphi(\vec{b}_{l_1}),\ldots,\varphi(\vec{b}_{l_r}) \quad \text{l.u.} \Leftrightarrow \vec{a}_{l_1},\ldots,\vec{a}_{l_r} \quad \text{l.u..} \tag{$*$}$$

Es gilt

$$\vec{0} = \sum_{j=1}^{r} \lambda_j \varphi(\vec{b}_{l_j}) = \sum_{j=1}^{r} \lambda_j \left(\sum_{k=1}^{m} \alpha_{kl_j} \vec{b}'_k \right) \quad \Leftrightarrow \quad \sum_{j=1}^{r} \alpha_{kl_j}\lambda_j = 0 \quad \text{für } k = 1,\ldots,m \quad (\text{da } \vec{b}'_1,\ldots,\vec{b}'_m \text{ l.u.})$$

$$\Leftrightarrow \quad \lambda_1\vec{a}_{l_1} + \ldots + \lambda_r\vec{a}_{l_r} = \vec{0}.$$

Damit ist $(*)$ bewiesen. Es folgt

$$\begin{aligned} \mathrm{rg}(\varphi) &= \dim\varphi(V) = \dim\left\langle\, \varphi(\vec{b}_1),\ldots,\varphi(\vec{b}_n) \,\right\rangle \\ &= \text{Maximalzahl von l.u. Vektoren unter den } \varphi(\vec{b}_1),\ldots,\varphi(\vec{b}_n) \\ &= \text{Maximalzahl von l.u. Vektoren unter den } \vec{a}_1,\ldots,\vec{a}_n \\ &= \text{Spaltenrang von } A = \mathrm{rg}(A). \end{aligned}$$

□

Satz 3.5.9. *Für Matrizen* $A \in K^{(m,n)}$ *und* $B \in K^{(n,p)}$ *gilt*

$$\mathrm{rg}(A) + \mathrm{rg}(B) - n \leq \mathrm{rg}(AB) \leq \min\{\mathrm{rg}(A), \mathrm{rg}(B)\}.$$

Beweis. Wähle Vektorräume V, V', V'' mit Dimensionen p, n, m und Basen $\mathcal{B}, \mathcal{B}', \mathcal{B}''$. Nach Satz 3.3.1 gibt es $\varphi \in L(V, V')$ und $\psi \in L(V', V'')$ mit $\mathcal{M}(\varphi, \mathcal{B}', \mathcal{B}) = B$ und $\mathcal{M}(\psi, \mathcal{B}'', \mathcal{B}') = A$, also gilt nach Satz 3.5.4 $\mathcal{M}(\psi \circ \varphi, \mathcal{B}'', \mathcal{B}) = A \cdot B$. Nach Satz 3.5.8 ist $\mathrm{rg}(A) = \mathrm{rg}(\psi)$, $\mathrm{rg}(B) = \mathrm{rg}(\varphi)$ und $\mathrm{rg}(AB) = \mathrm{rg}(\psi \circ \varphi)$, so dass die Behauptung aus Satz 3.3.3 folgt. □

Definition 3.5.9. Es sei $n \in \mathbb{N}$. Die Matrix $A \in K^{(n,n)}$ heißt regulär, wenn $rg(A) = n$ gilt. Die Menge aller regulären Matrizen von $K^{(n,n)}$ heißt $\mathrm{GL}(n, K)$. Andernfalls ($\mathrm{rg}(A) < n$) heißt A singulär.

Satz 3.5.10. *Es sei* $\dim V = n$, $\mathcal{B}$ *eine Basis von* V *und* $\varphi \in L(V, V)$. *Dann ist genau dann* $\varphi \in \mathrm{GL}(V)$, *wenn* $\mathcal{M}(\varphi, \mathcal{B}, \mathcal{B}) \in \mathrm{GL}(n, K)$ *ist.*

Beweis. Nach Satz 3.2.6 ist $\varphi \in GL(V) \Leftrightarrow rg(\varphi) = \dim V$.
Nach Satz 3.5.8 ist $rg(\varphi) = rg(\mathcal{M}(\varphi, \mathcal{B}, \mathcal{B}))$. □

Durch Einschränkung des Ringisomorphismus $\Phi\colon L(V; V) \to K^{(n,n)}$, $\varphi \to \mathcal{M}(\varphi, \mathcal{B}, \mathcal{B})$ auf $\mathrm{GL}(V)$ erhält man einen Gruppenisomorphismus $\psi\colon (\mathrm{GL}(V), \circ) \to (\mathrm{GL}(n, K), \cdot)$.

Satz 3.5.11. *Es sei* $\dim V = n$ *und* $\mathcal{B}$ *eine Basis von* V.
Die Abbildung $\Psi\colon \mathrm{GL}(V) \to \mathrm{GL}(n, K)$, $\varphi \to \mathcal{M}(\varphi, \mathcal{B}, \mathcal{B})$ *ist ein Gruppenisomorphismus.*
Die Gruppen $(\mathrm{GL}(V), \circ)$ *bzw.* $(\mathrm{GL}(n, K), \cdot)$ *sind isomorphe Gruppen mit Einselementen id bzw. mit* $\Psi(id) = E_n$.

Beweis. Zum Nachweis der Relationstreue von Ψ und der Gruppeneigenschaft von $\mathrm{GL}(n, K)$ benützen wir den Ringisomorphismus Φ von Satz 3.5.7, dessen Einschränkung auf $\mathrm{GL}(V)$ gerade Ψ ist.

1. Die Relationstreue $\Psi(\varphi \circ \psi) = \Psi(\varphi) \circ \Psi(\psi)$ gilt, da sie für Φ gilt.
   Für $A, B \in \mathrm{GL}(n, K)$ sind nach Satz 3.5.10 $\varphi := \Phi^{-1}(A) \in \mathrm{GL}(V)$ und $\psi := \Phi^{-1}(B) \in \mathrm{GL}(V)$.

2. Abgeschlossenheit von $\mathrm{GL}(n, K)$ bzgl. der Multiplikation:
   Wegen der Gruppeneigenschaft von $\mathrm{GL}(V)$ nach Satz 3.1.4 ist $\varphi \circ \psi \in \mathrm{GL}(V)$, und damit gilt wiederum nach Satz 3.5.10, dass $A \cdot B = \Phi(\varphi \circ \psi) \in \mathrm{GL}(n, K)$ gilt.

3. Neutrales Element E_n:
   Nach Satz 3.5.7 ist E_n das Einslement des Rings $K^{(n,n)}$. Also gilt für alle $A \in K^{(n,n)}$ und damit erst recht für alle $A \in \mathrm{GL}(V)$, dass $A \cdot E_n = E_n \cdot A = A$ gilt.

4. Existenz des Inversen:
   Wegen der Gruppeneigenschaft von $\mathrm{GL}(V)$ existiert das Inverse $\varphi^{-1} \in \mathrm{GL}(V)$ von φ, d.h. es ist $\varphi \circ \varphi^{-1} = \varphi^{-1} \circ \varphi = id$.
   Dann ist

   $$E_n = \Phi(id) = \Phi(\varphi) \cdot \Phi(\varphi^{-1}) = \Phi(\varphi^{-1}) \cdot \Phi(\varphi) = A \cdot \Phi(\varphi^{-1}) = \Phi(\varphi^{-1}) \cdot A.$$

   Also ist $\Phi(\varphi^{-1}) = A^{-1}$.

□

Definition 3.5.10. Eine quadratische Matrix $A \in K^{(n,n)}$ heißt invertierbar, wenn es eine Matrix $B \in K^{(n,n)}$ mit $B \cdot A = E_n$ oder $A \cdot B = E_n$ gibt.

Satz 3.5.12. *Eine Matrix $A \in K^{(n,n)}$ ist genau dann invertierbar, wenn A regulär ist. Dann besitzt A genau ein Inverses A^{-1} mit $A \cdot A^{-1} = A^{-1} \cdot A = E_n$. Die Matrix A^{-1} ist auch das einzige Links- und das einzige Rechtsinverse von A.*

Beweis. ”$\Leftarrow$”:
Es sei A regulär. Die Aussage folgt aus der Gruppeneigenschaft von $GL(n, K)$.
$\Rightarrow$”:
Es sei A invertierbar. Es existiere $B \in K^{(n,n)}$ mit $B \cdot A = E_n$. Nach Satz 3.5.9 ist

$$n = \mathrm{rg}(E_n) = \mathrm{rg}(B \cdot A) \leq \min\{rg(A), rg(B)\}.$$

Also ist $\mathrm{rg}(A) = \mathrm{rg}(B) = n$ und damit $A, B, \in \mathrm{GL}(n, K)$ und $B = A^{-1}$.
Entsprechendes gilt, falls $A \cdot B = E_n$ für $B \in K^{(n,n)}$. □

Satz 3.5.13. *Für eine beliebige (nicht notwendigerweise quadratische) Matrix A gilt $A \cdot E_n = A$ und $E_n \cdot A = A$, falls diese Produkte definiert sind.*

Beweis. Durch Nachrechnen. □

3.6 Berechnung des Rangs einer Matrix

Die Berechnung des Rangs und der Inversen einer Matrix beruht auf elementaren Matrixtransformationen. Diese bestehen aus elementaren Zeilenumformungen und elementaren Spaltenumformungen. Die elementaren Zeilenumformungen sind uns schon in Abschnitt 1.4 bei der Lösung linearer Gleichungssysteme mittels des Gaußschen Algorithmus begegnet. Die Schritte des Gaußschen Algorithmus bestehen aus den elementaren Zeilenumformungen und einer der elementaren Spaltenumformungen, nämlich der Vertauschung zweier Spalten.

Definition 3.6.1. Als elementare Umformungen einer Matrix $A \in K^{(m,n)}$ bezeichnet man:

- Elementare Zeilenumformungen:
  1. Multiplikation einer Zeile mit einem Skalar $\lambda \neq 0$ aus K,
  2. Addition einer mit $\lambda \in K$ multiplizierten Zeile zu einer anderen Zeile,
  3. Vertauschung zweier Zeilen.
- Elementare Spaltenumformungen:
  1'. Multiplikation einer Spalte mit einem Skalar $\lambda \neq 0$ aus K,
  2'. Addition einer mit $\lambda \in K$ multiplizierten Spalte zu einer anderen Spalte,
  3'. Vertauschung zweier Spalten.

Satz 3.6.1. *Durch elementare Umformungen lässt sich jede Matrix $A = (a_{ij}) \in K^{(m,n)}$ in eine Matrix $D_r^{(m,n)}$ der Form*

$$D_r^{(m,n)} = \left(\begin{array}{ccc|cc} 1 & & & & \\ & 1 & & & 0 \\ & & \ddots & & \\ 0 & & 1 & & \\ \hline & & & 0 & \\ & 0 & & & 0 \\ & & & & \ddots \end{array}\right) = (c_{ij}) \quad \textit{mit} \quad c_{ij} = \begin{cases} 1, & \textit{falls } i = j \leq r \\ 0 & \textit{sonst} \end{cases}$$

überführen.

Beweis. Wie in Abschnitt 1.4 zeigt man, dass A durch elementare Zeilenumformungen und möglicherweise Vertauschungen von Spalten in die Form

$$\left(\begin{array}{ccc|ccc} 1 & & & & & \\ & 1 & & \alpha_{1,r+1} & \cdots & \alpha_{1,n} \\ & & \ddots & \vdots & & \vdots \\ 0 & & 1 & \alpha_{r,r+1} & \cdots & \alpha_{r,n} \\ \hline & & & 0 & & \\ & 0 & & & 0 & \\ & & & & & \ddots \end{array}\right)$$

gebracht werden kann.Der einzige Unterschied besteht darin, dass die Elemente jetzt dem Körper K, nicht mehr notwendigerweise dem Körper $\mathbb{R}$ angehören. Die Teilmatrix

$$\begin{pmatrix} \alpha_{1,r+1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{r,r+1} & \cdots & \alpha_{r,n} \end{pmatrix}$$

kann dann durch Addition von passenden Vielfachen der ersten r Spalten in die Matrix $0^{(r,n-r)}$ umgeformt werden. $\square$

Satz 3.6.2. *Elementare Umformungen verändern weder den Zeilen- noch den Spaltenrang einer Matrix.*

Beweis. Wegen der Analogie zwischen Zeilen und Spalten genügt es zu zeigen, dass elementare Zeilenumformungen weder den Zeilen- noch den Spaltenrang einer Matrix verändern. Wir beschränken uns auf die Operation 2, der Beweis für die anderen Operationen verläuft analog. Wir addieren O.B.d.A. ein Vielfaches der ersten auf die zweite Zeile:

$$A = \begin{pmatrix} \vec{a}_1 \\ \vec{a}_2 \\ \vdots \\ \vec{a}_m \end{pmatrix} \overset{2}{\to} A' = \begin{pmatrix} \vec{a}_1 \\ \vec{a}_2 + \lambda\vec{a}_1 \\ \vdots \\ \vec{a}_m \end{pmatrix}.$$

Dann gilt $\langle\, \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m \,\rangle = \langle\, \vec{a}_1, \vec{a}_2 + \lambda\vec{a}_1, \ldots, \vec{a}_m \,\rangle$. Ist nämlich $\vec{v} = \mu_1\vec{a}_1 + \mu_2\vec{a}_2 + \cdots + \mu_m\vec{a}_m$, so gilt auch

$$\vec{v} = (\mu_1 - \mu_2\lambda)\vec{a}_1 + \mu_2(\vec{a}_2 + \lambda\vec{a}_1) + \mu_3\vec{a}_3 + \cdots + \mu_m\vec{a}_m.$$

Andererseits gilt für $\vec{w} = \nu_1\vec{a}_1 + \nu_2(\vec{a}_2 + \lambda\vec{a}_1) + \cdots + \nu_m\vec{a}_m$ auch

$$\vec{w} = (\nu_1 + \lambda\nu_2)\vec{a}_1 + \nu_2\vec{a}_2 + \cdots + \nu_m\vec{a}_m.$$

Also haben A und A' nicht nur gleichen Zeilenrang, ihre Zeilenvektoren spannen sogar denselben Unterraum von K^n auf. Sind

$$\vec{b}_{l_1} = \begin{pmatrix} a_{1,l_1} \\ a_{2,l_1} \\ \vdots \\ a_{m,l_1} \end{pmatrix}, \ \ldots, \ \vec{b}_{l,s} = \begin{pmatrix} a_{1,l_s} \\ a_{2,l_s} \\ \vdots \\ a_{m,l_s} \end{pmatrix}$$

beliebige Spaltenvektoren von A, dann sind die entsprechenden Spalten in A'

$$\vec{b}'_{l_1} = \begin{pmatrix} a_{1,l_1} \\ a_{2,l_1} + \lambda a_{1,l_1} \\ \vdots \\ a_{m,l_1} \end{pmatrix}, \ \ldots, \ \vec{b}'_{l,s} = \begin{pmatrix} a_{1,l_s} \\ a_{2,l_s} + \lambda a_{1,l_s} \\ \vdots \\ a_{m,l_s} \end{pmatrix}.$$

Jede lineare Relation

$$\mu_1 \vec{b}_{l_1} + \cdots + \mu_s \vec{b}_{l_s} = \vec{0}$$

ist zur entsprechenden Relation

$$\mu_1 \vec{b}'_{l_1} + \cdots + \mu_s \vec{b}'_{l_s} = \vec{0},$$

äquivalent, die $\vec{b}_l$ sind also genau dann linear unabhängig, wenn die entsprechenden $\vec{b}'_l$ es sind. Daher haben A und A' auch den gleichen Spaltenrang. □

Satz 3.6.3. *Für jede Matrix $A \in K^{(m,n)}$ gilt:*

$$\textit{Zeilenrang von } A = \textit{Spaltenrang von } A = \textit{Rang von } A.$$

Es ist $\mathrm{rg}(A) = r$ die Anzahl der Einsen aus der Matrix $D_r^{(m,n)}$ aus Satz 3.6.1.

Bemerkung 3.6.1. Die Zahl r hängt also nur von A ab, und nicht davon, mit welcher Serie elementarer Umformungen die Matrix $D_r^{(m,n)}$ gewonnen wurde.

Beweis von Satz 3.6.3. Nach Satz 3.6.2 hat $D_r^{(m,n)}$ denselben Zeilen- bzw. Spaltenrang wie A. Der Zeilen- sowie der Spaltenrang von $D_r^{(m,n)}$ ist aber offensichtlich r. □

3.7 Basiswechsel

Wir untersuchen nun, wie sich die Koordinaten eines Vektors $\vec{v} \in V$ beim Wechsel der Basis des Vektorraums V ändern. Außerdem untersuchen wir die damit eng verwandte Frage, wie sich die zu einer Abbildung $\varphi \in L(V, V')$ bzgl. der Basen $\mathcal{B}$, $\mathcal{B}'$ gehörende Matrix $\mathcal{M}(\varphi, \mathcal{B}', \mathcal{B})$ ändert, wenn die Basen $\mathcal{B}$, $\mathcal{B}'$ durch andere Basen ersetzt werden. Zunächst beweisen wir einen Satz über die Gesamtheit aller Basen eines Vektorraums:

Satz 3.7.1. *Es sei $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ eine Basis von V und $\varphi \in L(V, V)$. Es ist $\tilde{\mathcal{B}} = \{\varphi(\vec{b}_1), \ldots, \varphi(\vec{b}_n)\}$ genau dann eine Basis von V, wenn φ ein Automorphismus ist.*

Beweis. Es gilt

$$\tilde{\mathcal{B}} \text{ ist eine Basis} \Leftrightarrow \varphi(\vec{b}_1), \ldots, \varphi(\vec{b}_n) \text{ sind l.u.} \Leftrightarrow \mathrm{rg}(\varphi) = n \Leftrightarrow \varphi \text{ ist ein Automorphismus.}$$

□

Bemerkung 3.7.1. Es besteht also eine umkehrbar eindeutige Entsprechung zwischen

1. den Automorphismen von V,

2. den Basistransformationen von V und

3. den regulären $(n \times n)$-Matrizen über K.

Satz 3.7.2. *Es seien* $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ *und* $\mathcal{B}' = \{\vec{b}'_1, \ldots, \vec{b}'_n\}$ *Basen von* V *und* $\varphi \in \mathrm{GL}(V)$ *ein Automorphismus. Dann gilt* $\mathcal{M}(\varphi, \mathcal{B}', \mathcal{B}) = \mathcal{M}(\mathrm{id}_V, \mathcal{B}', \varphi(\mathcal{B}))$ *mit* $\varphi(\mathcal{B}) = \{\varphi(\vec{b}_1), \ldots, \varphi(\vec{b}_n)\}$*.*

Beweis. Die Menge $\varphi(\mathcal{B})$ ist nach Satz 3.7.1 eine Basis von V. Beide Matrizen sind gleich (α_{kl}) mit $1 \leq k \leq n$ bzw. $1 \leq l \leq n$ und

$$\varphi(\vec{b}_l) = \mathrm{id}_V(\varphi(\vec{b}_l)) = \sum_{k=1}^{n} \alpha_{kl} \vec{b}'_k$$

für $l = 1, \ldots, n$. □

Satz 3.7.3. *Es seien* $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ *und* $\varphi(\mathcal{B}) = \{\varphi(\vec{b}_1), \ldots, \varphi(\vec{b}_n)\}$ *Basen von* V *für* $\varphi \in \mathrm{GL}(V)$*. Sind* $\lambda_1, \ldots, \lambda_n$ *bzw.* $\lambda'_1, \ldots, \lambda'_n$ *die Koordinaten von* $\vec{v}$ *bzgl.* $\mathcal{B}$ *bzw. bzgl.* $\varphi(\mathcal{B})$*, d.h. gilt*

$$\vec{v} = \lambda_1 \vec{b}_1 + \cdots + \lambda_n \vec{b}_n \quad \textit{bzw.} \quad \vec{v} = \lambda'_1 \varphi(\vec{b}_1) + \cdots + \lambda'_n \varphi(\vec{b}_n),$$

so ist

$$\begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_n \end{pmatrix} = \mathcal{M}(\varphi, \mathcal{B}, \mathcal{B})^{-1} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Beweis. Nach Satz 3.5.2 gilt

$$\begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_n \end{pmatrix} = \mathcal{M}(\mathrm{id}_V, \varphi(\mathcal{B}), \mathcal{B}) \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}. \tag{$*$}$$

Nach Satz 3.5.4 ist andererseits

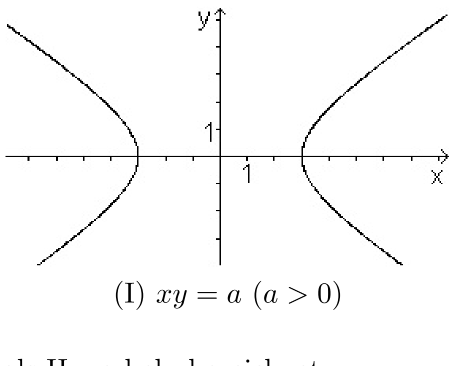
$$E_n = \mathcal{M}(\mathrm{id}_V, \mathcal{B}, \mathcal{B}) = \mathcal{M}(\mathrm{id}_V, \mathcal{B}, \varphi(\mathcal{B})) \cdot \mathcal{M}(\mathrm{id}_V, \varphi(\mathcal{B}), \mathcal{B}),$$

also $\mathcal{M}(\mathrm{id}_V, \varphi(\mathcal{B}), \mathcal{B}) = \mathcal{M}(\mathrm{id}_V, \mathcal{B}, \varphi(\mathcal{B}))^{-1}$ und daher nach Satz 3.7.2

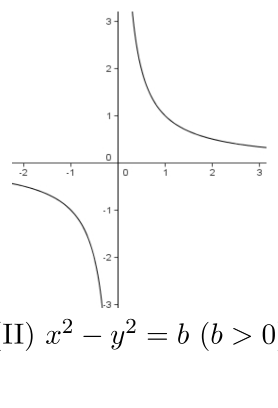
$$\mathcal{M}(\mathrm{id}_V, \varphi(\mathcal{B}), \mathcal{B}) = \mathcal{M}(\varphi, \mathcal{B}, \mathcal{B})^{-1}.$$

Daraus und aus $(*)$ folgt die Behauptung. □

Beispiel 3.7.1. Im Geometrieunterricht begegnet man der Hyperbel in zwei Formen als Kurven. Die Kurven in der xy-Ebene mit den Gleichungen



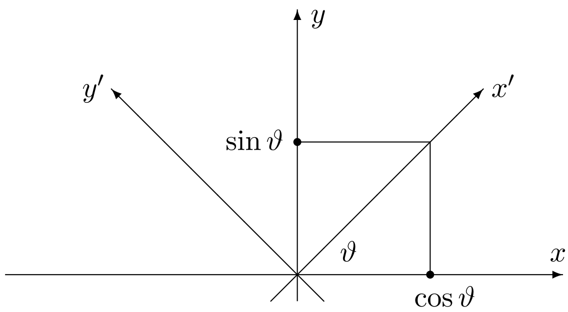
(I) $xy = a$ $(a > 0)$



(II) $x^2 - y^2 = b$ $(b > 0)$

werden jeweils als Hyperbeln bezeichnet.

Woher wissen wir, ob durch die Gleichungstypen (I) und (II) kongruente Kurven beschrieben werden? Die Skizzen legen die Vermutung nahe, dass Kurven vom Typ (II) durch eine Drehung um 45° (im Gegenuhrzeigersinn) in Kurven des Typs (I) übergeführt werden. Zum Beweis dieser Vermutung drehen wir das xy-Koordinatensystem um 45° in das $x'y'$-Koordinatensystem und untersuchen die Gleichung der Kurve $xy = a$ in $x'y'$-Koordinaten. Zunächst leiten wir die Transformationsgleichungen für eine Drehung des Koordinatensystems um einen Winkel ϑ im Uhrzeigersinn her.



Die x'-Achse wird vom Vektor $\vec{b}_1 = (\cos\vartheta, \sin\vartheta)$ bzw. die y'-Achse von $\vec{b}_2 = (-\sin\vartheta, \cos\vartheta)$ aufgespannt. Während die xy-Koordinaten gerade die Koeffizienten $\vec{v} = (x, y) \in \mathbb{R}^2$ in der Darstellung $\vec{v} = (x, y) = x\vec{e}_1 + y\vec{e}_2$ sind, sind die $x'y'$-Koordinaten von $\vec{v}$ die Koeffizienten x' und y' in der Darstellung $\vec{v} = x'\vec{b}_1 + y'\vec{b}_2$. Die Matrix der Drehung $\varphi \in L'(\mathbb{R}^2, \mathbb{R}^2)$, die $B = (\vec{e}_1, \vec{e}_2)$ in $B' = (\vec{b}_1, \vec{b}_2)$ überführt, ist

$$\mathcal{M}(\varphi, \mathcal{B}', \mathcal{B}) = \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix},$$

daher ist nach Satz 3.7.3

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix} \cdot \begin{pmatrix} x' \\ y' \end{pmatrix}$$

oder

$$\begin{cases} x = (\cos\vartheta)x' - (\sin\vartheta)y' \\ y = (\sin\vartheta)x' + (\cos\vartheta)y' \end{cases}$$

Für $\vartheta = \frac{\pi}{4}$ (=45°) erhalten wir $\sin\vartheta = \cos\vartheta = \frac{1}{\sqrt{2}}$, daher ist $xy = a$ zu $\frac{1}{2}(x' - y')(x' + y') = a$ oder $x'^2 - y'^2 = 2a$ äquivalent. Die Kurve $xy = a$ geht daher aus der Kurve $x^2 - y^2 = 2a$ durch eine Drehung um 45° hervor. Die Gleichungen (I) und (II) stellen in der Tat kongruente Kurven dar, falls $b = 2a$ ist.

Satz 3.7.4. *Es seien* $\dim(V), \dim(V') < \infty$ *und* $\mathcal{B}$ *bzw.* $\mathcal{B}'$ *Basen von* V *bzw.* V'*, sowie* $\varrho \in \mathrm{GL}(V)$ *bzw.* $\sigma \in \mathrm{GL}(V')$*. Sei* $\mathcal{X} = \mathcal{M}(\varrho, \mathcal{B}, \mathcal{B})$ *und* $\mathcal{Y} = \mathcal{M}(\sigma, \mathcal{B}', \mathcal{B}')$*. Dann gilt für alle* $\varphi \in L(V, V')$*:*

$$\mathcal{M}(\varphi, \sigma(\mathcal{B}'), \varrho(\mathcal{B})) = \mathcal{Y}^{-1} \cdot \mathcal{M}(\varphi, \mathcal{B}', \mathcal{B}) \cdot \mathcal{X}.$$

Beweis. Nach Satz 3.5.4 gilt

$$\begin{aligned} \mathcal{M}(\varphi, \sigma(\mathcal{B}'), \varrho(\mathcal{B})) &= \mathcal{M}(\mathrm{id}_{V'} \circ \varphi \circ \mathrm{id}_V, \sigma(\mathcal{B}'), \varrho(\mathcal{B})) \\ &= \mathcal{M}(\mathrm{id}_{V'} \circ \varphi, \sigma(\mathcal{B}'), \mathcal{B}) \cdot \mathcal{M}(\mathrm{id}_V, \mathcal{B}, \varrho(\mathcal{B})) \\ &= \mathcal{M}(\mathrm{id}_{V'}, \sigma(\mathcal{B}'), \mathcal{B}') \cdot \mathcal{M}(\varphi, \mathcal{B}', \mathcal{B}) \cdot \mathcal{M}(\mathrm{id}_V, \mathcal{B}, \varrho(\mathcal{B})). \end{aligned}$$

Nach Satz 3.7.2 folgt

$$\mathcal{M}(\mathrm{id}_{V'}, \sigma(\mathcal{B}'), \mathcal{B}') = \mathcal{M}(\mathrm{id}_{V'}, \mathcal{B}', \sigma(\mathcal{B}'))^{-1} = \mathcal{M}(\sigma, \mathcal{B}', \mathcal{B}')^{-1} = \mathcal{Y}^{-1}$$

und ebenso

$$\mathcal{M}(\mathrm{id}_V, \mathcal{B}, \varrho(\mathcal{B})) = \mathcal{M}(\varrho, \mathcal{B}, \mathcal{B}) = \mathcal{X}.$$

Damit ergibt sich die Behauptung. □

Wir erhalten als Spezialfall

Satz 3.7.5. *(Basiswechsel für lineare Abbildungen)*
Es sei $\dim(V) < \infty$, $\sigma \in \mathrm{GL}(V)$ *und* $\mathcal{B}$ *eine Basis von* V. *Dann gilt für alle* $\varphi \in L(V, V)$*:*

$$\mathcal{M}(\varphi, \sigma(\mathcal{B}), \sigma(\mathcal{B})) = \mathcal{X}^{-1} \cdot \mathcal{M}(\varphi, \mathcal{B}, \mathcal{B}) \cdot \mathcal{X}$$

mit $\mathcal{X} = \mathcal{M}(\sigma, \mathcal{B}, \mathcal{B})$.

Beispiel 3.7.2. Im Beispiel 3.5.6 betrachteten wir die lineare Abbildung $\varphi \in L(\mathbb{R}^3, \mathbb{R}^3)$ mit bzgl. $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ zugeordneter Matrix

$$\mathcal{M}(\varphi, \mathcal{B}, \mathcal{B}) = \begin{pmatrix} -3 & -2 & -4 \\ -3 & 0 & -3 \\ 8 & 4 & 9 \end{pmatrix}.$$

Wir betrachteten die neue Basis $\tilde{\mathcal{B}} = \{\vec{b}_1, \vec{b}_2, \vec{b}_3\}$ mit

$$\vec{b}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \vec{b}_2 = \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix} \quad \text{und} \quad \vec{b}_3 = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}.$$

Es ist $\tilde{\mathcal{B}} = \sigma(\mathcal{B})$ mit

$$\mathcal{X} = \mathcal{M}(\sigma, \mathcal{B}, \mathcal{B}) = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & -1 \\ -1 & -4 & 2 \end{pmatrix}.$$

Außerdem war

$$\mathcal{M}(\varphi, \tilde{\mathcal{B}}, \tilde{\mathcal{B}}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Satz 3.7.5 sagt uns daher, dass

$$\mathcal{X}^{-1} \cdot \begin{pmatrix} -3 & -2 & -4 \\ -3 & 0 & -3 \\ 8 & 4 & 9 \end{pmatrix} \cdot \mathcal{X} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

ist, was man durch Nachrechnen bestätigt. Wir haben die Matrix

$$\begin{pmatrix} -3 & -2 & -4 \\ -3 & 0 & -3 \\ 8 & 4 & 9 \end{pmatrix}$$

diagonalisiert. Im Kapitel über Eigenwerte werden wir das Verfahren zur Diagonalisierung beschreiben.

Ein wichtiger Spezialfall von $L(V, V')$ ist

Definition 3.7.1. Der Dualraum eines K-Vektorraums V ist $V^* = L(V, K)$, die $\varphi \in V^*$ nennt man Linearformen (im Fall $K = \mathbb{R}, \mathbb{C}$ auch lineare Funktionale).

Satz 3.7.6. *Es sei* $\dim(V) = n < \infty$, *dann ist* $V^* \cong V$. *Ist* $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ *eine Basis von* V, *so gibt es zu jedem* $\varphi \in V^*$ *eindeutig bestimmte* $\alpha_1, \dots, \alpha_n \in K$ *mit*

$$\varphi(\vec{v}) = \alpha_1\lambda_1 + \cdots + \alpha_n\lambda_n$$

für $\vec{v} = \lambda_1\vec{b}_1 + \cdots + \lambda_n\vec{b}_n$. *Die Abbildung* $\Psi\colon V^* \to K^n$, $\varphi \to (\alpha_1, \dots, \alpha_n)$ *ist ein Isomorphismus.*

Beweis. Nach Satz 3.5.3 ist $\dim(V^*) = n \cdot 1 = \dim(V)$, also gilt $V^* \cong V$ nach Satz 3.3.2. Bzgl. der Basis $\{1\}$ von K ist $\mathcal{M}(\varphi, \{1\}, \mathcal{B}) = (\alpha_1, \dots, \alpha_n) \in K^{(1\times n)}$, also nach Satz 3.5.2

$$\varphi(\vec{v}) = (\alpha_1, \dots, \alpha_n) \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \cdot 1$$

für $\vec{v} = \lambda_1\vec{b}_1 + \cdots + \lambda_n\vec{b}_n$. Die Linearität und Bijektivität der Abbildung Ψ prüft man leicht durch Nachrechnen. $\square$

Kapitel 4

Lineare Gleichungen

4.1 Theorie der Linearen Gleichungen

Das Verfahren des Gaußschen Algorithmus zur Lösung von Linearen Gleichungssystemen wurde für den Spezialfall des Körpers $\mathbb{R}$ der reellen Zahlen schon in Abschnitt 1.4 behandelt. Es lässt sich unmittelbar auf den Fall eines allgemeinen Körpers K übertragen.
In diesem Kapitel wenden wir uns der theoretischen Beschreibung der Lösungsmenge zu. Im folgenden sei K stets als Körper vorausgesetzt.

Wir verallgemeinern zunächst Definition 1.4.1:

Definition 4.1.1. Es sei $\mathcal{A} = (\alpha_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}} \in K^{(m,n)}$ sowie

$$\vec{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix} \in K^m \quad \text{und} \quad \vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n.$$

Dann heißt

$$\mathcal{A}\vec{x} = \vec{\beta} \tag{$*$}$$

ein lineares Gleichungssystem (LGS) in den Unbestimmten $x_1, \ldots, x_n$ und den Koeffizienten $\alpha_{ij} \in K$. Weiter heißt $\mathcal{A}$ die Koeffizientenmatrix und $\vec{\beta}$ die rechte Seite des Systems.

Fügt man zu $\mathcal{A}$ als $(n+1)$-te Spalte $\vec{\beta}$ hinzu, so erhält man die erweiterte Koeffizientenmatrix $(\mathcal{A}|\vec{\beta}) \in K^{(m,n+1)}$ des LGS $(*)$. Jedes $\vec{x} \in K^n$, für das $(*)$ gilt, heißt eine Lösung des LGS. Gibt es ein solches $\vec{x}$, so heißt $(*)$ lösbar, ansonsten unlösbar. Die Menge aller Lösungen heißt Lösungsmenge des Systems.
Ist $\vec{\beta} = \vec{0}$, so heißt das LGS homogen, ansonsten inhomogen. Ein homogenes LGS besitzt immer die triviale Lösung $\vec{x} = \vec{0}$.

Es erheben sich folgende Fragen:

1. Unter welchen Bedingungen an $\mathcal{A}$ und $\vec{\beta}$ ist $(*)$ lösbar?

2. Wie sieht die Lösungsmenge von $(*)$ aus?

Satz 4.1.1. *Es gilt:* $\mathcal{A}\vec{x} = \vec{\beta}$ *ist genau dann lösbar, wenn* $\operatorname{rg}(\mathcal{A}|\vec{\beta}) = \operatorname{rg}(\mathcal{A})$ *gilt.*

Beweis. Das System $\mathcal{A}\vec{x} = \vec{\beta}$ ist genau dann lösbar, wenn der Vektor $\vec{\beta}$ eine Linearkombination der Spalten $\vec{a}_1, \dots, \vec{a}_n$ von $\mathcal{A}$ ist und wenn

$$\begin{aligned}\vec{\beta} \in \langle \vec{a}_1, \dots, \vec{a}_n \rangle \quad &\Leftrightarrow \quad \left\langle \vec{a}_1, \dots, \vec{a}_n, \vec{\beta} \right\rangle = \langle \vec{a}_1, \dots, \vec{a}_n \rangle \Leftrightarrow \dim \left\langle \vec{a}_1, \dots, \vec{a}_n, \vec{\beta} \right\rangle = \dim \langle \vec{a}_1, \dots, \vec{a}_n \rangle \\ &\Leftrightarrow \quad \text{Spaltenrang von } (\mathcal{A}|\vec{\beta}) = \text{Spaltenrang von } \mathcal{A} \Leftrightarrow \operatorname{rg}(\mathcal{A}|\vec{\beta}) = \operatorname{rg}(\mathcal{A})\end{aligned}$$

gilt. □

Die Betrachtung mehrerer rechter Seiten führt auf die folgenden Begriffe:

Definition 4.1.2. Ein LGS $\mathcal{A}\vec{x} = \vec{\beta}$ mit $\mathcal{A} \in K^{(m,n)}$ heißt universell lösbar, wenn $(*)$ für jede rechte Seite $\vec{\beta} \in K^m$ lösbar ist bzw. heißt eindeutig lösbar, wenn $(*)$ für jede rechte Seite $\vec{\beta}$ höchstens eine (eventuell keine) Lösung $\vec{x} \in K^n$ besitzt.

Satz 4.1.2. *Ein LGS* $\mathcal{A}\vec{x} = \vec{\beta}$ *mit* $\mathcal{A} \in K^{(m,n)}$ *ist genau dann*

i) universell lösbar, wenn $\operatorname{rg}(\mathcal{A}) = m$ *bzw.*

ii) eindeutig lösbar, wenn $\operatorname{rg}(\mathcal{A}) = n$

ist.

Beweis. Betrachte $\varphi \in L(K^n, K^m)$ mit $\varphi(\vec{x}) = \mathcal{A}\vec{x}$. Dann gilt:

$$\mathcal{A}\vec{x} = \vec{\beta} \text{ universell lösbar} \Leftrightarrow \varphi \text{ surjektiv, d.h. } \varphi(K^n) = K^m \underset{\text{S. 3.5.8}}{\Leftrightarrow} \operatorname{rg}(\mathcal{A}) = \operatorname{rg}(\varphi) = \dim K^m = m.$$

Ebenso gilt

$$\begin{aligned}\mathcal{A}\vec{x} = \vec{\beta} \text{ eindeutig lösbar} \quad &\Leftrightarrow \quad \varphi \text{ injektiv} \underset{\text{S. 3.2.4}}{\Leftrightarrow} \operatorname{Kern}(\varphi) = \{\vec{0}\} \\ &\underset{\text{S. 3.2.5}}{\Leftrightarrow} \quad 0 = \operatorname{def}(\varphi) = \dim K^n - \operatorname{rg}(\varphi) = n - \operatorname{rg}(\mathcal{A}) \Leftrightarrow \operatorname{rg}(\mathcal{A}) = n.\end{aligned}$$

□

Satz 4.1.3. *Ein LGS* $\mathcal{A}\vec{x} = \vec{\beta}$ *mit* $\mathcal{A} \in K^{(m,n)}$ *ist genau dann universell und eindeutig lösbar, wenn* $m = n = \operatorname{rg}(\mathcal{A})$ *ist, d.h. wenn* $\mathcal{A}$ *quadratisch und regulär ist. Für ein LGS der Form* $\mathcal{A}\vec{x} = \vec{\beta}$ *mit* $\mathcal{A} \in K^{(n,n)}$ *(d.h. Anzahl der Unbekannten entspricht der Anzahl der Gleichungen) gilt:*

$$\mathcal{A}\vec{x} = \vec{\beta} \textit{ universell lösbar} \Leftrightarrow \mathcal{A}\vec{x} = \vec{\beta} \textit{ eindeutig lösbar.}$$

Beweis. Das folgt unmittelbar aus Satz 4.1.2. □

Satz 4.1.4. *Die Lösungsmenge* $H = \{\vec{x} \in K^n \colon \mathcal{A}\vec{x} = \vec{0}\}$ *eines homogenen LGS mit* $\mathcal{A} \in K^{(m,n)}$ *ist ein Unterraum des* K^n *der Dimension* $\dim H = n - \operatorname{rg}(\mathcal{A})$.

Beweis. Es ist $H = \operatorname{Kern}(\varphi)$ mit $\varphi \colon \vec{x} \to \mathcal{A}\vec{x}$, also gilt nach Satz 3.2.5

$$\dim H = n - \operatorname{rg}(\varphi) = n - \operatorname{rg}(\mathcal{A}).$$

□

Um die Lösungsmenge eines inhomogenen LGS beschreiben zu können, verallgemeinern wir den Begriff der linearen Mannigfaltigkeit von Definition 1.6.8 auf allgemeine Körper:

Definition 4.1.3. Eine Teilmenge M eines Vektorraums V heißt eine lineare Mannigfaltigkeit (oder auch affiner Unterraum) der Dimension m, wenn sie als

$$M = \{\vec{v}_0 + \vec{u} \colon \vec{u} \in U\}$$

für ein festes $\vec{v}_0 \in V$ und einem Untervektorraum U von V mit $\dim U = m$ geschrieben werden kann.

Satz 4.1.5. *Der Vektorraum U ist durch M eindeutig bestimmt, d.h. gilt*

$$M = \{\vec{v}_0 + \vec{u} \colon \vec{u} \in U_0\} = \{\vec{v}_1 + \vec{u} \colon \vec{u} \in U_1\},$$

so ist $U_0 = U_1$. Für $\vec{v}_0$ kann jeder Vektor aus M genommen werden.

Beweis. Es sei $\vec{u}_0 \in U_0$. Dann gilt $\vec{v}_0 + \vec{u}_0 = \vec{v}_1 + \vec{u}_1$ mit gewissen $\vec{v}_1 \in M$ und $\vec{u}_1 \in U_1$. Wegen $\vec{v}_0 = \vec{v}_0 + \vec{0} \in M$ ist $\vec{v}_0 = \vec{v}_1 + \vec{w}_1$ mit $\vec{w}_1 \in U_1$, also

$$\vec{v}_0 + \vec{u}_0 = \vec{v}_1 + \vec{w}_1 + \vec{u}_0 = \vec{v}_1 + \vec{u}_1 \Rightarrow \vec{u}_0 = \vec{u}_1 - \vec{w}_1 \in U_1$$

und somit $U_0 \subseteq U_1$. Genauso folgt $U_1 \subseteq U_0$, also $U_0 = U_1$.
Es sei nun $\vec{v}_0' \in M$. Dann ist

$$\vec{v}_0' = \vec{v}_0 + \vec{u}' \text{ (mit } \vec{u}' \in U) \Rightarrow M = \{\vec{v}_0 + \vec{u} \colon \vec{u} \in U\} = \{\vec{v}_0' + \vec{u} \colon \vec{u} \in U\}.$$

□

Beispiel 4.1.1. Den ersten Beispielen sind wir bereits in der Einleitung begegnet:
Die Geraden im $\mathbb{R}^2$ oder $\mathbb{R}^3$, die als $\{\vec{a} + t\vec{b} \colon t \in \mathbb{R}\}$ geschrieben werden können, bzw. die Ebenen im $\mathbb{R}^3$, in der Form $\{\vec{a} + s\vec{b} + t\vec{c} \colon s, t \in \mathbb{R}\}$ mit linear unabhängigen $\vec{b}, \vec{c}$, sind eindimensionale bzw. zweidimensionale Mannigfaltigkeiten.

Satz 4.1.6. *Eine lineare Mannigfaltigkeit eines Vektorraums V ist genau dann ein Untervektorraum von V, wenn sie den Nullvektor enthält.*

Beweis. Wenn M ein Untervektorraum ist, ist klar, dass $\vec{0} \in M$ ist. Andererseits gilt für $\vec{0} \in M$ nach Satz 4.1.5, dass $M = \{\vec{0} + \vec{u} \colon \vec{u} \in U\} = U$ ist. □

Satz 4.1.7. *Sind M und N lineare Mannigfaltigkeiten von V mit $\dim N < \infty$ und $M \subseteq N$, so ist $\dim M \leq \dim N$. Es gilt genau dann $\dim M = \dim N$, wenn $M = N$ ist.*

Beweis. Es sei $\vec{v}_0 \in M$. Nach Satz 4.1.5 gilt dann $M = \{\vec{v}_0 + \vec{u} \colon \vec{u} \in U_0\}$ und $N = \{\vec{v}_0 + \vec{u} \colon \vec{u} \in U_1\}$ mit gewissen Untervektorräumen U_0 und U_1 von V. Es folgt $U_0 \subseteq U_1$, und die Restbehauptung folgt aus Satz 2.4.7. □

Satz 4.1.8. *Ist $\varphi \in L(V, V')$ mit $\dim V, \dim V' < \infty$, und M eine lineare Mannigfaltigkeit in V, so ist $\varphi(M) = \{\varphi(\vec{v}) \colon \vec{v} \in M\}$ eine lineare Mannigfaltigkeit in V'. Ist φ ein Isomorphismus, so ist $\dim M = \dim \varphi(M)$.*

Beweis. Es ist

$$\varphi(M) = \{\varphi(\vec{v}_0) + \varphi(\vec{u}) \colon \vec{u} \in U\} = \{\varphi(\vec{v}_0) + \vec{w} \colon \vec{w} \in \varphi(U)\}.$$

□

Wir beschreiben nun die Lösungsmenge eines inhomogenen LGS:

Satz 4.1.9. *Ist das LGS*

$$\mathcal{A}\vec{x} = \vec{\beta} \tag{$*$}$$

mit $\mathcal{A} \in K^{(m,n)}$ *und* $\vec{\beta} \in K^m$ *lösbar, so ist die Lösungsmenge eine lineare Mannigfaltigkeit des* K^n *der Dimension* $n - \mathrm{rg}(\mathcal{A})$. *Man erhält alle Lösungen des LGS in der Form* $\vec{x} = \vec{x}_0 + \vec{x}_h$, *wenn* $\vec{x}_0$ *eine beliebige aber feste partikuläre (oder spezielle) Lösung des LGS ist, und* $\vec{x}_h$ *sämtliche Lösungen des zugehörigen homogenen LGS*

$$\mathcal{A}\vec{x} = \vec{0} \tag{$**$}$$

durchläuft. Die Lösungsmenge von $(*)$ *ist genau dann ein Vektorraum, wenn* $\vec{\beta} = \vec{0}$ *ist.*

Beweis. Es sei $\vec{x}_h$ eine Lösung von $(**)$. Dann ist $\vec{x} = \vec{x}_0 + \vec{x}_h$ eine Lösung von $(*)$, denn es gilt

$$\mathcal{A}\vec{x} = \mathcal{A}(\vec{x}_0 + \vec{x}_h) = \mathcal{A}\vec{x}_0 + \mathcal{A}\vec{x}_h = \vec{\beta} + \vec{0} = \vec{\beta}.$$

Jede Lösung $\vec{x}$ von $(*)$ hat die Form $\vec{x} = \vec{x}_0 + \vec{x}_h$, denn aus $\mathcal{A}\vec{x} = \vec{\beta}$ folgt $\mathcal{A}(\vec{x} - \vec{x}_0) = \vec{0}$, also ist $\vec{x} - \vec{x}_0 =: \vec{x}_n$ eine Lösung von $(**)$. Die Lösungsmenge von $(*)$ ist eine lineare Mannigfaltigkeit der Dimension $n - \mathrm{rg}(\mathcal{A})$, da nach Satz 4.1.4 die Lösungsmenge von $(**)$ ein Unterraum des K^n mit dieser Dimension ist. Die Lösungsmenge von $(*)$ ist nach Satz 4.1.6 genau dann ein Untervektorraum, wenn $\vec{x} = \vec{0}$ eine Lösung von $(*)$ ist. Dies ist jedoch äquivalent zu $\vec{\beta} = \mathcal{A}\vec{0} = \vec{0}$. □

Es gilt nun auch die Umkehrung von Satz 4.1.9: Jede lineare Mannigfaltigkeit des K^n lässt sich als Lösungsmenge eines linearen Gleichungssystems darstellen. Wir beweisen den etwas allgemeineren

Satz 4.1.10. *Es sei* M *eine lineare Mannigfaltigkeit des Vektorraums* V *über* K *mit* $\dim M = m$ *und* $\dim V = n$ *mit* $0 \le m \le n$ *und* $k = n - m$. *Es sei* $L(M) \subseteq V^*$ *die Menge aller Linearformen, die auf* M *konstant sind. Dann ist* $L(M)$ *ein* k*-dimensionaler Unterraum von* V^*. *Für jede Basis* $\{\varphi_1, \ldots, \varphi_k\}$ *von* $L(M)$ *gibt es* $c_1, \ldots, c_k \in K$, *so dass*

$$\vec{v} \in M \Leftrightarrow \begin{cases} \varphi_1(\vec{v}) = c_1 \\ \qquad\vdots \\ \varphi_k(\vec{v}) = c_k \end{cases}$$

Dann ist M *genau dann ein Untervektorraum von* V, *wenn* $c_1 = \cdots = c_k = 0$ *ist.*

Beweis. Sei zunächst U ein m-dimensionaler Untervektorraum von V, und $B = \{\vec{b}_1, \ldots, \vec{b}_n\}$ eine Basis von V. Nach Satz 3.7.6 gibt es zu jedem $\varphi \in V^*$ eindeutig bestimmte $\alpha_1, \ldots, \alpha_n \in K$ mit $\varphi(\vec{v}) = \alpha_1\lambda_1 + \cdots + \alpha_n\lambda_n$ für $\vec{v} = \lambda_1\vec{b}_1 + \cdots + \lambda_n\vec{b}_n \in V$. Die Abbildung $\Psi\colon \varphi \to (\alpha_1, \ldots, \alpha_n)$ ist nach Satz 3.7.1 ein Isomorphismus von V^* nach K^n. Die Vektoren

$$\begin{array}{ccccccc} \vec{v}_1 & = & \lambda_{11}\vec{b}_1 & + & \cdots & + & \lambda_{1n}\vec{b}_n \\ \vec{v}_2 & = & \lambda_{21}\vec{b}_1 & + & \cdots & + & \lambda_{2n}\vec{b}_n \\ \vdots & & \vdots & & & & \vdots \\ \vec{v}_m & = & \lambda_{m1}\vec{b}_1 & + & \cdots & + & \lambda_{mn}\vec{b}_n \end{array}$$

mögen eine Basis von U bilden, dann hat die Matrix

$$\mathcal{L} = \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1n} \\ \vdots & & \vdots \\ \lambda_{m1} & \cdots & \lambda_{mn} \end{pmatrix}$$

den Rang m. Wären nämlich die Zeilen von $\mathcal{L}$ linear abhängig, dann wären auch $\vec{v}_1, \ldots, \vec{v}_m$ linear abhängig, was der Basiseigenschaft widerspricht.

Es gilt für alle $\vec{v} \in U$

$$\varphi(\vec{v}) = 0 \Leftrightarrow \varphi(\vec{v}_1) = \ldots = \varphi(\vec{v}_n) = 0 \Leftrightarrow \begin{cases} \lambda_{11}\alpha_1 & + & \cdots & + & \lambda_{1n}\alpha_n & = & 0 \\ \lambda_{21}\alpha_1 & + & \cdots & + & \lambda_{2n}\alpha_n & = & 0 \\ \vdots & & & & \vdots & & \vdots \\ \lambda_{m1}\alpha_1 & + & \cdots & + & \lambda_{mn}\alpha_n & = & 0 \end{cases} \tag{$*$}$$

Das System $(*)$ fassen wir nun als LGS mit der Koeffizientenmatrix $\mathcal{L}$ und den Unbekannten $\alpha_1, \ldots, \alpha_n$ auf. Aufgrund von $\operatorname{rg}(\mathcal{L}) = m$ bildet nach Satz 4.1.9 die Lösungsmenge $\{(\alpha_1, \ldots, \alpha_n)\}$ von $(*)$ einen Unterraum $\tilde{L}$ des K^n der Dimension $k = n - m$. Die Menge $L(U) = \Psi^{-1}(\tilde{L})$ bildet einen k-dimensionalen Unterraum von V^*. Sei nun M eine m-dimensionale lineare Mannigfaltigkeit, also $M = \{\vec{v}_0 + \vec{u} \colon \vec{u} \in U\}$ sowie U ein Unterraum von V mit $\dim U = m$, so folgt

$$\varphi(\vec{v}) \text{ konstant auf } M \Leftrightarrow \forall \vec{u} \in U \colon \varphi(\vec{v}_0 + \vec{u}) = \varphi(\vec{v}_0) \Leftrightarrow \forall u \in U \colon \varphi(\vec{u}) = 0.$$

Diese φ bilden nun einen k-dimensionalen Unterraum von V^*. Es sei nun $\{\varphi_1, \ldots, \varphi_k\}$ eine Basis von $L(M)$ und

$$\varphi_i(\vec{v}) = c_i \tag{$**$}$$

mit $1 \le i \le k$ und für alle $\vec{v} \in M$. Dazu sei $\Psi(\varphi_i) = (\alpha_{i1}, \ldots, \alpha_{in})$ und

$$\mathcal{A} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{k1} & \cdots & \alpha_{kn} \end{pmatrix}.$$

Dann ist $\operatorname{rg}(\mathcal{A}) = k$. Für $\vec{v} = x_1\vec{b}_1 + \cdots + x_n\vec{b}_n$ ist $(**)$ zu

$$\mathcal{A} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix} \tag{$***$}$$

äquivalent. Nach Satz 4.1.9 ist die Lösungsmenge N von $(***)$ eine m-dimensionale Mannigfaltigkeit von V mit $M \subseteq N$. Nach Satz 4.1.2 ist $M = N$. □

Satz 4.1.11. *Es sei $A \in \mathrm{GL}(n, K)$. Man erhält A^{-1}, indem man an E_n simultan diesselben Zeilenumformungen vornimmt, die man verwendet, um A (gemäß dem Gaußschen Algorithmus) in E_n zu überführen.*

Beweis. Es sei

$$A = (a_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le n}} \quad \text{und} \quad A^{-1} = (y_{kl})_{\substack{1 \le k \le n \\ 1 \le l \le n}}$$

mit den Spaltenvektoren

$$\vec{y}_l = \begin{pmatrix} y_{1l} \\ \vdots \\ y_{nl} \end{pmatrix}.$$

Die Einheitsmatrix E_n werde durch die Zeilenumformungen, die A in E_n überführen, in die Matrix $B = (b_{gh})_{\substack{1 \le g \le n \\ 1 \le h \le n}}$ mit den Spaltenvektoren

$$\vec{b}_k = \begin{pmatrix} b_{1k} \\ \vdots \\ b_{nk} \end{pmatrix}$$

überführt.

Dann ist $\vec{x} = \vec{y}_l$ die Lösung des LGS $A\vec{x} = \vec{e}_l^{\,T}$, wobei $\vec{e}_1, \ldots, \vec{e}_n$ die Standardbasis des K^n bedeutet. Die Anwendung des Gaußschen Algorithmus führt die n Gleichungssysteme $A\vec{x} = \vec{e}_l^{\,T}$ mit $1 \leq l \leq n$ mit den Lösungen $\vec{x} = \vec{y}_l$ in die Gleichungssysteme $E_n\vec{x} = \vec{b}_l$ mit $1 \leq l \leq n$, mit den Lösungen $\vec{x} = \vec{b}_l$ über.
Es folgt $\vec{y}_l = \vec{b}_l$ und damit $B = A^{-1}$. □

Beispiel 4.1.2. Wir berechnen die Inverse der Matrix

$$A = \begin{pmatrix} 6 & 2 & 3 \\ 4 & 5 & -2 \\ 7 & 2 & 4 \end{pmatrix}$$

mit der folgenden Umformungskette:

A			E		
6	2	3	1	0	0
4	5	-2	0	1	0
7	2	4	0	0	1
-1	0	-1	1	0	-1
4	5	-2	0	1	0
7	2	4	0	0	1
1	0	1	-1	0	1
0	5	-6	4	1	-4
0	2	-3	7	0	-6
1	0	1	-1	0	1
0	1	0	-10	1	8
0	0	-3	27	-2	-22
3	0	3	-3	0	3
0	1	0	-10	1	8
0	0	-3	27	-2	-22
3	0	0	24	-2	-19
0	1	0	-10	1	8
0	0	-3	27	-2	-22
3	0	0	24	-2	-19
0	3	0	-30	3	24
0	0	3	-27	2	22
1	0	0			
0	1	0		A^{-1}	
0	0	1			

mit der Inversen

$$A^{-1} = \frac{1}{3} \cdot \begin{pmatrix} 24 & -2 & -19 \\ -30 & 3 & 24 \\ -27 & 2 & 22 \end{pmatrix}.$$

Man rechnet leicht die Probe

$$\frac{1}{3} \cdot \begin{pmatrix} 6 & 2 & 3 \\ 4 & 5 & -2 \\ 7 & 2 & 4 \end{pmatrix} \cdot \begin{pmatrix} 24 & -2 & -19 \\ -30 & 3 & 24 \\ -27 & 2 & 22 \end{pmatrix} = \frac{1}{3} \cdot \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} = E_3$$

nach.