ulm university universität uulm

Vorabskript zur Vorlesung

Lineare Algebra I und II

Wintersemester 2011/ 12
Sommersemester 2012

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

Institut für Zahlentheorie und Wahrscheinlichkeitstheorie
Universität Ulm

Inhaltsverzeichnis

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 Elementargeometerie 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 Anschuung 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.

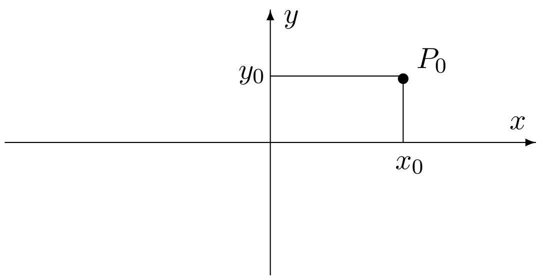
Das Konzept, das ganz am Anfang der Linearen Algebra steht, ist der Begriff des Vektorraumes. Auch in dieser Einleitung steht dieser Begriff im Mittelpunkt. Wir werden hier jedoch nur Vektorräume über dem Körper der reellen Zahlen- reelle Vektorräume- betrachten, die viele Aspekte, mit den aus der Elementargeometrie 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

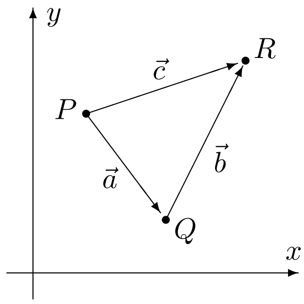
Den zu $\vec{v}$ gleichgerichteten, aber entgegengesetzten Vektor bezeichnen wir mit $-\vec{v}$ mit

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

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:

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

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:

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

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

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

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

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

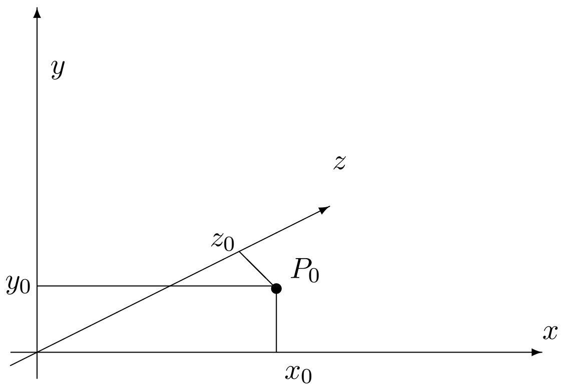
nach Subtraktion des 3-fachen der ersten Zeile von der zweiten Zeile. Dies führt zu $y = \frac{8}{11}$ und $x = \frac{31}{11}$, also zu dem Schnittpunkt $P_0 = \left(\frac{31}{11}, \frac{8}{11}\right)$.

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 <u>Rechtssystem</u> 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 <u>n- tupel</u> 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 <u>Punkte</u> 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 <u>Komponenten</u> 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

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

mit der Koeffizientenmatrix

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

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:

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

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).

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

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

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

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:

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

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.

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

Daraus resultiert dann

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

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

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

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

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

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

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

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

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

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.

"$\Rightarrow$":
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

$$b^{-1} = eb^{-1}. \tag{$*$}$$

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 $\langle \vec{v}_1 \rangle = \{t\vec{v}_1 : t \in \mathbb{R}\}$ 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:

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

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 $\mathbb{F}_2 = \{0, 1\}$. Die Verknüpfungen $+$ der Addition und $\cdot$ der Multiplikation sind durch folgende Tafeln definiert:

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

Beispiel 2.1.4. Es sei $\mathbb{F}_3 = \{0, 1, 2\}$. Die Verknüpfungen $+$ der Addition und $\cdot$ der Multiplikation sind durch folgende Tafeln definiert:

$$\begin{array}{c|ccc} + & 0 & 1 & 2 \\ \hline 0 & 0 & 1 & 2 \\ 1 & 1 & 2 & 0 \\ 2 & 2 & 0 & 1 \end{array} \quad \text{bzw.} \quad \begin{array}{c|cc} \cdot & 1 & 2 \\ \hline 1 & 1 & 2 \\ 2 & 2 & 1. \end{array}$$

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*

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

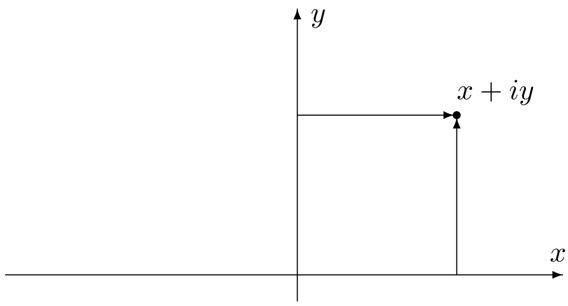
Beweis. Die Körperaxiome folgen durch Nachrechnen aus Definition 2.2.1.
Zum Nachweis der Inversen verwendet man

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

falls $x + iy \neq 0$. $\square$

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.

Skizze:



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 $\vec{x} = (x_1, \ldots, x_n)^T, \vec{y} = (y_1, \ldots, y_n)^T \in \mathbb{F}_q^n$ 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 linear abhängig, ein einzelner Vektor $\vec{v} \neq \vec{0}$ ist stets linear unabhängig.

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

iii) Sind $\vec{v}_1, \dots, \vec{v}_k$ linear unabhängig, so auch $\vec{v}_1, \dots, \vec{v}_m$ für $1 \leq m \leq k$.

iv) Ist $\vec{v}$ eine Linearkombination von $\vec{v}_1, \dots, \vec{v}_k$, so sind $\vec{v}_1, \dots, \vec{v}_k, \vec{v}$ linear abhängig.

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

vi) Sind $\vec{v}_1, \dots, \vec{v}_k$ linear unabhängig, aber $\vec{v}_1, \dots, \vec{v}_k, \vec{v}$ linear abhängig, so ist $\vec{v}$ eine Linearkombination der $\vec{v}_1, \dots, \vec{v}_k$.

Beweis. i) Es gilt $1 \cdot \vec{0} = \vec{0}$, also ist $\vec{0}$ linear abhängig, 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 $(\lambda_1, \dots, \lambda_k) \neq (0, \dots, 0)$ 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 $(\lambda_1, \dots, \lambda_l) \neq (0, \dots, 0)$, wenn man $\lambda_{k+1} = \ldots = \lambda_l = 0$ einsetzt.

iii) Wären $\vec{v}_1, \dots, \vec{v}_m$ linear abhängig, so nach (ii) auch $\vec{v}_1, \dots, \vec{v}_k$.

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 $(\lambda_1, \dots, \lambda_k) \neq (0, \dots, 0)$. 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 $(\lambda_1, \dots, \lambda_k, \lambda) \neq (0, \dots, 0)$. Dann muss $\lambda \neq 0$ gelten, da wegen der linearen Unabhängigkeit von $\vec{v}_1, \dots, \vec{v}_k$ 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 $\vec{v} = (-\lambda^{-1})\lambda_1\vec{v}_1 + \ldots + (-\lambda^{-1})\lambda_k\vec{v}_k$.

□

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 linear unabhängig, wenn je endlich viele verschiedene (!) Vektoren aus M linear unabhängig sind, andernfalls linear abhängig.

Bemerkung 2.4.1. Wir stellen einige Spezialfälle zusammen:

1. Die leere Menge ist linear unabhängig.
2. Ist $\vec{0} \in M$, so ist M linear abhängig.
3. Jede Teilmenge einer linear unabhängigen Menge ist linear unabhängig.
4. Jede in V liegende Obermenge einer linear abhängigen Menge ist linear abhängig.

5. Achtung: Im Falle $\vec{a}_1 = \vec{a}_2 \neq \vec{0}$ sind $\vec{a}_1$ und $\vec{a}_2$ linear abhängig, aber $\{\vec{a}_1, \vec{a}_2\}$ ist linear unabhängig (weil es tatsächlich die Menge $\{\vec{a}_1\}$ ist).

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

$$V = \langle\langle\, M\,\rangle\rangle \Leftrightarrow M \text{ ist linear unabhängig und } \langle\, M\,\rangle = V.$$

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 endlichdimensionale Vektorräume beschränken.

Definition 2.4.4. Gibt es eine maximale Zahl n von linear unabhängiger Vektoren in V, so heißt n die Dimension von V, geschrieben $\dim V$:

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

Gibt es kein solches n (existiert also zu jedem $k \in \mathbb{N}$ eine linear unabhängige Teilmenge $M \subset V$ mit $|M| = k$), so heißt V unendlichdimensional, 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 linear unabhängig, 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 linear abhängig, 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

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

Dann ist M linear unabhängig.
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 $\nu_i \le x < \nu_i + 1$ 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 linear unabhängige Teilmenge mit* n *Vektoren eine Basis von* V*.*

Beweis. Seien $\vec{v}_1, \ldots, \vec{v}_n \in V$ linear unabhängig und $\vec{v} \in V$ beliebig. Nach Definition der Dimension sind $\vec{v}_1, \ldots, \vec{v}_n, \vec{v}$ linear abhängig. 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* $\vec{v}_1, \ldots \vec{v}_n \in V$ *sind äquivalent:*

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

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 $(\lambda_1, \ldots, \lambda_n) \in K^n$*, oder:*
die durch $(*)$ *vermittelte Abbildung*

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

ist bijektiv.

Beweis. "⇒:"
Da $\{\vec{v}_1, \ldots, \vec{v}_n\}$ 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 $j = 1, \ldots, n$, also $\lambda_j = \mu_j$, d.h. die Darstellung $(*)$ ist eindeutig.
"⇐:"
Eine (also die einzige) Darstellung $(*)$ von $\vec{0}$ ist $\vec{0} = 0 \cdot \vec{v}_1 + \ldots + 0 \cdot \vec{v}_n$. Also sind $\vec{v}_1, \ldots, \vec{v}_n$ linear unabhängig. □

Satz 2.4.4. *Es sei* $V = \langle\langle \vec{v}_1, \ldots, \vec{v}_n \rangle\rangle$ *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 $(\lambda_1, \ldots, \lambda_n) \in K^n$ *sei* j *ein Index mit* $\lambda_j \neq 0$*. Dann ist auch*

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

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

i) $V = \langle \vec{w}, \vec{v}_2, \ldots, \vec{v}_n \rangle$:
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 $i = 2, \ldots, n$.
Sei nun $\vec{v} \in V$ beliebig, etwa $\vec{v} = \alpha_1 \vec{v}_1 + \ldots + \alpha_n \vec{v}_n$ mit $\alpha_1, \ldots, \alpha_n \in K$. Einsetzen von $(**)$ ergibt

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

ii) $\{\vec{w}, \vec{v}_2, \ldots, \vec{v}_n\}$ ist linear unabhängig:
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 $\vec{v}_1, \ldots, \vec{v}_n$ linear unabhängig sind, also ist $\mu_1 = 0$ wegen $\lambda_1 \neq 0$. Daraus folgt dann $\mu_2 = \ldots = \mu_n = 0$, also ist $\{\vec{w}, \vec{v}_2, \ldots, \vec{v}_n\}$ linear unabhängig.

□

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 $\vec{w}_1, \ldots, \vec{w}_k \in V$ linear unabhängig, 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 linear unabhängige 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
$$V = \langle\langle\, \vec{w}_1, \ldots, \vec{w}_k, \vec{v}_{k+1}, \ldots, \vec{v}_n \,\rangle\rangle\,. \tag{$*$}$$
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 linear unabhängig 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 $\vec{w}_1, \ldots, \vec{w}_k, \vec{w}_{k+1}$ linear abhängig, 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 *linear unabhängigen Vektoren besteht.*

Beweis. Für jede linear unabhängige 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 $n \leq \dim V$. Es folgt insgesamt $\dim V = n$. Nun sei $\mathcal{B}$ eine beliebige 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 linear unabhängige 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 $\{\vec{e}_1, \ldots, \vec{e}_n\}$ 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 $\{\vec{e}_1, \ldots, \vec{e}_n, i\vec{e}_1, \ldots, i\vec{e}_n\}$.

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 $\{\vec{v}_1, \ldots, \vec{v}_n\}$. 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 $\{\vec{a}_1, \ldots, \vec{a}_d\}$ eine Basis von $U_1 \cap U_2$ (falls $d = 0$ ist, ist dies die leere Menge). Wir ergänzen diese Basis nach Satz 2.4.5 zu je einer Basis von U_1 und U_2:

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

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

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 linear unabhängig:
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 linear unabhängig 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 $\mathcal{B}_1$ linear unabhängig ist. Also verschwinden sämtliche Koeffizienten $\alpha_i, \beta_j, \gamma_k$, womit $\mathcal{B}$ linear unabhängig 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

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

für feste $a_1,\ldots,a_n \in \mathbb{R}$ ist eine sogenannte Linearform.

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 Differentialoperator.

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. □

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 $\varphi \in L(V,V)$. □

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 $\psi \circ \varphi : V \to V$ 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}. □

3.2 Kern und Bild

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

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

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

iii) Sind $\varphi(\vec{v}_1), \ldots, \varphi(\vec{v}_n) \in V'$ linear unabhängig, 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).

□

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 <u>Bild</u> von φ. Die Menge $\mathrm{Kern}(\varphi) = \varphi^{-1}(\{\vec{0}'\}) = \{\vec{v} \in V \colon \varphi(\vec{v}) = \vec{0}'\}$ heißt <u>Kern</u> von φ. Nach Satz 3.2.2 sind dies Unterräume von V' bzw. V. Ferner heißt $\mathrm{rg}(\varphi) = \dim(\mathrm{Bild}(\varphi))$ der <u>Rang</u> und $\mathrm{def}(\varphi) = \dim(\mathrm{Kern}(\varphi))$ der <u>Defekt</u> von φ.

Beispiel 3.2.1. Wir betrachten die <u>Projektion</u> 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 $\vec{v}_1, \ldots, \vec{v}_n \in V$ linear unabhängig, so auch $\varphi(\vec{v}_1), \ldots, \varphi(\vec{v}_n)$.*

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

<u>(ii)⇒(iii):</u>
Es sei $\mathrm{Kern}(\varphi) = \{\vec{0}\}$, und $\vec{v}_1, \ldots, \vec{v}_n \in V$ linear unabhängig 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 $\vec{v}_1, \ldots, \vec{v}_n$ linear unabhängig sind.

(iii)⇒(i):
Es seien $\vec{v}_1, \vec{v}_2 \in V$ mit $\vec{v}_1 \neq \vec{v}_2$. Dann ist $\vec{v}_1 - \vec{v}_2 \neq \vec{0}$ linear unabhängig, also nach (iii) auch $\varphi(\vec{v}_1 - \vec{v}_2) = \varphi(\vec{v}_1) - \varphi(\vec{v}_2)$, und damit gilt $\varphi(\vec{v}_1) \neq \varphi(\vec{v}_2)$. □

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

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

Die beiden Extremfälle dieser Gleichheit sind

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

Beweis. Wir wählen eine Basis $\vec{v}_1, \ldots, \vec{v}_r$ von $\mathrm{Kern}(\varphi)$ 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 $\mathrm{rg}(\varphi) + \mathrm{def}(\varphi) = r + s = \dim V$, also die Behauptung.

i) $\varphi(\vec{w}_1), \ldots, \varphi(\vec{w}_s)$ sind linear unabhängig:
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

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

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 $\vec{v}_1, \ldots, \vec{v}_r, \vec{w}_1, \ldots, \vec{w}_s$ linear unabhängig 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 $\vec{v}_i \in \mathrm{Kern}(\varphi)$.

□

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* $\mathrm{rg}(\varphi) = \dim V$ *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 $\underset{\text{Satz 3.2.4}}{\Leftrightarrow}$ $\mathrm{def}(\varphi) = 0$ $\underset{\text{Satz 3.2.5}}{\Leftrightarrow}$ $\mathrm{rg}(\varphi) = \dim V$.

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

□

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*

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

Beweis. Die rechte Ungleichung folgt aus

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

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

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

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 $\{\vec{v}_1, \dots, \vec{v}_n\}$ ist dieser Isomorphismus φ durch

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

gegeben. Durch φ wird also der Vektor $\vec{v} = \lambda_1 \vec{v}_1 + \ldots + \lambda_n \vec{v}_n$ in $(\lambda_1, \dots, \lambda_n)$ 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: $\varphi(\vec{v}) = (\lambda_1, \dots, \lambda_n), \varphi(\vec{w}) = (\mu_1, \dots, \mu_n), \varphi(\vec{v}+\vec{w}) = (\lambda_1+\mu_1, \dots, \lambda_n+\mu_n) = \varphi(\vec{v})+\varphi(\vec{w})$

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

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

$\square$

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

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

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 $a_{11}, \ldots, a_{rr}$ 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

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

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 <u>Summe</u> 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 <u>Nullmatrix</u>

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

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 <u>Produkt</u> $C=AB$ 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:

$$C = \left(\begin{array}{c|cc} & & \\ \hline & & \end{array}\right) \leftarrow A = \left(\begin{array}{c} i\text{-te Zeile} \\ \hline \\ \end{array}\right), \quad B = \begin{array}{c} l\text{-te Spalte} \\ \left(\begin{array}{c|cc} & & \\ \end{array}\right) \end{array},$$

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 $\dim V' = m < \infty$, $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ eine Basis von V sowie $\mathcal{B}' = \{\vec{b}'_1, \dots, \vec{b}'_m\}$ 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 $l = 1, \dots, n$.
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 $\mathcal{B}'$ <u>zugeordnete</u> Matrix (auch <u>Darstellungsmatrix</u> 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 $\lambda_1, \dots, \lambda_n$ die <u>Koordinaten</u> von $\vec{v}$ bzgl. $\mathcal{B}$ (entsprechend sind $\lambda'_1, \dots, \lambda'_m$ 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 $k = 1, \ldots, m$, 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 $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ 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:

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

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 $\mathcal{B} = \{\vec{b}_1,\ldots,\vec{b}_n\}$ *eine Basis von* V *sowie* $\mathcal{B}' = \{\vec{b}'_1,\ldots,\vec{b}'_m\}$ *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

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

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

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

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

□

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 $1 \le i, j \le n$ mit dem Kroneckersymbol

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

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

$$\varphi(\vec{b}_{l_1}),\ldots,\varphi(\vec{b}_{l_r}) \text{ linear unabhängig} \Leftrightarrow \vec{a}_{l_1},\ldots,\vec{a}_{l_r} \text{ linear unabhängig.}. \qquad (*)$$

Es gilt

$$\vec{0} = \sum_{j=1}^{r} \lambda_j \varphi(\vec{b}_{l_j}) = \sum_{j=1}^{r} \lambda_j \left(\sum_{k=1}^{m} \alpha_{kl_j}\vec{b}'_k \right) \Leftrightarrow \sum_{j=1}^{r} \alpha_{kl_j}\lambda_j = 0 \Leftrightarrow \lambda_1\vec{a}_{l_1} + \ldots + \lambda_r\vec{a}_{l_r} = \vec{0},$$

für $k = 1,\ldots,m$ und da $\vec{b}'_1,\ldots,\vec{b}'_m$ linear unabhängig sind. Damit ist $(*)$ bewiesen. Es folgt

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

□

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, also wenn $\mathrm{rg}(A) < n$ gilt, 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 $B \in K^{(n,n)}$. $\square$

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. $\square$

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*

$$D_r^{(m,n)} = \left(\begin{array}{cccc|ccc} 1 & & & & & & \\ & 1 & & & & 0 & \\ & & \ddots & & & & \\ 0 & & & 1 & & & \\ \hline & & & & 0 & & \\ & & 0 & & & 0 & \\ & & & & & & \ddots \end{array}\right) = (c_{ij}) \quad \textit{mit} \quad c_{ij} = \begin{cases} 1, & \textit{falls } i = j \leq r \\ 0 & \textit{sonst} \end{cases}$$

überführen.

Beweis. Wie in Abschnitt 1.4 zeigt man, dass A durch elementare Zeilenumformungen und möglicherweise Vertauschungen von Spalten in die Form

$$\left(\begin{array}{cccc|ccc} 1 & & & & & & \\ & 1 & & & \alpha_{1,r+1} & \cdots & \alpha_{1,n} \\ & & \ddots & & \vdots & & \vdots \\ 0 & & & 1 & \alpha_{r,r+1} & \cdots & \alpha_{r,n} \\ \hline & & & & 0 & & \\ & & 0 & & & 0 & \\ & & & & & & \ddots \end{array}\right)$$

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. □

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:

$$A = \begin{pmatrix} \vec{a}_1 \\ \vec{a}_2 \\ \vdots \\ \vec{a}_m \end{pmatrix} \xrightarrow{2} A' = \begin{pmatrix} \vec{a}_1 \\ \vec{a}_2 + \lambda \vec{a}_1 \\ \vdots \\ \vec{a}_m \end{pmatrix}.$$

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

$$\begin{aligned}\tilde{\mathcal{B}} \text{ ist eine Basis} \quad &\Leftrightarrow \quad \varphi(\vec{b}_1), \dots, \varphi(\vec{b}_n) \text{ sind linear unabhängig} \Leftrightarrow \operatorname{rg}(\varphi) = n \\ &\Leftrightarrow \quad \varphi \text{ ist ein Automorphismus.}\end{aligned}$$

□

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 $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ und $\mathcal{B}' = \{\vec{b}'_1, \dots, \vec{b}'_n\}$ Basen von V und $\varphi \in \operatorname{GL}(V)$ ein Automorphismus. Dann gilt $\mathcal{M}(\varphi; \mathcal{B}', \mathcal{B}) = \mathcal{M}(\operatorname{id}_V; \mathcal{B}', \varphi(\mathcal{B}))$ mit $\varphi(\mathcal{B}) = \{\varphi(\vec{b}_1), \dots, \varphi(\vec{b}_n)\}$.*

Beweis. Die Menge $\varphi(\mathcal{B})$ ist nach Satz 3.7.1 eine Basis von V. Beide Matrizen sind gleich (α_{kl}) mit $1 \le k \le n$ bzw. $1 \le l \le n$ und

$$\varphi(\vec{b}_l) = \operatorname{id}_V(\varphi(\vec{b}_l)) = \sum_{k=1}^{n} \alpha_{kl} \vec{b}'_k$$

für $l = 1, \dots, n$. □

Satz 3.7.3. *Es seien $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ und $\varphi(\mathcal{B}) = \{\varphi(\vec{b}_1), \dots, \varphi(\vec{b}_n)\}$ Basen von V für $\varphi \in \operatorname{GL}(V)$. Sind $\lambda_1, \dots, \lambda_n$ bzw. $\lambda'_1, \dots, \lambda'_n$ die Koordinaten von $\vec{v}$ bzgl. $\mathcal{B}$ bzw. bzgl. $\varphi(\mathcal{B})$, d.h. gilt*

$$\vec{v} = \lambda_1 \vec{b}_1 + \dots + \lambda_n \vec{b}_n \quad \textit{bzw.} \quad \vec{v} = \lambda'_1 \varphi(\vec{b}_1) + \dots + \lambda'_n \varphi(\vec{b}_n),$$

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

$$\begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_n \end{pmatrix} = \mathcal{M}(\operatorname{id}_V; \varphi(\mathcal{B}), \mathcal{B}) \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}. \qquad (*)$$

Nach Satz 3.5.4 ist andererseits

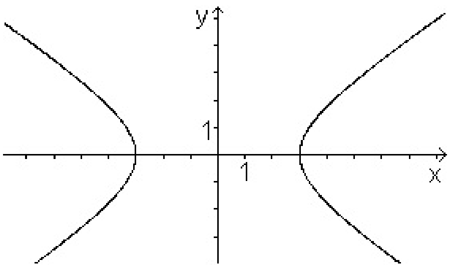
$$E_n = \mathcal{M}(\operatorname{id}_V; \mathcal{B}, \mathcal{B}) = \mathcal{M}(\operatorname{id}_V; \mathcal{B}, \varphi(\mathcal{B})) \cdot \mathcal{M}(\operatorname{id}_V; \varphi(\mathcal{B}), \mathcal{B}),$$

also $\mathcal{M}(\operatorname{id}_V; \varphi(\mathcal{B}), \mathcal{B}) = \mathcal{M}(\operatorname{id}_V; \mathcal{B}, \varphi(\mathcal{B}))^{-1}$ und daher nach Satz 3.7.2

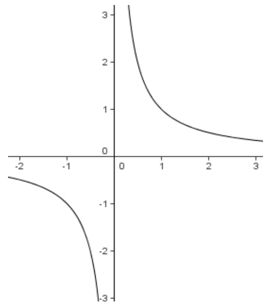
$$\mathcal{M}(\operatorname{id}_V; \varphi(\mathcal{B}), \mathcal{B}) = \mathcal{M}(\varphi; \mathcal{B}, \mathcal{B})^{-1}.$$

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

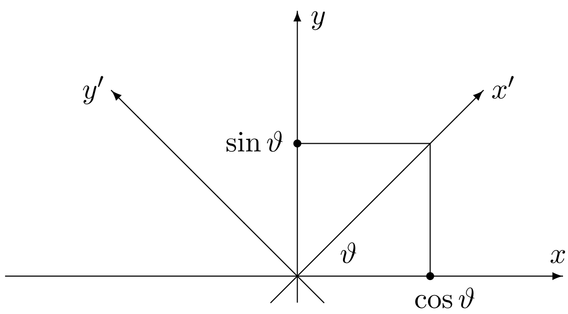


(I) $x^2 - y^2 = b$ $(b > 0)$



(II) $xy = a$ $(a > 0)$

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* $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ *eine Basis von* V, *so 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 $\vec{v} = \lambda_1\vec{b}_1 + \cdots + \lambda_n\vec{b}_n$. *Die Abbildung* $\Psi\colon V^* \to K^n$, $\varphi \to (\alpha_1, \ldots, \alpha_n)$ *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 $\mathcal{M}(\varphi; \{1\}, \mathcal{B}) = (\alpha_1, \ldots, \alpha_n) \in K^{(1\times n)}$, also nach Satz 3.5.2

$$\varphi(\vec{v}) = (\alpha_1, \ldots, \alpha_n) \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \cdot 1$$

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 $\vec{a}_1, \ldots, \vec{a}_n$ von $\mathcal{A}$ ist und wenn

$$\begin{aligned} \vec{\beta} \in \langle \vec{a}_1, \ldots, \vec{a}_n \rangle \quad &\Leftrightarrow \quad \left\langle \vec{a}_1, \ldots, \vec{a}_n, \vec{\beta} \right\rangle = \langle \vec{a}_1, \ldots, \vec{a}_n \rangle \Leftrightarrow \dim \left\langle \vec{a}_1, \ldots, \vec{a}_n, \vec{\beta} \right\rangle = \dim \langle \vec{a}_1, \ldots, \vec{a}_n \rangle \\ &\Leftrightarrow \quad \text{Spaltenrang von } (\mathcal{A}|\vec{\beta}) = \text{Spaltenrang von } \mathcal{A} \Leftrightarrow \operatorname{rg}(\mathcal{A}|\vec{\beta}) = \operatorname{rg}(\mathcal{A}) \end{aligned}$$

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:*

$$\mathcal{A}\vec{x} = \vec{\beta} \text{ universell lösbar} \Leftrightarrow \mathcal{A}\vec{x} = \vec{\beta} \text{ eindeutig lösbar.}$$

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* <u>*partikuläre*</u> *(oder* <u>*spezielle*</u>*) 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* $\{\varphi_1, \dots, \varphi_k\}$ *von* $L(M)$ *gibt es* $c_1, \dots, c_k \in K$, *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 $B = \{\vec{b}_1, \dots, \vec{b}_n\}$ eine Basis von V. Nach Satz 3.7.6 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 $\vec{v} = \lambda_1\vec{b}_1 + \cdots + \lambda_n\vec{b}_n \in V$. Die Abbildung $\Psi\colon \varphi \to (\alpha_1, \dots, \alpha_n)$ 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 $\vec{v}_1, \dots, \vec{v}_m$ linear abhängig, was der Basiseigenschaft widerspricht.

Es gilt für alle $\vec{v} \in U$

$$\varphi(\vec{v}) = 0 \Leftrightarrow \varphi(\vec{v}_1) = \ldots = \varphi(\vec{v}_n) = 0 \Leftrightarrow \begin{cases} \lambda_{11}\alpha_1 & + & \cdots & + & \lambda_{1n}\alpha_n & = & 0 \\ \lambda_{21}\alpha_1 & + & \cdots & + & \lambda_{2n}\alpha_n & = & 0 \\ \vdots & & & & \vdots & & \vdots \\ \lambda_{m1}\alpha_1 & + & \cdots & + & \lambda_{mn}\alpha_n & = & 0 \end{cases} \qquad (*)$$

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

$$\varphi_i(\vec{v}) = c_i \qquad (**)$$

mit $1 \leq i \leq k$ 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

$$\mathcal{A} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix} \qquad (***)$$

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 $M = N$. $\square$

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

$$A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \quad \text{und} \quad A^{-1} = (y_{kl})_{\substack{1 \leq k \leq n \\ 1 \leq l \leq n}}$$

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 $B = (b_{gh})_{\substack{1 \leq g \leq n \\ 1 \leq h \leq n}}$ 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.

Kapitel 5

Determinanten

5.1 Permutationen

Definition 5.1.1. Eine Permutation σ der Menge $N = \{1, 2, \ldots, n\}$ ist eine bijektive Abbildung $\sigma\colon \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$. Unter S_n verstehen wir die Gruppe der Permutationen von N mit der Komposition als Verknüpfung (siehe Beispiel 1.5.8 für den Fall $n = 3$).
Wir schreiben Permutationen in der Form

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}.$$

Eine Permutation $\tau \in S_n$, welche nur zwei Zahlen $i \neq j$ vertauscht, also $\tau(i) = j$, $\tau(j) = i$ und $\tau(k) = k$ für $k \notin \{i, j\}$, heißt Transposition. Wir schreiben sie in der Form $\tau = (i\ j)$.

Satz 5.1.1. *Es ist* $|S_n| = n!$

Beweis. Mit abzählbarer Kombinatorik lässt sich die Anzahl der Möglichkeiten bestimmen, n Elemente mit Berücksichtigung der Reihenfolge anzuordnen. Für das erste Element gibt es n Möglichkeiten, für das zweite noch $n - 1$, usw. Für das vorletzte verbleiben noch zwei Möglichkeiten, womit das letzte festgelegt ist. Also gilt

$$|S_n| = n \cdot (n-1) \cdots 2 \cdot 1 = n!.$$

□

Definition 5.1.2. Es sei $\sigma \in S_n$. Ein Paar (i, j) mit $1 \leq i < j \leq n$ heißt Fehlstandspaar (oder Inversion) der Permutation σ, falls $\sigma(i) > \sigma(j)$ ist. Die Anzahl der Fehlstandspaare von σ heißt Fehlstandszahl $\phi(\sigma)$. Die Parität oder das Signum von σ ist durch $\mathrm{sgn}(\sigma) = (-1)^{\phi(\sigma)}$ definiert. Dabei gilt

$$\begin{array}{lllllll} \sigma & \text{heißt} & \underline{\text{gerade}} & \Leftrightarrow & \mathrm{sgn}(\sigma) = 1 & \Leftrightarrow & \phi(\sigma) \ \text{gerade}, \\ \sigma & \text{heißt} & \underline{\text{ungerade}} & \Leftrightarrow & \mathrm{sgn}(\sigma) = -1 & \Leftrightarrow & \phi(\sigma) \ \text{ungerade}. \end{array}$$

Beispiel 5.1.1. Es sei $\sigma \in S_5$ durch

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 1 & 2 & 4 \end{pmatrix}.$$

gegeben. Dann hat σ die Fehlstandspaare $(1, 3)$,$(1, 4)$,$(2, 3)$,$(2, 4)$,$(2, 5)$. Es ist also $\phi(\sigma) = 5$ und $\mathrm{sgn}(\sigma) = (-1)^5 = -1$, womit σ eine ungerade Permutation ist.

Satz 5.1.2. *Für jedes $\sigma \in S_n$ gilt:*

i) Es kann σ als Produkt von Transpositionen geschrieben werden: $\sigma = \tau_1 \circ \cdots \circ \tau_r$.

ii) Es ist $\operatorname{sgn}(\sigma) = (-1)^r$.

iii) Insbesondere gilt $\operatorname{sgn}(\varrho \circ \sigma) = \operatorname{sgn}(\varrho) \cdot \operatorname{sgn}(\sigma)$ *für alle $\varrho, \sigma \in S_n$.*

Bemerkung 5.1.1. Die Anzahl r ist durch σ nicht eindeutig bestimmt. Es ist aufgrund von

$$\operatorname{sgn}(\sigma) = (-1)^r = \begin{cases} 1, & \text{falls } r \text{ gerade ist} \\ -1, & \text{falls } r \text{ ungerade ist} \end{cases}$$

aber bestimmt, ob r gerade oder ungerade ist.

Beweis. (Beweis von Satz 5.1.2)

i) Wir zeigen zunächst:
Ist $\sigma \in S_n$ mit $\phi(\sigma) = m$ und $\tau = (\sigma(h), \sigma(h+1)) \in S_n$ mit $1 \leq h \leq n-1$, dann gilt

$$\phi(\tau \circ \sigma) = \begin{cases} m-1, & \text{falls } (\sigma(h), \sigma(h+1)) \text{ ein Fehlstandspaar von } \sigma \text{ ist,} \\ m+1, & \text{falls } (\sigma(h), \sigma(h+1)) \text{ kein Fehlstandspaar von } \sigma \text{ ist.} \end{cases} \qquad (*)$$

Dazu vergleichen wir die Fehlstandspaare der Permutationen

$$\begin{aligned} \sigma &= \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix} \quad \text{und} \\ \varrho &= \tau \circ \sigma = \begin{pmatrix} 1 & \cdots & h-1 & h & h+1 & \cdots & n \\ \sigma(1) & \cdots & \sigma(h-1) & \sigma(h+1) & \sigma(h) & \cdots & \sigma(n) \end{pmatrix}. \end{aligned}$$

Ist $i \notin \{h, h+1\}$, so ist (i, h) genau dann ein Fehlstandspaar von σ, wenn $(i, h+1)$ ein Fehlstandspaar von ϱ ist. Ebenso ist (i, h) genau dann ein Fehlstandspaar von ϱ, wenn $(i, h+1)$ ein Fehlstandspaar von σ ist. Unter den Paaren $(i, j) \neq (h, h+1)$ haben ϱ und σ daher dieselbe Anzahl von Fehlstandspaaren. Ist $(h, h+1)$ ein Fehlstandspaar von σ, so ist es keines von ϱ. Ist $(h, h+1)$ kein Fehlstandspaar von σ, so ist eines von ϱ. Damit ist $(*)$ gezeigt.
Wir führen nun den Beweis von (i) durch Induktion nach der Zahl m der Fehlstandspaare. Ist $\phi(\sigma) = 0$, so muss $\sigma_1 \leq \sigma_2 \leq \ldots \leq \sigma(n)$ sein, also

$$\sigma = id = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}.$$

Die Identität kann als leeres Produkt von Transpositionen geschrieben werden, oder auch als $id = (1, 2) \circ (1, 2)$.
Nun sei die Behauptung (i) schon für die Fehlstandszahl $m-1$ bewiesen.
Es sei $\sigma \in S_n$ mit $\phi(\sigma) = m$ mit einem Fehlstandspaar $(h, h+1)$ sowie $\tau = (\sigma(h), \sigma(h+1))$. Wegen $(*)$ ist dann $\phi(\sigma \circ \tau) = m-1$. Nach Induktionsannahme ist $\tau \circ \sigma = \tau \circ \tau_1 \circ \ldots \circ \tau_k = \varrho$ für gewisse Transpositionen τ_i und somit wegen $\tau \circ \tau = id$

$$\sigma = \tau \circ \varrho = \tau \circ \tau_1 \circ \ldots \circ \tau_k.$$

ii) Sei $\tau = (i, j) \in S_n$ mit $i < j$. Die Vertauschung von i und j kann folgendermaßen erreicht werden: Durch $(j-1)$ Vertauschungen benachbarter Elemente wird i an die j-te Stelle, und j an die $(j-1)$-te Stelle gebracht. Danach wird j durch $(j-i-1)$ Vertauschungen mit benachbarten Elementen an die i-te Stelle gebracht.

Skizze:

$$
\begin{aligned}
\begin{pmatrix} 1 & 2 & \cdots & i & \cdots & j & \cdots & n \\ 1 & 2 & \cdots & i & \cdots & j & \cdots & n \end{pmatrix}
&\overset{(i\ i+1)}{\rightarrow} \begin{pmatrix} 1 & 2 & \cdots & i & i+1 & \cdots & j & \cdots & n \\ 1 & 2 & \cdots & i+1 & i & \cdots & j & \cdots & n \end{pmatrix} \\
&\overset{(i+1\ i+2)}{\rightarrow} \begin{pmatrix} 1 & 2 & \cdots & i & i+1 & i+2 & \cdots & j & \cdots & n \\ 1 & 2 & \cdots & i+1 & i+2 & i & \cdots & j & \cdots & n \end{pmatrix} \\
&\rightarrow \begin{pmatrix} 1 & 2 & \cdots & i & i+1 & \cdots & j-1 & j & \cdots & n \\ 1 & 2 & \cdots & i+1 & i+2 & \cdots & i & j & \cdots & n \end{pmatrix} \\
&\overset{(j-1\ j)}{\rightarrow} \begin{pmatrix} 1 & 2 & \cdots & i & i+1 & \cdots & j-1 & j & \cdots & n \\ 1 & 2 & \cdots & i+1 & i+2 & \cdots & j & i & \cdots & n \end{pmatrix} \\
&\overset{(j-1\ j-2)}{\rightarrow} \begin{pmatrix} 1 & 2 & \cdots & j-2 & j-1 & j & \cdots & n \\ 1 & 2 & \cdots & j & j-1 & i & \cdots & n \end{pmatrix} \\
&\rightarrow \begin{pmatrix} 1 & 2 & \cdots & i & \cdots & j & \cdots & n \\ 1 & 2 & \cdots & j & \cdots & i & \cdots & n \end{pmatrix}.
\end{aligned}
$$

Wir haben also die Produktdarstellung

$$\tau = (i, i+1) \circ \cdots \circ (j-1, j-2) \circ (j-1, j) \circ \ldots \circ (i+1, i+2) \circ (i, i+1),$$

ein Produkt von $2(j-i)-1$ Transpositionen der Form $(\sigma(h), \sigma(h+1))$, also einer ungeraden Anzahl solcher Transpositionen. Sei nun $\sigma = \tau_1 \circ \cdots \circ \tau_r$. Jedes τ_i schreiben wir, wie eben skizziert, als Produkt einer ungeraden Anzahl von Transpositionen der Form $(\sigma(h), \sigma(h+1))$. Dies führt zu einer Produktdarstellung $\sigma = \tilde{\tau}_1 \circ \cdots \circ \tilde{\tau}_s$, wobei jedes $\tilde{\tau}_j$ von der Form $(\sigma(h), \sigma(h+1))$ ist. s ist gerade falls r gerade ist, bzw. ungerade falls r ungerade ist. Nach $(*)$ unterscheiden sich $\phi(\varrho)$ und $\phi(\tilde{\tau} \circ \varrho)$ um ± 1. Daher ist $\phi(\sigma)$ gerade, falls s und somit r gerade ist, und andernfalls ungerade. Daraus folgt (ii).

iii) Sei $\sigma = \tau_1 \circ \cdots \circ \tau_r$ und $\varrho = \nu_1 \circ \cdots \circ \nu_s$ mit Transpositionen τ_i, ν_j. Dann ist

$$\sigma \circ \varrho = \tau_1 \circ \ldots \circ \tau_r \circ \nu_1 \circ \ldots \circ \nu_s.$$

Nach (ii) gilt dann $\operatorname{sgn}(\sigma) = (-1)^r$, $\operatorname{sgn}(\varrho) = (-1)^s$ sowie $\operatorname{sgn}(\sigma \circ \varrho) = (-1)^{r+s}$, woraus die Behauptung $\operatorname{sgn}(\sigma \circ \varrho) = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\varrho)$ folgt.

$\square$

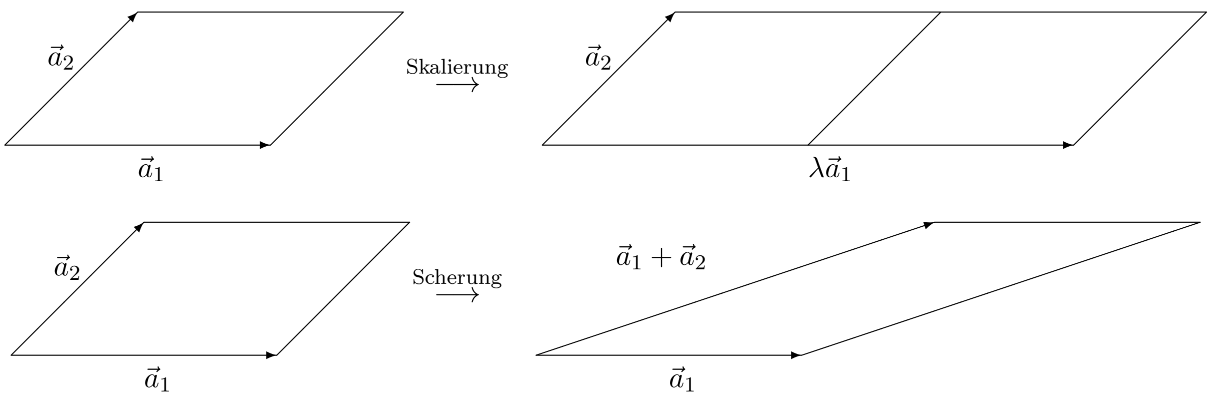
5.2 Determinantenfunktionen

Eine Motivation für die Einführung einer Determinantenfunktion liegt im Wunsch, das Volumen eines von n Vektoren $\vec{a}_1, \ldots, \vec{a}_n$ des $\mathbb{R}^n$ aufgespannten Parallelepipeds zu definieren. Wir beschränken uns bei der Illustration der Einfachheit halber auf den Fall $n = 2$, also den Flächeninhalt eines Parallelogramms. Es sei $|f(\vec{a}_1, \vec{a}_2)|$ die Fläche des von $\vec{a}_1, \vec{a}_2 \in \mathbb{R}^2$ aufgespannten Parallelogramms. Dann gilt offenbar

(H) Homogenität: $f(\lambda\vec{a}_1, \vec{a}_2) = f(\vec{a}_1, \lambda\vec{a}_2) = \lambda f(\vec{a}_1, \vec{a}_2)$ für alle $\lambda \in \mathbb{R}$.

(S) Scherungsinvarianz: $f(\vec{a}_1 + \vec{a}_2, \vec{a}_2) = f(\vec{a}_1, \vec{a}_1 + \vec{a}_2) = f(\vec{a}_1, \vec{a}_2)$.

Skizze:



Wir verallgemeinern diese Eigenschaften nun auf beliebige Abbildungen $f\colon V^n \to K$, wobei V ein n-dimensionaler Vektorraum über einem Körper K und $V^n = V \times V \times \ldots \times V$ ist.

Definition 5.2.1. Eine Abbildung $f\colon V^n \to K$ heißt homogen, falls für alle $\vec{a}_1, \ldots, \vec{a}_n \in V$ und $i = 1, \ldots, n$

$$f(\vec{a}_1, \ldots, \vec{a}_{i-1}, \lambda\vec{a}_i, \vec{a}_{i+1}, \ldots, \vec{a}_n) = \lambda f(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_n)$$

gilt. Die Abbildung f heißt scherungsinvariant, falls für alle i, k mit $i \neq k$ und alle $\vec{a}_1, \ldots, \vec{a}_n \in V$

$$f(\vec{a}_1, \ldots, \vec{a}_{i-1}, \vec{a}_i + \vec{a}_k, \vec{a}_{i+1}, \ldots, \vec{a}_n) = f(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_n)$$

gilt.

Im folgenden sei V ein K-Vektorraum.

Satz 5.2.1. *Es sei* $\dim V = n$ *und* $f\colon V^n \to K$ *eine homogene und scherungsinvariante Abbildung. Dann gilt:*

i) $f(\vec{a}_1, \ldots, \vec{a}_{i-1}, \vec{0}, \vec{a}_{i+1}, \ldots, \vec{a}_n) = 0.$

ii) Für alle $i \neq k$ *und alle* $\lambda \in K$ *ist* $f(\vec{a}_1, \ldots, \vec{a}_i + \lambda\vec{a}_k, \ldots, \vec{a}_n) = f(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_n)$.

iii) Sind $\vec{a}_1, \ldots, \vec{a}_n$ *linear abhängig, so ist* $f(\vec{a}_1, \ldots, \vec{a}_n) = 0$.

iv) Die Abbildung f *ist in jedem Argument additiv, d.h. für alle* i *gilt*

$$f(\vec{a}_1, \ldots, \vec{a}_i + \vec{b}_i, \ldots \vec{a}_n) = f(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_n) + f(\vec{a}_1, \ldots, \vec{b}_i, \ldots, \vec{a}_n).$$

v) Die Abbildung f *ist in jedem Argumentpaar alternierend, d.h. für alle* $i \neq j$ *ist*

$$f(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_j, \ldots, \vec{a}_n) = -f(\vec{a}_1, \ldots, \vec{a}_j, \ldots, \vec{a}_i, \ldots, \vec{a}_n).$$

Beweis. i) Wegen der Homogenität gilt

$$f(\vec{a}_1, \ldots, \vec{0}, \ldots, \vec{a}_n) = f(\vec{a}_1, \ldots, 0 \cdot \vec{0}, \ldots, \vec{a}_n) \underset{(\mathrm{H})}{=} 0 \cdot f(\vec{a}_1, \ldots, \vec{0}, \ldots, \vec{a}_n) = 0.$$

ii) Es sei ohne Einschränkung $\lambda \neq 0$. Dann gilt

$$\begin{aligned} f(\vec{a}_1,\dots,\vec{a}_i,\dots,\vec{a}_j,\dots,\vec{a}_n) &\underset{(H)}{=} \lambda^{-1} f(\vec{a}_1,\dots,\vec{a}_i,\dots,\lambda\vec{a}_j,\dots,\vec{a}_n) \\ &\underset{(S)}{=} \lambda^{-1} f(\vec{a}_1,\dots,\vec{a}_i+\lambda\vec{a}_j,\dots,\lambda\vec{a}_j,\dots,\vec{a}_n) \\ &\underset{(H)}{=} f(\vec{a}_1,\dots,\vec{a}_i+\lambda\vec{a}_j,\dots,\vec{a}_j,\dots,\vec{a}_n). \end{aligned}$$

iii) Sind $\vec{a}_1,\dots,\vec{a}_n$ linear abhängig, so ist $\lambda_1\vec{a}_1+\dots+\lambda_n\vec{a}_n=\vec{0}$ für $\lambda_1,\dots,\lambda_n \in K$, und mindestens ein λ_i nicht null; ohne Einschränkung sei dies $\lambda_1 \neq 0$. Dann erhält man

$$\begin{aligned} \lambda_1 f(\vec{a}_1,\dots,\vec{a}_n) &\underset{(H)}{=} f(\lambda_1\vec{a}_1,\dots,\vec{a}_n) \\ &\underset{(ii)}{=} f(\lambda_1\vec{a}_1+\lambda_2\vec{a}_2,\vec{a}_2,\dots,\vec{a}_n) \\ &\ \vdots \\ &\underset{(ii)}{=} f(\lambda_1\vec{a}_1+\dots+\lambda_n\vec{a}_n,\vec{a}_2,\dots,\vec{a}_n) \ = \ f(\vec{0},\vec{a}_2,\dots,\vec{a}_n)=0 \end{aligned}$$

und somit $f(\vec{a}_1,\dots,\vec{a}_n)=0$.

iv) Wir unterscheiden hier drei Fälle:

- $\vec{a}_1,\dots,\vec{a}_n$ sind linear unbhängig:
  Dann ist $\{\vec{a}_1,\dots,\vec{a}_n\}$ eine Basis von V, also $\vec{b}_i=\lambda_1\vec{a}_1+\dots+\lambda_n\vec{a}_n$ mit $\lambda_i \in K$. Es folgt

$$\begin{aligned} f(\vec{a}_1,\dots,\vec{a}_i+\vec{b}_i,\dots,\vec{a}_n) &= f(\vec{a}_1,\dots,\vec{a}_i+\lambda_1\vec{a}_1+\dots+\lambda_i\vec{a}_i+\dots+\lambda_n\vec{a}_n,\dots,\vec{a}_n) \\ &\underset{(ii)}{=} f(\vec{a}_1,\dots,\vec{a}_i+\lambda_i\vec{a}_i,\dots,\vec{a}_n) \\ &\underset{(H)}{=} (1+\lambda_i) f(\vec{a}_1,\dots,\vec{a}_i,\dots,\vec{a}_n) \\ &= f(\vec{a}_1,\dots,\vec{a}_i,\dots,\vec{a}_n)+\lambda_i f(\vec{a}_1,\dots,\vec{a}_i,\dots,\vec{a}_n) \\ &\underset{(H)}{=} f(\vec{a}_1,\dots,\vec{a}_i,\dots,\vec{a}_n)+f(\vec{a}_1,\dots,\lambda_i\vec{a}_i,\dots,\vec{a}_n) \\ &\underset{(ii)}{=} f(\vec{a}_1,\dots,\vec{a}_i,\dots,\vec{a}_n) \\ &\quad +f(\vec{a}_1,\dots,\lambda_1\vec{a}_1+\dots+\lambda_i\vec{a}_i+\dots+\lambda_n\vec{a}_n,\dots,\vec{a}_n) \\ &= f(\vec{a}_1,\dots,\vec{a}_i,\dots,\vec{a}_n)+f(\vec{a}_1,\dots,\vec{b}_i,\dots,\vec{a}_n). \end{aligned}$$

- $\vec{a}_1,\dots,\vec{a}_{i-1},\vec{a}_{i+1},\dots,\vec{a}_n$ sind linear unabhängig, aber $\vec{a}_1,\dots,\vec{a}_n$ sind linear abhängig:
  Dann ist $\vec{a}_i$ eine Linearkombination von $\vec{a}_1,\dots,\vec{a}_{i-1},\vec{a}_{i+1},\dots,\vec{a}_n$, also gilt

$$f(\vec{a}_1,\dots,\vec{a}_i+\vec{b}_i,\dots,\vec{a}_n) \underset{(ii)}{=} f(\vec{a}_1,\dots,\vec{b}_i,\dots,\vec{a}_n)=f(\vec{a}_1,\dots,\vec{a}_i,\dots,\vec{a}_n)+f(\vec{a}_1,\dots,\vec{b}_i,\dots,\vec{a}_n),$$

  da nach (iii) $f(\vec{a}_1,\dots,\vec{a}_i,\dots,\vec{a}_n)=0$ ist.
- $\vec{a}_1,\dots,\vec{a}_{i-1},\vec{a}_{i+1},\dots,\vec{a}_n$ sind linear abhängig:
  In diesem Fall ist die Aussage (iv) richtig, weil alle f-Werte nach (iii) verschwinden.

v) Es gilt

$$\begin{aligned} f(\vec{a}_1,\dots,\vec{a}_i,\dots,\vec{a}_j,\dots,\vec{a}_n) &\underset{(H)}{=} -f(\vec{a}_1,\dots,\vec{a}_i,\dots,-\vec{a}_j,\dots,\vec{a}_n) \\ &\underset{(ii)}{=} -f(\vec{a}_1,\dots,\vec{a}_i+\vec{a}_j,\dots,-\vec{a}_j,\dots,\vec{a}_n) \\ &\underset{(ii)}{=} -f(\vec{a}_1,\dots,\vec{a}_i+\vec{a}_j,\dots,\vec{a}_i,\dots,\vec{a}_n) \\ &\underset{(ii)}{=} -f(\vec{a}_1,\dots,\vec{a}_j,\dots,\vec{a}_i,\dots,\vec{a}_n). \end{aligned}$$

□

Definition 5.2.2. Es sei $\dim V = n$ und $f\colon V^n \to K$ eine Abbildung.

i) Dann heißt f n-fache Linearform, wenn f in jedem Argument linear ist, d.h. wenn

$$f(\vec{a}_1, \ldots, \lambda\vec{a}_i + \mu\vec{b}_i, \ldots, \vec{a}_n) = \lambda f(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_n) + \mu f(\vec{a}_1, \ldots, \vec{b}_i, \ldots, \vec{a}_n)$$

für alle $\vec{a}_1, \ldots, \vec{a}_n, \vec{b}_i \in V$ und $\lambda, \mu \in K$ gilt.

ii) Weiter heißt f alternierend, falls für alle $i \neq j$ und alle $\vec{a}_1, \ldots, \vec{a}_n \in V$

$$f(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_j, \ldots, \vec{a}_n) = -f(\vec{a}_1, \ldots, \vec{a}_j, \ldots, \vec{a}_i, \ldots, \vec{a}_n)$$

gilt.

Satz 5.2.2. *Es sei $\dim V = n$ und $f\colon V^n \to K$ eine Abbildung. Dann sind folgende Aussagen äquivalent:*

i) f ist homogen und scherungsinvariant.

ii) f ist eine n-fache alternierende Linearform.

Beweis. $(i) \Rightarrow (ii)$:
Aus der Homogenität und der Additivität (Satz 5.2.1 (iv)) folgt die Linearität. Nach 5.2.1 (v) ist f auch alternierend.
$(ii) \Rightarrow (i)$:
Es sei f eine n-fache alternierende Linearform. Dann ist

$$f(\vec{a}_1, \ldots, \vec{a}_i + \vec{a}_j, \ldots, \vec{a}_n) = f(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_j, \ldots, \vec{a}_n) + f(\vec{a}_1, \ldots, \vec{a}_j, \ldots, \vec{a}_j, \ldots, \vec{a}_n).$$

Da f alternierend ist, folgt $f(\vec{a}_1, \ldots, \vec{a}_j, \ldots, \vec{a}_j, \ldots, \vec{a}_n) = 0$ und damit die Scherungsinvarianz von f. Die Homogenität von f ergibt sich unmittelbar aus der Linearität. □

Vertauschung zweier Argumente einer alternierenden Linearform ändert den Wert von f um den Faktor -1. Wir wollen nun untersuchen, wie sich der Wert bei einer beliebigen Umordnung der Argumente verhält:

Satz 5.2.3. *Es sei f eine n-fache alternierende Linearform auf V und $\sigma \in S_n$. Für alle $\vec{a}_1, \ldots, \vec{a}_n \in V$ ist dann*

$$f(\vec{a}_{\sigma(1)}, \ldots, \vec{a}_{\sigma(n)}) = \operatorname{sgn}(\sigma) \cdot f(\vec{a}_1, \ldots, \vec{a}_n).$$

Beispiel 5.2.1. Es sei f eine 3-fache alternierende Linearform und $\dim V = 3$. Für $\vec{a}_1, \vec{a}_2, \vec{a}_3 \in V$ ist

$$f(\vec{a}_2, \vec{a}_3, \vec{a}_1) = f(\vec{a}_{\sigma(1)}, \vec{a}_{\sigma(2)}, \vec{a}_{\sigma(3)}) \quad \text{mit} \quad \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

Die Permutation σ hat die Fehlstandspaare $(1,3)$ und $(2,3)$, also $\operatorname{sgn}(\sigma) = 1$. Daher ist

$$f(\vec{a}_2, \vec{a}_3, \vec{a}_1) = \operatorname{sgn}(\sigma) f(\vec{a}_1, \vec{a}_2, \vec{a}_3) \;=\; f(\vec{a}_1, \vec{a}_2, \vec{a}_3).$$

Beweis. (Beweis von Satz 5.2.3)
Wir schreiben σ als Produkt von Transpositionen $\sigma = \tau_1 \circ \cdots \circ \tau_r$, was nach Satz 5.2.1 möglich ist. Für $1 \leq j \leq r$ setzen wir $\sigma_j = \tau_j \circ \cdots \circ \tau_r$. Nun ist

$$f(\vec{a}_{\sigma(1)}, \vec{a}_{\sigma(2)}, \ldots, \vec{a}_{\sigma(n)}) = f(\vec{a}_{(\tau_1 \circ \sigma_2)(1)}, \vec{a}_{(\tau_1 \circ \sigma_2)(2)}, \ldots, \vec{a}_{(\tau_1 \circ \sigma_2)(n)}) = -f(\vec{a}_{\sigma_2(1)}, \vec{a}_{\sigma_2(2)}, \ldots, \vec{a}_{\sigma_2(n)}),$$

da f alternierend ist, und τ_1 zwei Argumente tauscht. Wiederholtes Abspalten der τ_j ergibt

$$\begin{aligned} f(\vec{a}_{\sigma(1)}, \vec{a}_{\sigma(2)}, \ldots, \vec{a}_{\sigma(n)}) &= (-1)^1 f(\vec{a}_{\sigma_2(1)}, \vec{a}_{\sigma_2(2)}, \ldots, \vec{a}_{\sigma_2(n)}) \\ &= (-1)^2 f(\vec{a}_{\sigma_3(1)}, \vec{a}_{\sigma_3(2)}, \ldots, \vec{a}_{\sigma_3(n)}) \\ &\vdots \\ &= (-1)^r f(\vec{a}_{\sigma_n(1)}, \ldots, \vec{a}_{\sigma_n(n)}) = (-1)^r f(\vec{a}_1, \ldots, \vec{a}_n). \end{aligned}$$

Nach Satz 5.2.1 ist $(-1)^r = \phi(\sigma)$. □

Satz 5.2.4. *Es sei f eine n-fache alternierende Linearform auf V. Dann gilt:*

i) Sind $\vec{a}_1, \ldots, \vec{a}_n, \vec{b}_1, \ldots, \vec{b}_n \in V$ und $\lambda_{ik} \in K$ mit $\vec{a}_i = \sum_{k=1}^{n} \lambda_{ik} \vec{b}_k$, dann ist

$$f(\vec{a}_1, \ldots, \vec{a}_n) = f(\vec{b}_1, \ldots, \vec{b}_n) \cdot \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \lambda_{1,\sigma(1)} \cdots \lambda_{n,\sigma(n)}.$$

ii) Ist $f \neq 0$, so gilt

$$f(\vec{b}_1, \ldots, \vec{b}_n) \neq 0 \Leftrightarrow \vec{b}_1, \ldots, \vec{b}_n \textit{ sind linear unabhängig.}$$

Beweis. i) Wir wenden wiederholt die Linearität von f an:

$$\begin{aligned} f(\vec{a}_1, \ldots, \vec{a}_n) &= f\left(\sum_{k=1}^{n} \lambda_{1,k} \vec{b}_k, \vec{a}_2, \ldots, \vec{a}_n\right) = \sum_{k=1}^{n} \lambda_{1,k} f(\vec{b}_k, \vec{a}_2, \ldots, \vec{a}_n) \\ &= \sum_{k_1=1}^{n} \lambda_{1,k_1} f\left(\vec{b}_{k_1}, \sum_{k_2=1}^{n} \lambda_{2,k_2} \vec{b}_{k_2}, \vec{a}_3, \ldots, \vec{a}_n\right) \\ &= \sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \lambda_{1,k_1} \lambda_{2,k_2} f(\vec{b}_{k_1}, \vec{b}_{k_2}, \vec{a}_3, \ldots, \vec{a}_n) \\ &\vdots \\ &= \sum_{k_1=1}^{n} \cdots \sum_{k_n=1}^{n} \lambda_{1,k_1} \cdots \lambda_{n,k_n} f(\vec{b}_{k_1}, \vec{b}_{k_2}, \ldots, \vec{b}_{k_n}). \end{aligned}$$

Da f alternierend ist, sind alle Summanden null, für die zwei k_i übereinstimmen. Wir brauchen also nur über die n-Tupel $(k_1, \ldots, k_n)$ zu summieren, für die $(k_1, \ldots, k_n) = (\sigma(1), \ldots, \sigma(n))$ für ein $\sigma \in S_n$ ist. Also ist

$$\begin{aligned} f(\vec{a}_1, \ldots, \vec{a}_n) &= \sum_{\sigma \in S_n} \lambda_{1,\sigma(1)} \cdots \lambda_{n,\sigma(n)} f(\vec{b}_{\sigma(1)}, \vec{b}_{\sigma(2)}, \ldots, \vec{b}_{\sigma(n)}) \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \lambda_{1,\sigma(1)} \cdots \lambda_{n,\sigma(n)} f(\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n) \end{aligned}$$

nach Satz 5.2.3.

ii) Sind $\vec{b}_1, \ldots, \vec{b}_n$ linear unabhängig, so bilden sie eine Basis von V. Für beliebige $\vec{a}_1, \ldots, \vec{a}_n \in V$ gibt es dann $\lambda_{ik} \in K$ für $1 \leq i, k \leq n$ mit $\vec{a}_i = \sum \lambda_{ik} \vec{b}$. Aus (i) folgt dann

$$f(\vec{a}_1, \ldots, \vec{a}_n) = f(\vec{b}_1, \ldots, \vec{b}_n) \cdot \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \lambda_{1,\sigma(1)} \cdots \lambda_{n,\sigma(n)}.$$

Wäre $f(\vec{b}_1, \ldots, \vec{b}_n) = 0$, so gilt auch $f(\vec{a}_1, \ldots, \vec{a}_n) = 0$. Da die $\vec{a}_1, \ldots, \vec{a}_n \in V$ aber beliebig sind, ist f bereits die Nullabbildung. Sind nun andererseits $\vec{b}_1, \ldots, \vec{b}_n$ linear abhängig, so ist nach Satz 5.2.1 (iii) dann $f(\vec{b}_1, \ldots, \vec{b}_n) = 0$.

□

Satz 5.2.5. *Es sei* $\dim V = n$. *Dann bildet die Menge der n-fachen alternierenden Linearformen auf V mit werteweiser Addition und Skalarmultiplikation einen K-Vektorraum der Dimension* 1.

Beweis. Es sei $\mathcal{L}$ die Menge aller n-fachen Linearformen auf V. Man rechnet leicht nach, dass mit $f, g \in \mathcal{L}$ auch $\lambda f + \mu g \in \mathcal{L}$ ist. Weiter ist $\mathcal{L}$ auch nicht leer, da die Nullform mit $N(\vec{a}_1, \ldots, \vec{a}_n) = 0$ für alle $\vec{a}_1, \ldots, \vec{a}_n \in V$ stets in $\mathcal{L}$ liegt. Nach dem Unterraumkriterium ist $\mathcal{L}$ damit ein Untervektorraum von $\operatorname{Abb}(V^n, K)$. Die Abbildung $f_0 \colon V^n \to K$ werde folgendermaßen definiert: Es sei $\{\vec{b}_1, \ldots, \vec{b}_n\}$ eine fest gewählte Basis von V und

$$f_0(\vec{a}_1, \ldots, \vec{a}_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \lambda_{1,\sigma(1)} \cdots \lambda_{n,\sigma(n)}, \ \vec{a}_i = \sum_{k=1}^{n} \lambda_{ik} \vec{b}_k.$$

Wir zeigen, dass f_0 ein nichttriviales Element von $\mathcal{L}$ ist:

- f_0 ist n-fache Linearform:
  Es seien $\vec{a}_1, \ldots, \vec{a}_n \in V$ wie oben gegeben, und es sei für $1 \leq i \leq n$

$$\vec{a}_i' = \sum_{k=1}^{n} \lambda_{ik}' \vec{b}_k.$$

  Dann gilt für $\mu, \nu \in K$

$$\begin{aligned} f_0(\vec{a}_1, \ldots, \vec{a}_{i-1}, \mu \vec{a}_i + \nu \vec{a}_i', \vec{a}_{i+1}, \ldots, \vec{a}_n) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \lambda_{1,\sigma(1)} \cdots (\mu \lambda_{i,\sigma(i)} + \nu \lambda_{i,\sigma(i)}') \cdots \lambda_{n,\sigma(n)} \\ &= \mu \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \lambda_{1,\sigma(1)} \cdots \lambda_{i,\sigma(i)} \cdots \lambda_{n,\sigma(n)} \\ &\quad + \nu \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \lambda_{1,\sigma(1)} \cdots \lambda_{i,\sigma(i)}' \cdots \lambda_{n,\sigma(n)} \\ &= \mu f_0(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_n) + \nu f_0(\vec{a}_1, \ldots, \vec{a}_i', \ldots, \vec{a}_n). \end{aligned}$$

- f_0 ist alternierend:
  Für $i \neq j$ betrachte man

$$f_0(\vec{a}_1, \ldots, \vec{a}_j, \ldots, \vec{a}_i, \ldots, \vec{a}_n) = f_0(\vec{a}_{\tau(1)}, \ldots, \vec{a}_{\tau(i)}, \ldots, \vec{a}_{\tau(j)}, \ldots, \vec{a}_{\tau(n)})$$

  mit der Transposition $\tau = (i, j)$. Es ist

$$\begin{aligned} f_0(\vec{a}_1, \ldots, \vec{a}_j, \ldots, \vec{a}_i, \ldots, \vec{a}_n) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \lambda_{\tau(1),\sigma(1)} \cdots \lambda_{\tau(n),\sigma(n)} \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \lambda_{\tau(1),(\sigma\circ\tau)(\tau(1))} \cdots \lambda_{\tau(n),(\sigma\circ\tau)(\tau(n))} \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \lambda_{1,(\sigma\circ\tau)(1)} \cdots \lambda_{n,(\sigma\circ\tau)(n)}. \end{aligned}$$

Wegen der Gruppeneigenschaft von S_n durchläuft $\tilde{\sigma} = \sigma \circ \tau$ alle Permutationen aus S_n wenn σ dies tut. Die Abbildung $\sigma \mapsto \tilde{\sigma}$ ist nämlich bijektiv: ist $\sigma \circ \tau = \sigma' \circ \tau$, so folgt nach Multiplikation mit τ von rechts $\sigma = \sigma'$ (Injektivität), und für $\tilde{\sigma} \in S_n$ ist $\sigma = \tilde{\sigma} \circ \tau$ wegen $\tau \circ \tau = id$ stets ein Urbild (Surjektivität). Also kann man die Summe umindizieren, es folgt

$$\begin{aligned}
f_0(\vec{a}_1, \ldots, \vec{a}_j, \ldots, \vec{a}_i, \ldots, \vec{a}_n) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \lambda_{1,(\sigma\circ\tau)(1)} \cdots \lambda_{n,(\sigma\circ\tau)(n)} \\
&= \sum_{\tilde{\sigma} \in S_n} \operatorname{sgn}(\tilde{\sigma} \circ \tau) \lambda_{1,\tilde{\sigma}(1)} \cdots \lambda_{n,\tilde{\sigma}(n)} \\
&\underset{\text{Satz 5.2.1}}{=} -\sum_{\tilde{\sigma} \in S_n} \operatorname{sgn}(\tilde{\sigma}) \lambda_{1,\tilde{\sigma}(1)} \cdots \lambda_{n,\tilde{\sigma}(n)} \\
&= -f_0(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_j, \ldots, \vec{a}_n).
\end{aligned}$$

Also ist f_0 alternierend.

- f_0 ist nicht die Nullform:
  Für die spezielle Wahl $\vec{a}_i = \vec{b}_i$ für $i = 1, \ldots, n$ ist $\lambda_{ik} = \delta_{ik}$ das Kroneckersymbol, und wir erhalten
  $$f_0(\vec{b}_1, \ldots, \vec{b}_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \delta_{1,\sigma(1)} \cdots \delta_{n,\sigma(n)}.$$
  Ist σ nicht die Identität, so gibt es ein i mit $\sigma(i) \neq i$ und der zugehörige Summand verschwindet wegen $\delta_{i,\sigma(i)} = 0$. Nur der Summand für $\sigma = id$ bleibt mit $\delta_{1,1} = \cdots = \delta_{n,n} = 1$ übrig. Also ist $f_0(\vec{b}_1, \ldots, \vec{b}_n) = 1 \neq 0$.

Ist nun $f \in \mathcal{L}$ beliebig, so ist nach Satz 5.2.4 für beliebige Vektoren $\vec{a}_1, \ldots, \vec{a}_n \in V$

$$f(\vec{a}_1, \ldots, \vec{a}_n) = f(\vec{b}_1, \ldots, \vec{b}_n) \cdot \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \lambda_{1,\sigma(1)} \cdots \lambda_{n,\sigma(n)} = f(\vec{b}_1, \ldots, \vec{b}_n) \cdot f_0(\vec{a}_1, \ldots, \vec{a}_n),$$

wobei $\mu = f(\vec{b}_1, \ldots, \vec{b}_n) \in K$ ein von f abhängiges Skalar ist, also ist $f = \mu \cdot f_0$, und es folgt insgesamt $\mathcal{L} = \langle\langle\, f_0 \,\rangle\rangle$ und $\dim \mathcal{L} = 1$. □

Definition 5.2.3. Eine Abbildung $D\colon V^n \to K$ heißt Determinantenfunktion, falls D eine n-fache alternierende Linearform, aber nicht die Nullform ist.

Satz 5.2.6. *Es sei* $\dim V = n$ *und* $\{\vec{b}_1, \ldots, \vec{b}_n\}$ *eine Basis von* V. *Dann gibt es für jedes* $d \in K\backslash\{0\}$ *genau eine Determinantenfunktion* $D\colon V^n \to K$ *mit* $D(\vec{b}_1, \ldots, \vec{b}_n) = d$.

Beweis. Es sei $f_0 \in \mathcal{L}$ mit $f_0 \neq 0$ beliebig, was nach Satz 5.2.5 auch existiert. Dann ist nach Satz 5.2.4 (ii) speziell $f_0(\vec{b}_1, \ldots, \vec{b}_n) = \eta \neq 0$, da $\vec{b}_1, \ldots, \vec{b}_n$ linear unabhängig sind. Wir setzen $D = d\eta^{-1} f_0 \in \mathcal{L}$. Offenbar ist dann $D(\vec{b}_1, \ldots, \vec{b}_n) = d$. Es sei nun $\tilde{D} \in \mathcal{L}$ mit $\tilde{D}(\vec{b}_1, \ldots, \vec{b}_n) = d$ gegeben. Weiter ist nach Satz 5.2.5 dann $\tilde{D} = \mu f_0$ für ein $\mu \in K$, also $\tilde{D}(\vec{b}_1, \ldots, \vec{b}_n) = \mu f_0(\vec{b}_1, \ldots, b_n) = \mu\eta$. Es folgt $\mu\eta = d$ bzw. $\mu = d\eta^{-1}$, also $\tilde{D} = d\eta^{-1} f_0 = D$. □

Satz 5.2.7. *Ist* $D\colon V^n \to K$ *eine Determinantenfunktion auf* V, *so gilt*

$$\vec{a}_1, \ldots, \vec{a}_n \text{ linear unabhängig} \Leftrightarrow D(\vec{a}_1, \ldots, \vec{a}_n) \neq 0.$$

Beweis. Dies folgt unmittelbar aus Satz 5.2.4 (ii). □

5.3 Die natürliche Determinantenfunktion des K^n

Definition 5.3.1. Es sei $\{\vec{e}_1, \dots, \vec{e}_n\}$ die Standardbasis des $V = K^n$ (geschrieben als Zeilenvektoren). Dann verstehen wir unter der <u>natürlichen Determinantenfunktion</u> des K^n die nach Satz 5.2.6 eindeutig bestimmte Determinantenfunktion $D_n \colon V^n \to K$ mit $D_n(\vec{e}_1, \dots, \vec{e}_n) = 1$.

Satz 5.3.1. *(Leibniz-Formel)*
Es seien $\vec{a}_1, \dots, \vec{a}_n \in K^n$ mit

$$\vec{a}_i = (a_{i1} \cdots a_{in}) = \sum_{k=1}^{n} a_{ik}\vec{e}_k$$

gegeben. Dann ist

$$D_n(\vec{a}_1, \dots, \vec{a}_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}.$$

Beweis. Nach Satz 5.2.4 ist

$$D_n(\vec{a}_1, \dots, \vec{a}_n) = D_n(\vec{e}_1, \dots, \vec{e}_n) \cdot \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$$

mit $D_n(\vec{e}_1, \dots, \vec{e}_n) = 1$. $\square$

Definition 5.3.2. Es sei $\mathcal{A} = (a_{ik}) \in K^{(n,n)}$. Dann ist die <u>Determinante</u> von $\mathcal{A}$ der Wert

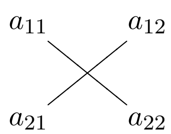
$$\det(A) = |A| = D_n(\vec{a}_1, \dots, \vec{a}_n)$$

mit den Zeilen $\vec{a}_i = (a_{i1} \cdots a_{in})$ von $\mathcal{A}$.

Beispiel 5.3.1. Im Fall $n = 2$ ist die Determinante nach der Leibniz-Formel durch

$$\det(\mathcal{A}) = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \sum_{\sigma \in S_2} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} = \underbrace{a_{11} \cdot a_{22}}_{\substack{\sigma = id \\ \mathrm{sgn}=1}} - \underbrace{a_{12} \cdot a_{21}}_{\substack{\sigma = (1,2) \\ \mathrm{sgn}=-1}}$$

gegeben. Dazu gehört das Schema:



Dabei ist das Produkt über die Querdiagonale mit dem Vorfaktor -1 zu versehen. Dann kann $|\det(\mathcal{A})|$ als der Flächeninhalt des von $\vec{a}_1 = (a_{11}\ a_{12})$ und $\vec{a}_2 = (a_{21}\ a_{22})$ aufgespannten Parallelogramms angesehen werden.

Beispiel 5.3.2. Im Fall $n = 3$ ist die Determinante

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \sum_{\sigma \in S_3} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} a_{3,\sigma(3)}.$$

Die sechs Elemente $(6 = 3!)$ von S_3 sind

$$\underset{\mathrm{sgn}=1}{\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}}, \underset{\mathrm{sgn}=-1}{\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}}, \underset{\mathrm{sgn}=-1}{\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}}, \underset{\mathrm{sgn}=1}{\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}}, \underset{\mathrm{sgn}=1}{\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}}, \underset{\mathrm{sgn}=-1}{\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}}.$$

Also ist

$$\det(\mathcal{A}) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}.$$

Dies kann schematisch durch die Sarrussche Regel dargestellt werden:

$$\begin{array}{ccccc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}$$

Dabei sind die Produkte über die Querdiagonalen mit dem Vorfaktor -1 zu versehen, und $|\det(\mathcal{A})|$ kann als das Volumen des von $\vec{a}_1 = (a_{11}\ a_{12}\ a_{13})$, $\vec{a}_2 = (a_{21}\ a_{22}\ a_{23})$ und $\vec{a}_3 = (a_{31}\ a_{32}\ a_{33})$ aufgespannten Parallelepipeds angesehen werden.

Satz 5.3.2. *Für $\mathcal{A} \in K^{(n,n)}$ gilt $\det(\mathcal{A}) = \det(\mathcal{A}^T)$.*

Beweis. Es sei $\mathcal{A} = (a_{ik})$ mit $1 \leq i, k \leq n$. Dann ist

$$\det(\mathcal{A}) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \text{ und } \det(\mathcal{A}^T) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n}.$$

Es ist $a_{1,\sigma(1)} \cdots a_{n,\sigma_n} = a_{\sigma^{-1}(1),1} \cdots a_{\sigma^{-1}(n),n}$, weswegen

$$\det(\mathcal{A}) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma^{-1}(1),1} \cdots a_{\sigma^{-1}(n),n} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma^{-1}) a_{\sigma^{-1}(1),1} \cdots a_{\sigma^{-1}(n),n}$$

gilt. Durchläuft σ alle Elemente von S_n, so auch $\varrho = \sigma^{-1}$, also ist nach Umindizierung der Summe

$$\det(\mathcal{A}) = \sum_{\varrho \in S_n} \operatorname{sgn}(\varrho) a_{\varrho(1),1} \cdots a_{\varrho(n),n} = \det(\mathcal{A}^T).$$

□

Satz 5.3.3. *Für Determinanten von Matrizen gelten folgende Regeln:*

i) Für alle $\lambda \in K$ gilt

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ \lambda a_{i1} & \cdots & \lambda a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \lambda \cdot \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$

sowie

$$\det \begin{pmatrix} a_{11} & \cdots & \lambda a_{1i} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & \lambda a_{ni} & \cdots & a_{nn} \end{pmatrix} = \lambda \cdot \det \begin{pmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{ni} & \cdots & a_{nn} \end{pmatrix}.$$

ii)

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1}+b_{i1} & \cdots & a_{in}+b_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} + \det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ b_{i1} & \cdots & b_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$

sowie

$$\begin{aligned}\det\begin{pmatrix} a_{11} & \cdots & a_{1i}+b_{1i} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{ni}+b_{ni} & \cdots & a_{nn} \end{pmatrix} \quad &= \quad \det\begin{pmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{ni} & \cdots & a_{nn} \end{pmatrix} \\ &\quad + \det\begin{pmatrix} a_{11} & \cdots & b_{1i} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & b_{ni} & \cdots & a_{nn} \end{pmatrix}.\end{aligned}$$

iii) Besitzt $\mathcal{A}$ eine Zeile oder Spalte, die nur aus Nullen besteht, so ist $\det(\mathcal{A}) = 0$.

iv) Die Determinante einer Matrix ändert sich nicht, wenn man zu einer Zeile (oder Spalte) das Vielfache einer Zeile (oder Spalte) addiert.

v) Die Determinante einer Matrix ändert beim Vertauschen zweier Zeilen (oder zweier Spalten) das Vorzeichen.

vi) Es gilt

$$\begin{aligned}\det(\mathcal{A}) \neq 0 \quad &\Leftrightarrow \quad \textit{die Zeilenvektoren sind linear unabhängig} \\ &\Leftrightarrow \quad \textit{die Spaltenvektoren sind linear unabhängig} \Leftrightarrow \mathcal{A} \textit{ ist regulär.}\end{aligned}$$

Beweis. Per Definition ist die Determinante einer Matrix eine n-fache alternierende Linearform der Zeilenvektoren der Matrix. Nach Satz 5.3.2 ist sie auch eine n-fache alternierende Linearform der Spaltenvektoren. Alle Aussagen folgen daher aus den entsprechenden Aussagen für alternierende Linearformen. □

Satz 5.3.4. *Es ist*

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1,n-1} & a_{1n} \\ \vdots & & \vdots & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & \cdots & 0 & 1 \end{pmatrix} = \det\begin{pmatrix} a_{11} & \cdots & a_{1,n-1} \\ \vdots & & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} \end{pmatrix}.$$

Beweis. Es sei $\mathcal{A} = (a_{ik})$. Dann ist nach nach Satz 5.3.3

$$\det(\mathcal{A}) \underset{\text{Satz 5.3.1}}{=} \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}.$$

Ist $\sigma \in S_n$ eine Permutation mit $\sigma(n) \neq n$, so ist $\sigma(j) = n$ mit $j < n$, so verschwindet der zugehörige Summand, weil er den Faktor $a_{j,\sigma(j)} = a_{j,n} = 0$ enthält. Die Determinante von $\mathcal{A}$ ist dann

$$\begin{aligned}\sum_{\substack{\sigma\in S_n\\ \sigma(n)=n}} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \quad &= \quad \sum_{\varrho\in S_{n-1}} \operatorname{sgn}(\varrho) a_{1,\varrho(1)} \cdots a_{n-1,\varrho(n-1)} \cdot \underbrace{a_{n,n}}_{=1} \\ &\underset{\text{S. 5.3.1}}{=} \det \begin{pmatrix} a_{11} & \cdots & a_{1,n-1} \\ \vdots & & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} \end{pmatrix},\end{aligned}$$

da jedes $\sigma \in S_n$ mit $\sigma(n) = n$ als Permutation von $\{1, \ldots, n-1\}$ aufgefasst werden kann (ohne Änderung der Fehlstandszahl). $\square$

Definition 5.3.3. Es sei $n \geq 2$ und $\mathcal{A} \in K^{(n,n)}$. Mit $\mathcal{A}_{ij}$ bezeichnen wir die Matrix, die aus $\mathcal{A}$ durch Streichung der i-ten Zeile und der j-ten Spalte entsteht:

$$\mathcal{A}_{ij} = \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}.$$

Satz 5.3.5. *Es sei $n \geq 2$ und $\mathcal{A} \in K^{(n,n)}$. Dann gilt*

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j} & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix} = (-1)^{i+j} \cdot \det(\mathcal{A}_{ij}).$$

Beweis. Durch $n-j$ Vertauschungen benachbarter Spalten erhalten wir

$$\det(\mathcal{A}) = (-1)^{j-1} \cdot \det \begin{pmatrix} a_{11} & \cdots & a_{1n} & a_{1j} \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} & a_{nj} \end{pmatrix}.$$

Weitere $n-i$ Vertauschungen benachbarter Zeilen ergeben

$$\det(\mathcal{A}) = (-1)^{n-j}(-1)^{n-i} \cdot \left(\begin{array}{ccc|c} & & & * \\ & \mathcal{A}_{ij} & & \vdots \\ & & & * \\ \hline 0 & \cdots & 0 & 1 \end{array}\right).$$

Aus $(-1)^{n-i+n-j} = (-1)^{i+j}$ und Satz 5.3.4 folgt die Behauptung. $\square$

Satz 5.3.6. *(Entwicklungssatz von Laplace)*
Es sei $n \geq 2$ und $\mathcal{A} \in K^{(n,n)}$.
Für jedes feste i gilt die Entwicklung nach der i-ten Zeile:

$$det(\mathcal{A}) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(\mathcal{A}_{ij}).$$

Für jedes feste j gilt die Entwicklung nach der j-ten Spalte:

$$\det(\mathcal{A}) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(\mathcal{A}_{ij}).$$

Beweis. Es ist $\det(\mathcal{A}) = D_n(\vec{a}_1, \ldots, \vec{a}_n)$ mit den Zeilen $\vec{a}_k = (a_{k1} \ \cdots \ a_{kn})$ für $1 \leq k \leq n$. Die Entwicklung nach der i-ten Zeile folgt aus

$$\begin{aligned} D(\vec{a}_1, \ldots, \vec{a}_n) &= D_n\left(\vec{a}_1, \ldots , \sum_{j=1}^{n} a_{ij}\vec{e}_j \ , \ldots, \vec{a}_n\right) = \sum_{j=1}^{n} a_{ij} D_n(\vec{a}_1, \ldots, \vec{e}_j, \ldots, \vec{a}_n) \\ &= \sum_{j=1}^{n} a_{ij} \det \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix} \underset{\text{S. 5.3.5}}{=} \sum_{j=1}^{n} a_{ij} (-1)^{i+j} \det(\mathcal{A}_{ij}). \end{aligned}$$

Die Entwicklung nach der j-ten Spalte wird durch Übergang zur Transponierten gezeigt. □

Beispiel 5.3.3. Wir bestimmen die Determinante von

$$\mathcal{A} = \begin{pmatrix} 1 & 2 & 1 & 3 \\ 1 & 0 & 2 & 0 \\ 3 & 6 & 4 & 9 \\ 5 & 8 & 7 & 13 \end{pmatrix}.$$

Wegen des Auftretens von zwei Nullen in der zweiten Zeile empfiehlt es sich, die Determinante nach der zweiten Zeile zu entwickeln:

$$\det(\mathcal{A}) = (-1) \cdot 1 \cdot \det \begin{pmatrix} 2 & 1 & 3 \\ 6 & 4 & 9 \\ 8 & 7 & 13 \end{pmatrix} + (-1) \cdot 2 \cdot \det \begin{pmatrix} 1 & 2 & 3 \\ 3 & 6 & 9 \\ 5 & 8 & 13 \end{pmatrix}.$$

Die zweite Determinante ist null, da die ersten beiden Zeilen linear abhängig sind. Also ist

$$\begin{aligned} \det(\mathcal{A}) &= -\det \begin{pmatrix} 2 & 1 & 3 \\ 6 & 4 & 9 \\ 8 & 7 & 13 \end{pmatrix} \underset{\substack{(II)-3\cdot(I) \\ (III)-4\cdot(I)}}{=} -\det \begin{pmatrix} 2 & 1 & 3 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix} \overset{\text{2. Zeile}}{\underset{\text{entwickeln}}{=}} -(-1)^2 \det \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix} \\ &= -(2 \cdot 1 - 0 \cdot 3) = -2. \end{aligned}$$

Determinanten können auch dazu verwendet werden, explizite Formeln für die Lösungen eindeutig lösbarerer quadratischer Linearer Gleichungssysteme zu erhalten:

Satz 5.3.7. *(Cramersche Regel)*
Es sei $\mathcal{A} \in K^{(n,n)}$ *und* $\vec{b} = (b_1, \dots, b_n)^T \in K^n$ *(ein Spaltenvektor). Das LGS* $\mathcal{A}\vec{x} = \vec{b}$ *besitzt genau dann eine eindeutige Lösung, wenn* $\det(\mathcal{A}) \neq 0$ *ist. Ist dies der Fall, dann ist die Lösung* $\vec{x} = (x_1, \dots, x_n)^T$ *durch*

$$x_k = \frac{1}{\det(\mathcal{A})} \cdot \det \begin{pmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{pmatrix},$$

wobei die k*-te Spalte von* $\mathcal{A}$ *durch* $\vec{b}$ *ersetzt wird.*

Beweis. Nach Satz 4.1.3 ist $\mathcal{A}\vec{x} = \vec{b}$ genau dann eindeutig lösbar, wenn $\mathcal{A}$ regulär ist. Das ist nach Satz 5.3.3 (vi) zu $\det(\mathcal{A}) \neq 0$ äquivalent. In diesem Fall gilt für die eindeutig bestimmte Lösung $\vec{x} = (x_1, \dots, x_n)^T$

$$\vec{b} = x_1\vec{a}_1 + \dots + x_n\vec{a}_n,$$

wobei

$$\vec{a}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}$$

die j-te Spalte von $\mathcal{A}$ ist. Dann gilt

$$\begin{aligned} \det \begin{pmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{pmatrix} &= D_n(\vec{a}_1, \dots, \vec{a}_{k-1}, \vec{b}, \vec{a}_{k+1}, \dots, \vec{a}_n) \\ &= D_n\left(\vec{a}_1, \dots, \vec{a}_{k_1}, \sum_{j=1}^{n} x_j\vec{a}_j, \vec{a}_{k+1}, \dots, \vec{a}_n\right) \\ &= \sum_{j=1}^{n} x_j D_n(\vec{a}_1, \dots, \vec{a}_{k_1}, \vec{a}_j, \vec{a}_{k+1}, \dots, \vec{a}_n). \end{aligned}$$

Nun ist $D_n(\vec{a}_1, \dots, \vec{a}_{k-1}, \vec{a}_j, \vec{a}_{k+1}, \dots, \vec{a}_n) = 0$ für $j \neq k$, da dann zwei gleiche Argumente vorliegen. Der verbleibende Summand ist $x_k \det(\mathcal{A})$, also ist

$$\det \begin{pmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{pmatrix} = x_k \cdot \det(\mathcal{A}),$$

woraus die Behauptung folgt, weil man diese Gleichung wegen $\det(\mathcal{A}) \neq 0$ durch $\det(\mathcal{A})$ teilen darf. □

5.4 Der Multiplikationssatz

Satz 5.4.1. *(Determinantenmultiplikationssatz)*
Für $\mathcal{A}, \mathcal{B} \in K^{(n,n)}$ *gilt* $\det(\mathcal{A} \cdot \mathcal{B}) = \det(\mathcal{A}) \cdot \det(\mathcal{B})$*. Ist* $\mathcal{A}$ *regulär, so gilt* $\det(\mathcal{A}^{-1}) = \det(\mathcal{A})^{-1}$*.*

Beweis. Übungsaufgabe. □

Eine wichtige Folgerung aus diesem Satz ist, dass ein Basiswechsel die Determinante einer Darstellungsmatrix nicht ändert: Ist $\varphi \in L(V,V)$ für einen endlichdimensionalen Vektorraum V, und sind $\mathcal{B}_1$ und $\mathcal{B}_2$ zwei Basen von V, so ist $\mathcal{M}(\varphi;\mathcal{B}_2,\mathcal{B}_2) = \mathcal{X}^{-1}\cdot\mathcal{M}(\varphi;\mathcal{B}_1,\mathcal{B}_1)\cdot\mathcal{X}$ für eine reguläre Matrix $\mathcal{X} \in \mathrm{GL}(n,K)$ nach Satz 3.7.5. Nach dem Multiplikationssatz ist

$$\begin{aligned}\det(\mathcal{M}(\varphi;\mathcal{B}_2,\mathcal{B}_2)) &= \det(\mathcal{X}^{-1}\cdot\mathcal{M}(\varphi;\mathcal{B}_1,\mathcal{B}_1)\cdot\mathcal{X})\\ &= \det(\mathcal{X})^{-1}\cdot\det(\mathcal{M}(\varphi;\mathcal{B}_1,\mathcal{B}_1))\cdot\det(\mathcal{X}) = \det(\mathcal{M}(\varphi;\mathcal{B}_1,\mathcal{B}_1)).\end{aligned}$$

Definition 5.4.1. Es sei V ein n-dimensionaler K-Vektorraum. Die Determinante von φ ist

$$\det(\varphi) = \det(\mathcal{M}(\varphi;\mathcal{B},\mathcal{B}))$$

für eine Basis $\mathcal{B}$ von V. Die Definition hängt nach der vorigen Rechnung nicht davon ab, welche Basis man wählt.

Der Multiplikationssatz für Endomorphismen lautet

Satz 5.4.2. *Sind $\varphi,\psi \in L(V,V)$, so gilt $\det(\varphi\circ\psi) = \det(\varphi)\cdot\det(\psi)$. Es gilt*

$$\varphi \textit{ bijektiv} \;\Leftrightarrow\; \det(\varphi) \neq 0.$$

In diesem Fall ist $\det(\varphi^{-1}) = \det(\varphi)^{-1}$.

5.5 Inverse Matrix

Aus der Cramerschen Regel ergibt sich nun ein Ausdruck für die Inverse einer regulären Matrix.

Definition 5.5.1. Es sei $n \geq 2$ und $\mathcal{A} \in K^{(n,n)}$.
Die Matrix $\mathcal{A}_{ij}$ von Definition 5.3.3, die aus $\mathcal{A}$ durch Streichung der i- ten Zeile und der j- ten Spalte entsteht, bezeichnen wir auch als Minor der Matrix $\mathcal{A}$. Weiter heißt $\tilde{a}_{ij} = (-1)^{i+j}\det\mathcal{A}_{ij}$ Kofaktor, und die Transponierte der Matrix der Kofaktoren von $\mathcal{A}$ heißt Adjunkte von $\mathcal{A}$:

$$\operatorname{adj}(\mathcal{A}) = \begin{pmatrix}\tilde{a}_{11} & \tilde{a}_{21} & \dots & \tilde{a}_{n1}\\ \tilde{a}_{12} & \tilde{a}_{22} & \dots & \tilde{a}_{n2}\\ \vdots & \vdots & \ddots & \vdots\\ \tilde{a}_{1n} & \tilde{a}_{2n} & \dots & \tilde{a}_{nn}\end{pmatrix}.$$

Satz 5.5.1. *Es sei $n \in \mathbb{N}$ und $\mathcal{A} \in GL(n,K)$. Dann gilt*

$$\mathcal{A}^{-1} = \frac{1}{\det\mathcal{A}}\cdot\operatorname{adj}(\mathcal{A}).$$

Beweis. Es sei

$$\mathcal{A}^{-1} = \begin{pmatrix}x_{11} & \dots & x_{n1}\\ \vdots & & \vdots\\ x_{1n} & \dots & x_{nn}\end{pmatrix} = (\vec{x}_1,\dots,\vec{x}_n)$$

mit den Spaltenvektoren

$$\vec{x}_j = \begin{pmatrix}x_{1j}\\ \vdots\\ x_{ij}\\ \vdots\\ x-nj\end{pmatrix}.$$

Aus $\mathcal{A}\mathcal{A}^{-1} = E_n$ folgt, dass die n LGS $\mathcal{A}\vec{x}_j = \vec{e}_j$ für $1 \leq j \leq n$ erfüllt sein müssen.
Nach der Cramerschen Regel (Satz 5.3.7) folgt

$$x_{ij} = \frac{1}{\det(\mathcal{A})} \cdot \det \begin{pmatrix} a_{11} & \cdots & 0 & \cdots & a_{n1} \\ \vdots & & \vdots & & \vdots \\ \vdots & & 1 & & \vdots \\ \vdots & & \vdots & & \vdots \\ a_{1n} & \cdots & 0 & \cdots & a_{nn} \end{pmatrix},$$

wobei die i- te Spalte von $\mathcal{A}$ durch $\vec{e}_j$ ersetzt wird. Entwicklung nach der i- ten Spalte (nach Laplace, Satz 5.3.6) ergibt

$$\det \begin{pmatrix} a_{11} & \cdots & 0 & \cdots & a_{n1} \\ \vdots & & \vdots & & \vdots \\ \vdots & & 1 & & \vdots \\ \vdots & & \vdots & & \vdots \\ a_{1n} & \cdots & 0 & \cdots & a_{nn} \end{pmatrix} = (-1)^{i+j} \det \mathcal{A}_{ij} = \tilde{a}_{ji}.$$

$\square$

Kapitel 6

Euklidische und unitäre Vektorräume

6.1 Euklidische und unitäre Vektorräume

Definition 6.1.1. Es sei V ein Vektorraum über $\mathbb{R}$. Eine Abbildung $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{R}$ heißt inneres Produkt (oder Skalarprodukt) auf V, wenn

(I1) $\langle \vec{x} + \vec{y}, \vec{z} \rangle = \langle \vec{x}, \vec{z} \rangle + \langle \vec{y}, \vec{z} \rangle$ für alle $\vec{x}, \vec{y}, \vec{z} \in V$ (Additivität)

(I2) $\langle \lambda \vec{x}, \vec{y} \rangle = \lambda \langle \vec{x}, \vec{y} \rangle$ für alle $\vec{x}, \vec{y} \in V$ und $\lambda \in \mathbb{R}$ (Homogenität)

(I3) $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$ für alle $\vec{x}, \vec{y} \in V$ (Symmetrie)

(I4) $\langle \vec{x}, \vec{x} \rangle \geq 0$ für alle $\vec{x} \in V$ und $\langle \vec{x}, \vec{x} \rangle = 0 \Leftrightarrow \vec{x} = \vec{0}$ (positive Definitheit)

gelten.

Definition 6.1.2. Ein Euklidischer Vektorraum ist ein Paar $(V, \langle \cdot, \cdot \rangle)$, wobei V ein Vektorraum über $\mathbb{R}$ und $\langle \cdot, \cdot \rangle$ ein inneres Produkt auf V im Sinne von Definition 6.1.1 ist.

Definition 6.1.3. Es sei V ein Vektorraum über einem Körper K.
Eine Abbildung $b(\cdot, \cdot) \colon V \times V \to K$ heißt Bilinearform auf V, wenn

(B1) $b(\lambda \vec{x} + \mu \vec{y}, \vec{z}) = \lambda b(\vec{x}, \vec{z}) + \mu b(\vec{y}, \vec{z})$ für alle $\vec{x}, \vec{y}, \vec{z} \in V$ und $\lambda, \mu \in K$
(Linearität in der ersten Variablen)

(B2) $b(\vec{x}, \lambda \vec{y} + \mu \vec{z}) = \lambda b(\vec{x}, \vec{y}) + \mu b(\vec{x}, \vec{z})$ für alle $\vec{x}, \vec{y}, \vec{z} \in V$ und $\lambda, \mu \in K$
(Linearität in der zweiten Variablen)

gelten.

Definition 6.1.4. Es sei V ein Vektorraum über einem Körper K und $b(\cdot, \cdot) \colon V \times V \to K$ eine Bilinearform auf V.

i) Die Abbildung b heißt symmetrisch, wenn $b(\vec{x}, \vec{y}) = b(\vec{y}, \vec{x})$ für alle $\vec{x}, \vec{y} \in V$ gilt.

ii) Im Fall $K = \mathbb{R}$ heißt b positiv definit, wenn $b(\vec{x}, \vec{x}) \geq 0$ für alle $\vec{x} \in V$ und genau dann $b(\vec{x}, \vec{x}) = 0$ gilt, wenn $\vec{x} = \vec{0}$ gilt.

Bemerkung 6.1.1. Ein Euklidischer Vektorraum ist also ein reeller Vektorraum mit einer symmetrsichen, positiv definiten Bilinearform.

Definition 6.1.5. Es sei V ein Vektorraum über $\mathbb{C}$. Eine Abbildung $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{C}$ heißt inneres Produkt (oder Skalarprodukt) auf V, wenn

(I1') $\langle \vec{x} + \vec{y}, \vec{z} \rangle = \langle \vec{x}, \vec{z} \rangle + \langle \vec{y}, \vec{z} \rangle$ für alle $\vec{x}, \vec{y}, \vec{z} \in V$ (Additivität)

(I2') $\langle \lambda \vec{x}, \vec{y} \rangle = \lambda \langle \vec{x}, \vec{y} \rangle$ für alle $\vec{x}, \vec{y} \in V$ und $\lambda \in \mathbb{C}$ (Homogenität in der ersten Variablen)

(I3') $\langle \vec{x}, \vec{y} \rangle = \overline{\langle \vec{y}, \vec{x} \rangle}$ für alle $\vec{x}, \vec{y} \in V$ (Hermitizität)

(I4') $\langle \vec{x}, \vec{x} \rangle \geq 0$ für alle $\vec{x} \in V$ und $\langle \vec{x}, \vec{x} \rangle = 0 \Leftrightarrow \vec{x} = \vec{0}$ (positive Definitheit)

gelten.

Definition 6.1.6. Ein unitärer Vektorraum ist ein Paar $(V, \langle \cdot, \cdot \rangle)$, wobei V ein Vektorraum über $\mathbb{C}$ und $\langle \cdot, \cdot \rangle$ ein inneres Produkt auf V im Sinne von Definition 6.1.5 ist.

Bemerkung 6.1.2. Diese Abbildung $\langle \cdot, \cdot \rangle$ ist zwar in der ersten Variablen, nicht aber in der zweiten Variablen linear, da aus $(I2')$ und $(I3')$ dann $\langle \vec{x}, \lambda \vec{y} \rangle = \overline{\lambda} \langle \vec{x}, \vec{y} \rangle$ folgt.
Damit ist $\langle \cdot, \cdot \rangle$ keine Bilinearform, sondern eine Sesquilinearform ($1\frac{1}{2}$- fache Linearform).

Beispiel 6.1.1. Es sei $n \in \mathbb{N}$.
Der Vektorraum $V = \mathbb{R}^n$ wird durch $\langle \vec{x}, \vec{y} \rangle := \sum_{i=1}^{n} x_i y_i$ für $\vec{x}, \vec{y} \in \mathbb{R}^n$ mit $\vec{x} = (x_1, \ldots, x_n)^T$ und $\vec{y} = (y_1, \ldots, y_n)^T$ zu einem Euklidischen Vektorraum.

Beispiel 6.1.2. Es sei $n \in \mathbb{N}$.
Der Vektorraum $V = \mathbb{C}^n$ wird durch $\langle \vec{x}, \vec{y} \rangle := \sum_{i=1}^{n} x_i \overline{y_i}$ für $\vec{x}, \vec{y} \in \mathbb{C}^n$ mit $\vec{x} = (x_1, \ldots, x_n)^T$ und $\vec{y} = (y_1, \ldots, y_n)^T$ zu einem unitären Vektorraum.

Definition 6.1.7. Die inneren Produkte $\langle \cdot, \cdot \rangle$ auf dem $\mathbb{R}^n$ in Beispiel 6.1.1 bzw. auf dem $\mathbb{C}^n$ in Beispiel 6.1.2 heißen kanonisches inneres Produkt auf dem $\mathbb{R}^n$ bzw. auf dem $\mathbb{C}^n$.

Es gibt nun eine umkehrbar eindeutige Beziehung zwischen Bilinearformen auf dem $\mathbb{R}^n$ bzw. Sesquilinearformen auf dem $\mathbb{C}^n$ einerseits und quadratischen Matrizen aus dem $\mathbb{R}^{(n,n)}$ bzw. aus dem $\mathbb{C}^{(n,n)}$ andererseits.

Definition 6.1.8. Es sei K ein Körper und $n \in \mathbb{N}$. Die Matrix $\mathcal{A} \in K^{(n,n)}$ heißt symmetrisch, wenn $\mathcal{A} = \mathcal{A}^T$ gilt.

Bemerkung 6.1.3. Ist $\mathcal{A} = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \in K^{(n,n)}$, so ist $\mathcal{A}$ offenbar genau dann symmetrisch, wenn $a_{ij} = a_{ji}$ für alle $1 \leq i, j \leq n$ gilt.

Definition 6.1.9. Es sei $n \in \mathbb{N}$. Eine Matrix $\mathcal{A} \in \mathbb{R}^{(n,n)}$ heißt

i) positiv definit, wenn $\vec{x}^T \mathcal{A} \vec{x} > 0$ für alle $\vec{x} \in \mathbb{R}^n \backslash \{\vec{0}\}$

ii) negativ definit, wenn $\vec{x}^T \mathcal{A} \vec{x} < 0$ für alle $\vec{x} \in \mathbb{R}^n \backslash \{\vec{0}\}$

iii) positiv semidefinit, wenn $\vec{x}^T \mathcal{A} \vec{x} \geq 0$ für alle $\vec{x} \in \mathbb{R}^n$

iv) negativ semidefinit, wenn $\vec{x}^T \mathcal{A} \vec{x} \leq 0$ für alle $\vec{x} \in \mathbb{R}^n$

gilt.

Satz 6.1.1. *Es sei $n \in \mathbb{N}$, K ein Körper und V ein Vektorraum über K mit* $\dim V = n$.

i) Es sei $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ eine Basis von V und $b(\cdot,\cdot)\colon V \times V \to K$ eine Bilinearform auf V. Dann existiert genau eine Matrix $\mathcal{A} \in K^{(n,n)}$, so dass

$$b(\vec{x}, \vec{y}) = (\vec{x}')^T \mathcal{A} \vec{y}' \tag{$*$}$$

mit den Koordinatenvektoren $\vec{x}'$ von $\vec{x}$ bzw. $\vec{y}'$ von $\vec{y}$ bzgl. $\mathcal{B}$, d.h. $\vec{x}' = (x_1', \ldots, x_n')^T$, wobei $\vec{x} = \sum_{j=1}^n x_j' \vec{b}_j$, bzw. $\vec{y}' = (y_1', \ldots, y_n')^T$, wobei $\vec{y} = \sum_{j=1}^n y_j' \vec{b}_j$ gilt.
Es ist dann $\mathcal{A} = (b(\vec{b}_i, \vec{b}_j))_{\substack{1 \le i \le n \\ 1 \le j \le n}}$.

ii) Für jede Matrix $\mathcal{A} \in K^{(n,n)}$ ist durch $(*)$ eine Bilinearform auf V bestimmt.

iii) Die Abbildung b ist genau dann symmetrisch, wenn $\mathcal{A}$ symmetrisch ist.

iv) Ist $K = \mathbb{R}$, so ist b genau dann positiv definit, wenn $\mathcal{A}$ positiv definit ist.

v) Ist $K = \mathbb{R}$, so ist b ein inneres Produkt auf V und (V, b) damit genau dann ein Euklidischer Vektorraum, wenn $\mathcal{A}$ symmetrisch und positiv definit ist.

Beweis. Durch Nachrechnen. □

Definition 6.1.10. i) Für die Matrix in Satz 6.1.1 (i) schreiben wir $\mathcal{A} = \mathcal{M}_{\mathcal{B}}(b)$. Sie heißt die Matrix der Bilinearform b (bzgl. der Basis $\mathcal{B}$).

ii) Für die Bilinearform b in Satz 6.1.1 (ii) schreiben wir $b = b_{\mathcal{B}}(\mathcal{A})$. Sie heißt die zur Matrix $\mathcal{A}$ (bzgl. der Basis $\mathcal{B}$) gehörende Bilinearform.

Satz 6.1.2. *(Bilinearformen bei Basiswechsel)*
Es sei $n \in \mathbb{N}$, K ein Körper und V ein Vektorraum über K mit $\dim V = n$.
Weiter seien $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ und $\tilde{\mathcal{B}} = \{\sigma(\vec{b}_1), \ldots, \sigma(\vec{b}_n)\}$ Basen von V mit $\sigma \in GL(V)$ und b eine Bilinearform auf V. Dann gilt

$$\mathcal{M}_{\tilde{\mathcal{B}}}(b) = \mathcal{X}^T \cdot \mathcal{M}_{\mathcal{B}}(b) \cdot \mathcal{X}.$$

Beweis. Es seien

$$\begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix} \quad \text{bzw.} \quad \begin{pmatrix} x_1'' \\ \vdots \\ x_n'' \end{pmatrix}$$

die Koordinatenvektoren von $\vec{x}$ bzgl. $\mathcal{B}$ bzw. $\tilde{\mathcal{B}}$ und

$$\begin{pmatrix} y_1' \\ \vdots \\ y_n' \end{pmatrix} \quad \text{bzw.} \quad \begin{pmatrix} y_1'' \\ \vdots \\ y_n'' \end{pmatrix}$$

die Koordinatenvektoren von $\vec{y}$ bzgl. $\mathcal{B}$ bzw. $\tilde{\mathcal{B}}$. Nach Satz 3.7.3 ist

$$\begin{pmatrix} x_1'' \\ \vdots \\ x_n'' \end{pmatrix} = \mathcal{X}^{-1} \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} y_1'' \\ \vdots \\ y_n'' \end{pmatrix} = \mathcal{X}^{-1} \begin{pmatrix} y_1' \\ \vdots \\ y_n' \end{pmatrix}$$

und damit

$$(x_1', \ldots, x_n')\mathcal{M}_{\mathcal{B}}(b) \begin{pmatrix} y_1' \\ \vdots \\ y_n' \end{pmatrix} = (x_1'', \ldots, x_n'')\mathcal{X}^T \mathcal{M}_{\mathcal{B}}(b)\mathcal{X} \begin{pmatrix} y_1'' \\ \vdots \\ y_n'' \end{pmatrix}.$$

Vergleich mit $(*)$ in 6.1.1 ergibt die Behauptung. □

Eine eng verbundene Frage ist die folgende: Welche Automorphismen $\varphi \in GL(V)$ lassen eine gegebene Bilinearform invariant, d.h. für welche φ ist $b(\varphi(\vec{x}), \varphi(\vec{y})) = b(\vec{x}, \vec{y})$ für alle $\vec{x}, \vec{y} \in V$?

Definition 6.1.11. Es sei V ein Vektorraum über $\mathbb{R}$ und b eine Bilinearform auf V. Ein Automorphismus $\varphi \in GL(V)$ heißt Automorphismus von b, falls $b(\varphi(\vec{x}), \varphi(\vec{y})) = b(\vec{x}, \vec{y})$ für alle $\vec{x}, \vec{y} \in V$ gilt. Die Menge aller Automorphismen von b bezeichnen wir mit $\mathrm{Aut}(b)$.

Satz 6.1.3. *Es bildet* $\mathrm{Aut}(b)$ *eine Gruppe bzgl. der Komposition von Abbildungen.*

Beweis. Wir wenden das Untergruppenkriterium (Satz 1.5.4) an.
Wegen $id \in \mathrm{Aut}(b)$ gilt $\mathrm{Aut}(b) \neq \emptyset$. Weiter folgt

$$\varphi \in \mathrm{Aut}(b) \Rightarrow b(\varphi^{-1}(\vec{x}), \varphi^{-1}(\vec{y})) = b(\varphi(\varphi^{-1}(\vec{x})), \varphi(\varphi^{-1}(\vec{y}))) = b(\vec{x}, \vec{y})$$

für alle $\vec{x}, \vec{y} \in V$. Also gilt $\varphi^{-1} \in \mathrm{Aut}(b)$.
Es seien nun $\varphi, \psi \in \mathrm{Aut}(b)$. Wegen $\varphi^{-1} \in \mathrm{Aut}(b)$ gilt

$$b(\vec{x}, \vec{y}) = b(\varphi^{-1}(\vec{x}), \varphi^{-1}(\vec{y})) = b(\psi(\varphi^{-1}(\vec{x})), \psi(\varphi^{-1}(\vec{y}))) \Rightarrow \psi \circ \varphi^{-1} \in \mathrm{Aut}(b)$$

□

Für endlichdimensionale Vektorräume lassen sich nun die Automorphismen einer Bilinearform mittels Satz 6.1.1 durch ihre Darstellungsmatrizen charakterisieren.

Satz 6.1.4. *Es sei $n \in \mathbb{N}$, K ein Körper und V ein Vektorraum über K mit* $\dim V = n$. *Weiter sei $\mathcal{B}$ eine Baiss von V, b eine Bilinearform auf V, $\mathcal{A} = \mathcal{M}_{\mathcal{B}}(b)$ die Matrix von b bzgl. $\mathcal{B}$ und $\varphi \in GL(V)$. Dann gilt genau dann $\varphi \in$* $\mathrm{Aut}(b)$, *wenn für die Darstelllungsmatrix $\mathcal{X} = \mathcal{M}(\varphi; \mathcal{B}, \mathcal{B})$*

$$\mathcal{X}^T \cdot \mathcal{A} \cdot \mathcal{X} = \mathcal{A}$$

gilt.

Beweis. Es seien $\vec{x}, \vec{y} \in V$mit den Koordinatenvektoren $\vec{x}', \vec{y}' \in K^n$, womit $\varphi(\vec{x})$ bzw. $\varphi(\vec{y})$ die Koordinatenvektoren $\mathcal{X}\vec{x}'$ bzw. $\mathcal{X}\vec{y}'$ haben. Nach Satz 6.1.1 (ii) gilt

$$b(\varphi(\vec{x}), \varphi(\vec{y})) = (\vec{x}')^T \mathcal{X}^T \mathcal{A} \mathcal{X} \vec{y}' \quad \text{und} \quad b(\vec{x}, \vec{y}) = (\vec{x}')^T \mathcal{A} \vec{y}'.$$

Da die Matrizen durch die Bilinearformen eindeutig bestimmt sind, folgt die Behauptung. □

Von besonderer Wichtigkeit ist der Spezialfall des kanonischen inneren Produktes auf dem $\mathbb{R}^n$:
Für $\vec{x}, \vec{y} \in \mathbb{R}^n$ mit

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{und} \quad \vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

ist $b(\vec{x}, \vec{y}) = \langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^n x_i y_i = \vec{x}^T E_n \vec{y}$ mit der Einheitsmatrix E_n.

Es sei $\mathcal{B} \in GL(n, \mathbb{R})$. Nach Satz 6.1.4 gilt für alle $\vec{x}, \vec{y} \in \mathbb{R}^n$ genau dann

$$\langle \mathcal{B}\vec{x}, \mathcal{B}\vec{y} \rangle = \langle \vec{x}, \vec{y} \rangle,$$

wenn

$$\mathcal{B}^T \mathcal{B} = E_n \tag{$**$}$$

gilt. Matrizen, die $(**)$ erfüllen, nennt man <u>orthogonal</u>. Der Grund für diese Namensgebung wird in Kürze klar werden.

Definition 6.1.12. Es sei $n \in \mathbb{N}$. Eine Matrix $\mathcal{A} \in \mathbb{R}^{(n,n)}$ heißt <u>orthogonal</u>, wenn $\mathcal{A}^T \mathcal{A} = E_n$ gilt.

Bemerkung 6.1.4. Die Eigenschaft $\mathcal{A}^T \mathcal{A} = E_n$ bedeutet, dass $\mathcal{A}^{-1} = \mathcal{A}^T$ ist. Da das Linksinverse stets auch Rechtsinverses ist, folgt auch $\mathcal{A}\mathcal{A}^T = E_n$.

Satz 6.1.5. *Es sei $n \in \mathbb{N}$, und $\mathcal{A} \in \mathbb{R}^{(n,n)}$ sei orthogonal. Dann gilt* $\det \mathcal{A} \in \{-1, 1\}$.

Beweis. Wegen $\det \mathcal{A}^T = \det \mathcal{A}$ folgt nach dem Determinantenmultiplikationssatz (Satz 5.4.1) und nach Satz 5.3.2

$$(\det \mathcal{A})^2 = \det \mathcal{A}^T \cdot \det \mathcal{A} = \det E_n = 1.$$

□

Definition 6.1.13. Es sei $n \in \mathbb{N}$. Dann setzen wir

$$\begin{aligned} O(n) &:= \{\mathcal{A} \in \mathbb{R}^{(n,n)} \colon \mathcal{A}^T \mathcal{A} = E_n\} \\ SO(n) &:= \{\mathcal{A} \in O(n) \colon \det \mathcal{A} = 1\} \end{aligned}$$

Nach Satz 6.1.3 und Satz 6.1.4 bildet $O(n)$ eine Gruppe bzgl. der Komposition von Abbildungen. Nach dem Determinantenmultiplikationssatz (Satz 5.4.1) ist $SO(n)$ eine Untergruppe von $O(n)$.

Definition 6.1.14. Es heißt $O(n)$ <u>orthogonale Gruppe</u> und $SO(n)$ <u>spezielle orthogonale Gruppe</u>.

Wir kommen nun zur Theorie der Sesquilinearformen und der inneren Produkte auf komplexen Vektorräumen. Es gibt eine enge Analogie zur obigen Theorie der Bilinearformen und inneren Produkte auf reellen Vektorräumen. Wir geben sie dehalb nur in gekürzter Form wider.

Definition 6.1.15. Es sei V ein Vektorraum über $\mathbb{C}$. Eine Abbildung $s(\cdot,\cdot) \colon V \times V \to \mathbb{C}$ heißt <u>Sesquilinearform</u> auf V, wenn

(S1) $s(\lambda\vec{x} + \mu\vec{y}, \vec{z}) = \lambda s(\vec{x}, \vec{z}) + \mu s(\vec{y}, \vec{z})$ für alle $\vec{x}, \vec{y}, \vec{z} \in V$ und $\lambda, \mu \in \mathbb{C}$
(Linearität in der ersten Variablen)

(S2) $s(\vec{x}, \lambda\vec{y} + \mu\vec{z}) = \overline{\lambda} s(\vec{x}, \vec{y}) + \overline{\mu} s(\vec{x}, \vec{z})$ für alle $\vec{x}, \vec{y}, \vec{z} \in V$ und $\lambda, \mu \in \mathbb{C}$
(Semilinearität in der zweiten Variablen)

gelten.

Definition 6.1.16. Es sei $n \in \mathbb{N}$ und $\mathcal{A} \in \mathbb{C}^{(n,n)}$ mit $\mathcal{A} = (a_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le n}}$. Dann heißt $\mathcal{A}^* = (\overline{a_{ji}})_{\substack{1 \le i \le n \\ 1 \le j \le n}}$ die <u>Adjungierte</u> von $\mathcal{A}$.

Bemerkung 6.1.5. Die Adjungierte ist für $\mathcal{A} \in \mathbb{R}(n,n)$ die Transponierte von $\mathcal{A}$ und für $\mathcal{A} \in \mathbb{C}^{(n,n)}$ die Konjugierte der Transponierten.

Definition 6.1.17. Es sei $n \in \mathbb{N}$. Eine Matrix $\mathcal{A} \in \mathbb{C}^{(n,n)}$ heißt (nach Charles Hermite, 1822-1901) hermitesch oder selbstadjungiert, wenn $\mathcal{A}^* = \mathcal{A}$ ist.

Definition 6.1.18. Es sei $n \in \mathbb{N}$. Eine Matrix $\mathcal{A} \in \mathbb{C}^{(n,n)}$ heißt

i) positiv definit, wenn $\vec{x}^*\mathcal{A}\vec{x} > 0$ für alle $\vec{x} \in \mathbb{C}^n \backslash \{\vec{0}\}$

ii) negativ definit, wenn $\vec{x}^*\mathcal{A}\vec{x} < 0$ für alle $\vec{x} \in \mathbb{C}^n \backslash \{\vec{0}\}$

iii) positiv semidefinit, wenn $\vec{x}^*\mathcal{A}\vec{x} \geq 0$ für alle $\vec{x} \in \mathbb{C}^n$

iv) negativ semidefinit, wenn $\vec{x}^*\mathcal{A}\vec{x} \leq 0$ für alle $\vec{x} \in \mathbb{C}^n$

gilt.

Die Sätze 6.1.1 bis 6.1.5 und die mit ihnen verbundenen Definitionen 6.1.8 bis 6.1.12 haben nun alle ihre Entsprechung für Sesquilinearformen und inneren Produkte auf komplexen Vektorräumen:
An die Stelle der Darstellung $b(\vec{x}, \vec{y}) = (\vec{x}')^T \mathcal{A}\vec{y}'$ der Bilinearform b von Satz 6.1.1 tritt nun die Darstellung $s(\vec{x}, \vec{y}) = (\vec{x}')^T \mathcal{A}\vec{y}'$ der Sesquilinearform s.
Da die entsprechenden Sätze, ihre Beweise und die Definitionen alle durch offensichtliche Änderungen aus den Sätzen 6.1.1 bis 6.1.5 und den Definitionen 6.1.8 bis 6.1.12 hervorgehen, verzichten wir auf die Einzelheiten.

Wir geben lediglich folgende Definitionen und Sätze ohne Beweise an:

Definition 6.1.19. Es sei $n \in \mathbb{N}$. Eine Matrix $\mathcal{A} \in \mathbb{C}^{(n,n)}$ heißt unitär, wenn $\mathcal{A}^*\mathcal{A} = E_n$ gilt.

Satz 6.1.6. *Es sei $n \in \mathbb{N}$, und $\mathcal{A} \in \mathbb{C}^{(n,n)}$ sei unitär. Dann gilt $|\det \mathcal{A}| = 1$.*

Bemerkung 6.1.6. Die Menge der möglichen Werte der Determinante einer unitären Matrix $\mathcal{A}$ ist damit unendlich. Die Werte von $\det \mathcal{A}$ liegen auf dem Einheitskreis der komplexen Zahlenebene.

Definition 6.1.20. Es sei $n \in \mathbb{N}$. Die Menge $U(n)$ der unitären Matrizen bildet eine Gruppe bzgl. der Matrixmultiplikation. Die Menge $SU(n)$ der Matrizen aus $U(n)$ mit Determinante 1 bildet eine Untergruppe von $U(n)$.
Für $\mathcal{A} \in U(n)$ gilt

$$\mathcal{A}^*\mathcal{A} = \mathcal{A}\mathcal{A}^* = E_n \quad \text{und} \quad \mathcal{A}^{-1} = \mathcal{A}^*.$$

Es gilt genau dann $\mathcal{A} \in U(n)$, wenn

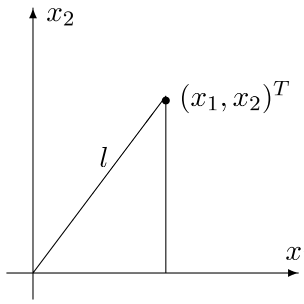
$$\langle \vec{x}, \vec{y} \rangle = \langle \mathcal{A}\vec{x}, \mathcal{A}\vec{y} \rangle$$

für alle $\vec{x}, \vec{y} \in \mathbb{C}^n$ gilt, wobei $\langle \cdot, \cdot \rangle$ das kanonische innere Produkt auf $\mathbb{C}^n$ bedeutet.

6.2 Längen und Winkel, Orthogonalität

Mittels des inneren Produktes auf Euklidischen Vektorräumen können nun Längen von und Winkel zwischen Vektoren definiert werden. Anstelle des Begriffs der Länge wird meist der Begriff Norm verwendet. Die Definitionen dieser Begriffe werden auch für unitäre Vektorräume verwendet. Diese Begriffe haben ihren Ursprung in der Elementargeometrie.

Beispiel 6.2.1. Es sei $\vec{x} = (x_1, x_2)^T \in \mathbb{R}^2$.



Nach dem Satz des Pythagoras ist der Abstand l des Punktes $(x_1, x_2)^T$ vom Urspung gerade

$$l = \sqrt{x_1^2 + x_2^2} \quad \text{oder} \quad l = \sqrt{\langle \vec{x}, \vec{x} \rangle},$$

wobei $\langle \cdot, \cdot \rangle$ das kanonische innere Produkt auf $\mathbb{R}^2$ bedeutet.

Dies liefert die Motivation für folgende

Definition 6.2.1. Es sei $(V, \langle \cdot, \cdot \rangle)$ ein Euklidischer oder unitärer Vektorraum und $\vec{x} \in V$. Unter der Norm $||\vec{x}||$ des Vektors $\vec{x}$ versteht man $||\vec{x}|| = \sqrt{\langle \vec{x}, \vec{x} \rangle}$.

Satz 6.2.1. *(Cauchy- Schwarzsche- Ungleichung (CSU))*
Es sei $(V, \langle \cdot, \cdot \rangle)$ ein Euklidischer oder unitärer Vektorraum und $\vec{x}, \vec{y} \in V$. Dann gilt

$$|\langle \vec{x}, \vec{y} \rangle| \leq ||\vec{x}|| \cdot ||\vec{y}||,$$

und das Gleichheitszeichen gilt genau dann, wenn $\vec{x}$ und $\vec{y}$ linear abhängig sind.

Beweis. Da die Behauptung für $\vec{y} = \vec{0}$ offenbar richtig ist, können wir $\vec{y} \neq \vec{0}$ annehmen. Wir setzen

$$t = -\frac{\langle \vec{x}, \vec{y} \rangle}{\langle \vec{y}, \vec{y} \rangle}.$$

Wegen der positiven Definitheit (Definition 6.1.1 (I4)) folgt

$$0 \leq \langle \vec{x} + t\vec{y}, \vec{x} + t\vec{y} \rangle = ||\vec{x}||^2 - \frac{|\langle \vec{x}, \vec{y} \rangle|^2}{||\vec{y}||^2}$$

mit Gleichheit genau für $\vec{x} + t\vec{y} = \vec{0}$. Multiplikation mit $||\vec{y}||^2$ liefert die Behauptung. □

Satz 6.2.2. *Es sei V ein Euklidischer oder unitärer Vektorraum sowie $\vec{x}, \vec{y} \in V$ und $\lambda \in K$ mit $K = \mathbb{R}$ oder $K = \mathbb{C}$. Dann gilt*

i) $||\vec{x}|| \geq 0$ *und* $||\vec{x}|| = 0$ *genau dann, wenn* $\vec{x} = \vec{0}$

ii) $||\lambda \vec{x}|| = |\lambda| \cdot ||\vec{x}||$

iii) $||\vec{x} + \vec{y}|| \leq ||\vec{x}|| + ||\vec{y}||$ *(Dreiecksungleichung)*

Beweis. Die Aussagen (i) und (ii) folgen unmittelbar aus den Definitionen 6.1.1 und 6.2.1.
Zu (iii): Für Euklidische Vektorräume gilt

$$\begin{aligned} ||\vec{x}+\vec{y}||^2 &= \langle \vec{x}+\vec{y}, \vec{x}+\vec{y}\rangle = ||\vec{x}||^2 + 2\langle \vec{x}, \vec{y}\rangle + ||\vec{y}||^2 \underset{(CSU)}{\leq} ||\vec{x}||^2 + 2||\vec{x}|| \cdot ||\vec{y}|| + ||\vec{y}||^2 \\ &= (||\vec{x}|| + ||\vec{y}||)^2. \end{aligned}$$

Für unitäre Vektorräume folgt der Beweis etwas aufwändiger aber ähnlich. □

In der Elementargeometrie ist der Begriff der Kongruenz (Deckungsgleichheit) von besonderer Bedeutung. Zwei Teilmengen des $\mathbb{R}^2$ heißen kongruent, wenn sie durch eine Bewegung (Kongruenz) ineinander überführt werden können.

Definition 6.2.2. Es sei $n \in \mathbb{N}$.

i) Unter dem Abstand $d(\vec{x}, \vec{y})$ zwischen den Vektoren $\vec{x}, \vec{y} \in \mathbb{R}^n$ versteht man $d(\vec{x}, \vec{y}) := ||\vec{x} - \vec{y}||$.

ii) Eine Abbildung $f\colon \mathbb{R}^n \to \mathbb{R}^n$ mit $d(f(\vec{x}), f(\vec{y})) = d(\vec{x}, \vec{y})$ für alle $\vec{x}, \vec{y} \in \mathbb{R}^n$ heißt Bewegung.

iii) Eine lineare Abbildung $\varphi \in L(\mathbb{R}^n, \mathbb{R}^n)$ heißt orthogonal, wenn $||\varphi(\vec{x})|| = ||\vec{x}||$ für alle $\vec{x} \in \mathbb{R}^n$ gilt.

Bemerkung 6.2.1. Man kann zeigen, dass die Bewegungen f des $\mathbb{R}^n$ mit $f(\vec{0}) = \vec{0}$ gerade die orthogonalen Abbildungen sind.

Satz 6.2.3. *Es sei $n \in \mathbb{N}$ und $\varphi \in L(\mathbb{R}^n, \mathbb{R}^n)$. Folgende Aussagen sind äquivalent:*

i) Die Abbildung φ ist orthogonal.

ii) Die Abbildung φ erhält das innere Produkt, d.h. es gilt $\langle \vec{x}, \vec{y}\rangle = \langle \varphi(\vec{x}), \varphi(\vec{y})\rangle$ für alle $\vec{x}, \vec{y} \in \mathbb{R}^n$.

iii) Es existierte eine Matrix $\mathcal{A} \in O(n)$ mit $\varphi(\vec{x}) = \mathcal{A}\vec{x}$ für alle $\vec{x} \in \mathbb{R}^n$.

Beweis. i) $\to$ *ii*):
Es seien $\vec{x}, \vec{y} \in \mathbb{R}^n$. Es ist

$$\begin{aligned} ||\varphi(\vec{x}+\vec{y})||^2 &= \langle \varphi(\vec{x}+\vec{y}), \varphi(\vec{x}+\vec{y})\rangle = ||\varphi(\vec{x})||^2 + 2\langle \varphi(\vec{x}), \varphi(\vec{y})\rangle + ||\varphi(\vec{y})||^2 \\ &= ||\vec{x}||^2 + 2\langle \varphi(\vec{x}), \varphi(\vec{y})\rangle + ||\vec{y}||^2 \end{aligned}$$

und andererseits

$$||\varphi(\vec{x}+\vec{y})||^2 = ||\vec{x}+\vec{y}||^2 = ||\vec{x}||^2 + 2\langle \vec{x}, \vec{y}\rangle + ||\vec{y}||^2.$$

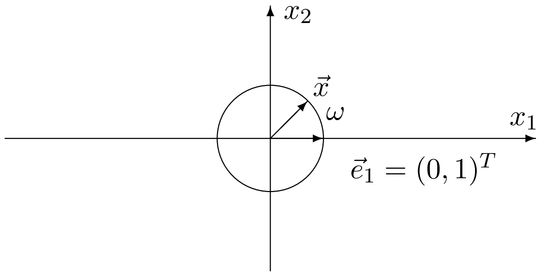
Zusammenfassen beider Aussagen liefert $\langle \varphi(\vec{x}), \varphi(\vec{y})\rangle = \langle \vec{x}, \vec{y}\rangle$.

ii) $\to$ *iii*):
Es sei $\mathcal{A}$ die Darstellungsmatrix von φ bzgl. der kanonischen Basis $\mathcal{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$. Nach Definition 6.1.11 ist $\varphi \in \mathrm{Aut}(\langle \cdot, \cdot \rangle)$ und nach Satz 6.1.4 gilt $\mathcal{A}^T E_n \mathcal{A} = E_n$, also $\mathcal{A}^T\mathcal{A} = E_n$, womit $\mathcal{A} \in O(n)$ ist.

iii) $\to$ *i*):
Es ist $||\varphi(\vec{x})||^2 = \langle \mathcal{A}\vec{x}, \mathcal{A}\vec{x}\rangle = \langle \vec{x}, \vec{x}\rangle = ||\vec{x}||^2$, womit die Abbildung φ orthogonal ist.

□

Zur Motivation der Definition des Winkels zwischen zwei Vektoren betrachten wir den Einheitskreis $\mathbb{E} = \{\vec{x} \in \mathbb{R}^2 \colon ||\vec{x}|| = 1\}$ im $\mathbb{R}^2$. Dieser kann durch die aus der Analysis bekannten Funktionen Cosinus und Sinus folgendermaßen parametrisiert werden: $\mathbb{E} = \{(\cos t, \sin t)^T \colon 0 \leq t < 2\pi\}$.
Es sei $\vec{x} = (\cos\omega, \sin\omega)^T \in \mathbb{E}$. Der Einfachheit halber beschränken wir uns auf den Fall $\sin\omega \geq 0$, d.h. $\omega \in [0, \pi]$.
Das Bogenmaß des Winkels $\measuredangle(\vec{e}_1, \vec{x})$ zwischen dem Einheitsvektor $\vec{e}_1$ und dem Vektor $\vec{x}$ ist durch die Länge der Kurve (des Bogens) $\gamma \colon [0, \omega] \to \mathbb{R}^2$, $t \to (\cos t, \sin t)$ gegeben.



Diese Länge l erhalten wir über

$$l = \int_0^\omega \left|\left|\frac{d}{dt}\gamma(t)\right|\right| dt = \int_0^\omega \sqrt{\left(\frac{d}{dt}\cos t\right)^2 + \left(\frac{d}{dt}\sin t\right)^2}\, dt = \int_0^\omega \sin^2 t + \cos^2 t \, dt = \int_0^\omega 1 \, dt = \omega.$$

Also sollte $\measuredangle(\vec{e}_1, \vec{x})$ durch

$$\cos(\measuredangle(\vec{e}_1, \vec{x})) = \cos\omega = x_1 = \langle \vec{e}_1, \vec{x} \rangle$$

mit $0 \leq \measuredangle(\vec{e}_1, \vec{x}) \leq \pi$ definiert werden.
Diese Definition sollte auch für jedes Paar $\vec{x}$ und $\vec{y}$ von Vektoren aus $\mathbb{E}$ gelten, die aus $\vec{e}_1$ und $\vec{x}$ durch eine orthogonale Abbildung hervorgehen. Da das innere Produkt erhalten bleibt, sollte auch hier der Winkel $\measuredangle(\vec{x}, \vec{y})$ durch

$$\cos(\measuredangle(\vec{x}, \vec{y})) = \langle \vec{x}, \vec{y} \rangle$$

mit $0 \leq \measuredangle(\vec{x}, \vec{y}) \leq \pi$ definiert werden.
Schließlich sollte der Winkel erhalten bleiben, wenn sowohl $\vec{x}$ als auch $\vec{y}$ mit beliebigen positiven Faktoren multipliziert werden:

$$\measuredangle(\lambda\vec{x}, \mu\vec{y}) = \measuredangle(\vec{x}, \vec{y})$$

für alle $\lambda, \mu > 0$. Dies wird durch die Festlegung

$$\langle \vec{x}, \vec{y} \rangle = ||\vec{x}|| \cdot ||\vec{y}|| \cdot \cos(\measuredangle(\vec{x}, \vec{y}))$$

mit $0 \leq \measuredangle(\vec{x}, \vec{y}) \leq \pi$ gewährleistet.

Diese gesamten der Elementargeometrie und der Analysis entstammenden Überlegungen lassen folgende Definition als sinnvoll erscheinen:

Definition 6.2.3. Es sei $(V, \langle\cdot,\cdot\rangle)$ ein Euklidischer oder unitärer Vektorraum und $\vec{x}, \vec{y} \in V - \{\vec{0}\}$. Dann ist der Winkel $\measuredangle(\vec{x}, \vec{y})$ zwischen $\vec{x}$ und $\vec{y}$ durch

$$\langle \vec{x}, \vec{y} \rangle = ||\vec{x}|| \cdot ||\vec{y}|| \cdot \cos(\measuredangle(\vec{x}, \vec{y}))$$

mit $0 \leq \measuredangle(\vec{x}, \vec{y}) \leq \pi$ definiert.
Die Vektoren $\vec{x}$ und $\vec{y}$ heißen orthogonal (senkrecht) zueinander (Schreibweise: $\vec{x} \perp \vec{y}$), wenn $\langle \vec{x}, \vec{y} \rangle = 0$ gilt.

Bemerkung 6.2.2. Aus der Definition des Winkels folgt, dass zwei Vektoren $\vec{x}$ und $\vec{y}$ genau dann senkrecht zueinander stehen, also $\vec{x} \perp \vec{y}$ gilt, wenn $\measuredangle(\vec{x}, \vec{y}) = \frac{\pi}{2}$ ist.

6.3 Orthonormalbasen, Gram- Schmidt- Verfahren

Definition 6.3.1. Es sei $(V, \langle\cdot,\cdot\rangle)$ ein Euklidischer oder unitärer Vektorraum.
Eine Teilmenge $M \subset V$ heißt Orthogonalsystem, wenn $\vec{0} \notin M$ und für $\vec{x}, \vec{y} \in M$ mit $\vec{x} \neq \vec{y}$ dann $\vec{x} \perp \vec{y}$ folgt. Gilt außerdem noch $||\vec{x}|| = 1$ für alle $\vec{x} \in M$, so heißt M Orthonormalsystem (ONS).

Satz 6.3.1. *Es sei $(V, \langle\cdot,\cdot\rangle)$ ein Euklidischer oder unitärer Vektorraum und $M = \{\vec{v}_1, \ldots, \vec{v}_k\}$ ein ONS. Dann gilt*

i) Jedes Orthogonalsystem von V ist linear unabhängig.

ii) Für $\vec{x} \in \langle M\rangle$ gilt

$$\vec{x} = \sum_{i=1}^{k} \langle \vec{x}, \vec{v}_i\rangle \cdot \vec{v}_i.$$

iii) Es seien $\vec{x}, \vec{y} \in \langle M\rangle$ mit $\vec{x} = \sum_{i=1}^{k} \lambda_i \vec{v}_i$ und $\vec{y} = \sum_{i=1}^{k} \mu_i \vec{v}_i$. Dann ist

$$\langle \vec{x}, \vec{y}\rangle = \sum_{i=1}^{k} \lambda_i \overline{\mu_i}.$$

Beweis. i) Es sei T ein Orthogonalsystem und $\vec{w}_1, \ldots, \vec{w}_k \in T$ sowie $\sum_{i=1}^{k} \lambda_i \vec{w}_i = \vec{0}$ mit $\lambda_i \in \mathbb{R}$ (bzw. $\mathbb{C}$). Dann gilt für alle $j = 1, \ldots, n$

$$0 = \langle \vec{0}, \vec{w}_j\rangle = \left\langle \sum_{i=1}^{k} \lambda_i \vec{w}_i, \vec{w}_j \right\rangle = \sum_{i=1}^{k} \lambda_i \langle \vec{w}_i, \vec{w}_j\rangle = \lambda_j \langle \vec{w}_j, \vec{w}_j\rangle.$$

Wegen $\langle \vec{w}_j, \vec{w}_j\rangle = ||\vec{w}_j||^2 \neq 0$ folgt $\lambda_j = 0$.

ii) Wegen $\vec{x} \in \langle M\rangle$ gibt es $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ (bzw. $\mathbb{C}$) mit $\vec{x} = \sum_{i=1}^{k} \lambda_i \vec{v}_i$. Es folgt dann

$$\langle \vec{x}, \vec{v}_j\rangle = \left\langle \sum_{i=1}^{k} \lambda_i \vec{v}_i, \vec{v}_j \right\rangle = \sum_{i=1}^{k} \lambda_i \langle \vec{v}_i, \vec{v}_j\rangle = \sum_{i=1}^{k} \lambda_i \delta_{ij} = \lambda_j.$$

iii) Es gilt

$$\begin{aligned}
\langle \vec{x}, \vec{y}\rangle &= \left\langle \sum_{i=1}^{k} \lambda_i \vec{v}_i, \sum_{j=1}^{k} \mu_j \vec{v}_j \right\rangle = \sum_{i=1}^{k} \lambda_i \left\langle \vec{v}_i, \sum_{j=1}^{k} \mu_j \vec{v}_j \right\rangle = \sum_{i=1}^{k} \lambda_i \overline{\left\langle \sum_{j=1}^{k} \mu_j \vec{v}_j, \vec{v}_i \right\rangle} \\
&= \sum_{i=1}^{k} \lambda_i \sum_{j=1}^{k} \overline{\mu_j} \cdot \underbrace{\overline{\langle \vec{v}_j, \vec{v}_i\rangle}}_{=\delta_{ij}} = \sum_{i=1}^{k} \lambda_i \overline{\mu_i}.
\end{aligned}$$

□

Definition 6.3.2. Es sei $(V, \langle\cdot,\cdot\rangle)$ ein Euklidischer oder unitärer Vektorraum. Eine Basis $\mathcal{B}$ von V heißt Orthonormalbasis (ONB) von V, wenn die Menge ihrer Elemente ein ONS bildet.

Satz 6.3.2. *Es sei $n \in \mathbb{N}$.*

i) Es sei $\mathcal{A} \in \mathbb{R}^{(n,n)}$, und es seien $\vec{b}_1, \ldots, \vec{b}_n$ die Spalten- oder Zeilenvektoren von $\mathcal{A}$. Dann ist $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ genau dann eine ONB von $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$, wenn $\mathcal{A} \in O(n)$ ist.

ii) Es sei $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ eine ONB von $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ und $\mathcal{B}' = \{\vec{b}'_1, \ldots, \vec{b}'_n\}$ mit $\vec{b}'_j \in \mathbb{R}^n$ eine weitere Basis. Dann gibt es genau ein $\mathcal{A} \in \mathbb{R}^{(n,n)}$ mit $\vec{b}'_j = \mathcal{A}\vec{b}_j$ für alle $1 \leq j \leq n$.
Die Basis $\mathcal{B}'$ ist genau dann eine ONB des $\mathbb{R}^n$, wenn $\mathcal{A} \in O(n)$ gilt.

iii) Die Aussagen (i) und (ii) bleiben richtig, wenn $\mathbb{R}$ durch $\mathbb{C}$ und $O(n)$ durch $U(n)$ ersetzt werden.

Beweis. i) Es ist $\mathcal{A}^T\mathcal{A} = (\langle \vec{b}_i, \vec{b}_j \rangle)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$. Somit ist $\mathcal{B}$ genau dann eine ONB, wenn $\mathcal{A}^T\mathcal{A} = E_n$ ist.
Die Aussage für Zeilenvektoren folgt durch Betrachtung von $\mathcal{A}\mathcal{A}^T$.

ii) Es sei $\mathcal{C}$ bzw. $\mathcal{C}'$ die Matrix mit den Spaltenvektoren $\vec{b}_1, \ldots, \vec{b}_n$ bzw. $\vec{b}'_1, \ldots, \vec{b}'_n$. Die Matrix $\mathcal{C}$ ist regulär. Es gilt dann $\mathcal{C}'\mathcal{C}^{-1}\vec{b}_j = \mathcal{C}'\vec{e}_j = \vec{b}'_j$. Also ist $\mathcal{A} = \mathcal{C}'\mathcal{C}^{-1}$.
Für den zweiten Teil gilt

$$\mathcal{B}' \text{ ist ONB} \Leftrightarrow \langle \vec{b}'_i, \vec{b}'_k \rangle = \delta_{ik} \Leftrightarrow \langle \mathcal{A}\vec{b}_i, \mathcal{A}\vec{b}_k \rangle = \delta_{ik}.$$

Damit ist nun noch die Äquivalenz $\langle \mathcal{A}\vec{b}_i, \mathcal{A}\vec{b}_k \rangle = \delta_{ik} \Leftrightarrow \mathcal{A} \in O(n)$ zu zeigen:
Für $\mathcal{A} \in O(n)$ gilt $\langle \mathcal{A}\vec{b}_i, \mathcal{A}\vec{b}_k \rangle = \langle \vec{b}_i, \vec{b}_k \rangle = \delta_{ik}$.
Nun gelte $\langle \mathcal{A}\vec{b}_i, \mathcal{A}\vec{b}_k \rangle = \delta_{ik}$. Dann ist $\{\mathcal{A}\vec{b}_1, \ldots, \mathcal{A}\vec{b}_n\}$ eine ONB des $\mathbb{R}^n$. Es seien $\vec{x}, \vec{y} \in \mathbb{R}^n$ mit $\vec{x} = \sum_{i=1}^n \lambda_i \vec{b}_i$ und $\vec{y} = \sum_{i=1}^n \mu_i \vec{b}_i$. Somit gilt

$$\langle \mathcal{A}\vec{x}, \mathcal{A}\vec{y} \rangle = \left\langle \sum_{i=1}^n \lambda_i \mathcal{A}\vec{b}_i, \sum_{j=1}^n \mu_j \mathcal{A}\vec{b}_j \right\rangle = \sum_{i=1}^n \lambda_i \sum_{j=1}^n \mu_j \underbrace{\langle \mathcal{A}\vec{b}_i, \mathcal{A}\vec{b}_j \rangle}_{=\delta_{ij}} = \sum_{i=1}^n \lambda_i \mu_i = \langle \vec{x}, \vec{y} \rangle$$

nach Satz 6.3.1 (iii). Also ist $\mathcal{A} \in O(n)$.

iii) Die neuen Aussagen folgen ebenfalls durch Ersetzen von $\mathbb{R}$ durch $\mathbb{C}$ und von $O(n)$ durch $U(n)$ in den Beweisen.

□

Definition 6.3.3. Es sei $0 \leq \psi \leq 2\pi$ und $\mathcal{B} = \{\vec{e}_1, \vec{e}_2\}$. Dann heißt $\varphi \in L(\mathbb{R}^2, \mathbb{R}^2)$ mit

$$\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{pmatrix}$$

die <u>Drehung</u> um den Winkel ψ. Die Matrix

$$D(\psi) = \begin{pmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{pmatrix}$$

heißt <u>Drehmatrix</u> (zum Winkel ψ). Weiter heißt $\varphi \in L(\mathbb{R}^2, \mathbb{R}^2)$ mit

$$\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} \cos\psi & \sin\psi \\ \sin\psi & -\cos\psi \end{pmatrix} =: D_{Sp}(\psi)$$

<u>Drehspiegelung</u>.

Satz 6.3.3. *Es gilt*

$$\begin{array}{rcl} SO(2) & = & \{D(\psi)\colon 0 \le \psi \le 2\pi\} \quad \textit{und} \\ O(2) & = & SO(2) \cup \{D_{Sp}(\psi)\colon 0 \le \psi \le 2\pi\}. \end{array}$$

Beweis. Es sei

$$\mathcal{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in O(2).$$

Nach Satz 6.3.2 (i) bilden die Spalten von $\mathcal{A}$ eine ONB des $\mathbb{R}^2$, also gilt

$$\left|\left|\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}\right|\right| = 1 \quad \text{bzw.} \quad a_{11}^2 + a_{21}^2 = 1.$$

Es gibt somit ψ mit $0 \le \psi \le 2\pi$ und $a_{11} = \cos\psi$ und $a_{21} = \sin\psi$. Aus

$$\left\langle \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} \right\rangle = 0 \quad \text{und} \quad \left|\left|\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}\right|\right| = 1$$

folgt

$$\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} = \pm \begin{pmatrix} -\sin\psi \\ \cos\psi \end{pmatrix}.$$

Der Rest der Behauptung folgt durch Betrachtung von $\det\mathcal{A}$. □

Bemerkung 6.3.1. Aus den Additionstheoremen

$$\begin{array}{rcl} \cos(\psi_1 + \psi_2) & = & \cos\psi_1 \cos\psi_2 - \sin\psi_1 \sin\psi_2 \quad \text{und} \\ \sin(\psi_1 + \psi_2) & = & \cos\psi_1 \sin\psi_2 + \sin\psi_1 \cos\psi_2 \end{array}$$

folgt

$$\begin{pmatrix} \cos\psi_1 & -\sin\psi_1 \\ \sin\psi_1 & \cos\psi_1 \end{pmatrix} \cdot \begin{pmatrix} \cos\psi_2 & -\sin\psi_2 \\ \sin\psi_2 & \cos\psi_2 \end{pmatrix} = \begin{pmatrix} \cos(\psi_1 + \psi_2) & -\sin(\psi_1 + \psi_2) \\ \sin(\psi_1 + \psi_2) & \cos(\psi_1 + \psi_2) \end{pmatrix}.$$

Damit ist die Abbildung $\Phi\colon \mathbb{R} \to SO(2)$,

$$\psi \to \begin{pmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{pmatrix}$$

eine relationstreue Abbildung der Gruppe $(\mathbb{R}, +)$ auf die Gruppe $(SO(2), \cdot)$, ein Homomorphismus. Allerdings ist Φ kein Isomorphismus, da er nicht bijektiv ist.

Wir diskutieren nun die Existenz von Orthonormalbasen. Es zeigt sich, dass für endlichdimensionale Euklidische bzw. unitäre Vektorräume stets ONB existieren.

Satz 6.3.4. *(Verfahren von Gram- Schmidt)*
Es sei $(V, \langle\cdot,\cdot\rangle)$ ein Euklidischer oder unitärer Vektorraum mit Basis $\mathcal{B} = \{\vec{w}_1, \ldots, \vec{w}_n\}$. Dann lässt sich eine ONB von $(V, \langle\cdot,\cdot\rangle)$ folgendermaßen gewinnen:
Setze $\vec{b}_1 := \vec{w}_1 \cdot ||\vec{w}_1||^{-1}$.
Sind $\vec{b}_1, \ldots, \vec{b}_{k-1}$ bereits definiert, so setze man

$$\begin{array}{rcl} \vec{v}_k & := & \vec{w}_k - \sum_{i=1}^{k-1} \langle \vec{w}_k, \vec{b}_i \rangle \cdot \vec{b}_i \quad \textit{und} \\ \vec{b}_k & := & \vec{v}_k \cdot ||\vec{v}_k||^{-1}. \end{array}$$

Dann ist $\{\vec{b}_1, \ldots, \vec{b}_n\}$ eine ONB von $(V, \langle\cdot,\cdot\rangle)$.
Für $1 \le k \le n$ ist $\{\vec{b}_1, \ldots, \vec{b}_k\}$ eine ONB des Unterraums $\langle \vec{w}_1, \ldots, \vec{w}_k \rangle$.

Beweis. Es ist $\vec{v}_k \neq \vec{0}$ wegen

$$\sum_{i=1}^{k-1} \langle \vec{w}_k, \vec{b}_i \rangle \cdot \vec{b}_i \in \langle \vec{w}_1, \dots, \vec{w}_{k-1} \rangle$$

und $\vec{w}_k \notin \langle \vec{w}_1, \dots, \vec{w}_{k-1} \rangle$ wegen der Basiseigenschaft.
Damit ist gewährleistet, dass die Schritte des Verfahrens definiert sind. Es ist klar, dass $\langle \vec{b}_k, \vec{b}_k \rangle = 1$ ist. Für $j < k$ folgt

$$\begin{aligned} \langle \vec{b}_k, \vec{b}_j \rangle &= \left\langle \frac{\vec{v}_k}{||\vec{v}_k||}, \vec{b}_j \right\rangle = \left\langle \frac{\vec{w}_k - \sum_{i=1}^{k-1} \langle \vec{w}_k, \vec{b}_i \rangle \cdot \vec{b}_i}{||\vec{v}_k||}, \vec{b}_j \right\rangle = ||\vec{v}_k||^{-1} \cdot \left(\langle \vec{w}_k, \vec{b}_j \rangle - \sum_{i=1}^{k-1} \langle \vec{w}_k, \vec{b}_i \rangle \cdot \underbrace{\langle \vec{b}_i, \vec{b}_j \rangle}_{=\delta_{ij}} \right) \\ &= ||\vec{v}_k||^{-1} \cdot \left(\langle \vec{w}_k, \vec{b}_j \rangle - \langle \vec{w}_k, \vec{b}_j \rangle \right) = 0. \end{aligned}$$

□

Korollar 6.3.1. *Jeder endlichdimensionale Vektorraum besitzt eine ONB.*

6.4 Projektionen

Definition 6.4.1. Es sei V ein Vektorraum über einem Körper K.
Eine lineare Abbildung $P \in L(V, V)$ heißt Projektion, wenn $P \circ P = P$ gilt.

Beispiel 6.4.1. Es sei $n \in \mathbb{N}$, V ein Vektorraum mit $\dim V = n$ und U ein Unterraum von V mit $\dim U = m \leq n$. Weiter sei $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ eine Basis von V, so dass $\mathcal{B}_U = \{\vec{b}_1, \dots, \vec{b}_m\}$ eine Basis von U ist. Jedes $\vec{v} \in V$ hat dann genau eine Darstellung der Form

$$\vec{v} = \lambda_1 \vec{b}_1 + \ldots + \lambda_m \vec{b}_m + \lambda_{m+1} \vec{b}_{m+1} + \ldots + \lambda_n \vec{b}_n$$

mit $\lambda_i \in K$. Wir setzen $P_U(\vec{v}) := \lambda_1 \vec{b}_1 + \ldots + \lambda_m \vec{b}_m$. Offenbar ist P_U eine Projektion.

Wir interessieren uns hauptsächlich für orthogonale Projektionen.
Im folgenden sei V stets als Euklidischer oder unitärer Vektorraum vorausgesetzt.

Definition 6.4.2. Es sei U ein Unterraum von V. Unter $U^\perp$ (lies: U senkrecht) versteht man

$$U^\perp := \{\vec{v} \in V : \vec{v} \perp \vec{u} \ \forall \vec{u} \in U\}.$$

Die Menge $U^\perp$ heißt dann auch orthogonales Komplement von U.

Satz 6.4.1. *Es sei U ein Unterraum von V.*

i) Die Menge $U^\perp$ ist ein Unterraum von V.

ii) Es gilt $U \cap U^\perp = \{\vec{0}\}$.

iii) Es gilt $U \subset (U^\perp)^\perp$.

iv) Ist $\dim U < \infty$, so ist $(U^\perp)^\perp = U$.

v) Ist $\dim V < \infty$, so ist $\dim U + \dim U^\perp = \dim V$.

Beweis. i) Wir prüfen das Unterraumkriterium (Satz 2.3.2):
Es gilt $U^\perp \neq \emptyset$ wegen $\vec{0} \in U^\perp$. Weiter seien $\vec{x}, \vec{y} \in U^\perp$ und $\lambda, \mu \in \mathbb{R}$ (bzw. $\mathbb{C}$). Dann gilt $\langle \vec{x}, \vec{u} \rangle = \langle \vec{y}, \vec{u} \rangle = 0$ für alle $\vec{u} \in U$. Somit folgt

$$\langle \lambda\vec{x} + \mu\vec{y}, \vec{u} \rangle = \lambda\langle \vec{x}, \vec{u} \rangle + \mu\langle \vec{y}, \vec{u} \rangle = 0$$

für alle $\vec{u} \in U$. Also gilt auch $\lambda\vec{x} + \mu\vec{y} \in U^\perp$.

ii) Es sei $\vec{x} \in U \cap U^\perp$. Aus $\langle \vec{x}, \vec{x} \rangle = 0$ folgt dann aber $\vec{x} = \vec{0}$.

iii) Es sei $\vec{w} \in U$. Dann gilt $\vec{w} \perp \vec{v}$ für alle $\vec{v} \in U^\perp$. Also gilt auch $\vec{w} \in (U^\perp)^\perp$.

iv) Es sei $\dim U = n$ und $\vec{x} \in (U^\perp)^\perp$, und wir nehmen an, es gelte $\vec{x} \notin U$.
Wir setzen $W = \langle U \cup \{\vec{x}\} \rangle$, womit $\dim W = \dim U + 1$ gilt. Nach Satz 6.3.4 (Gram- Schmidt) gibt es eine Orthonormalbasis $\{\vec{b}_1, \ldots, \vec{b}_n, \vec{b}_{n+1}\}$ von W, so dass $\{\vec{b}_1, \ldots, \vec{b}_n\}$ eine ONB von U ist. Dann ist $\vec{x} = \lambda_1\vec{b}_1 + \ldots + \lambda_n\vec{b}_n + \lambda_{n+1}\vec{b}_{n+1}$ mit $\lambda_i \in \mathbb{R}$ (bzw. $\mathbb{C}$). Wegen $\vec{b}_{n+1} \in U^\perp$ folgt $\langle \lambda_{n+1}\vec{b}_{n+1}, \lambda_1\vec{b}_1 + \ldots + \lambda_n\vec{b}_n \rangle = 0$ und damit

$$\langle \lambda_{n+1}\vec{b}_{n+1}, \lambda_{n+1}\vec{b}_{n+1} \rangle = \langle \lambda_{n+1}\vec{b}_{n+1}, \lambda_1\vec{b}_1 + \ldots + \lambda_n\vec{b}_n + \lambda_{n+1}\vec{b}_{n+1} \rangle = \langle \lambda_{n+1}\vec{b}_{n+1}, \vec{x} \rangle = 0,$$

da $\vec{x} \in (U^\perp)^\perp$ gilt. Wegen der positiven Definitheit von $\langle \cdot, \cdot \rangle$ folgt aber $\lambda_{n+1} = 0$, also $\vec{x} \in U$, im Widerspruch zur Annahme.
Somit gilt $(U^\perp)^\perp \subset U$. Mit ii) folgt die Behauptung.

v) Nach Satz 6.3.4 gibt es eine ONB $\{\vec{b}_1, \ldots, \vec{b}_m\}$ von U, die zu einer ONB $\{\vec{b}_1, \ldots, \vec{b}_m, \vec{b}_{m+1}, \ldots, \vec{b}_n\}$ von V ergänzt werden kann. Es sei $\vec{x} = \lambda_1\vec{b}_1 + \ldots + \lambda_m\vec{b}_m + \lambda_{m+1}\vec{b}_{m+1}$ mit $\lambda_{m+1}\vec{b}_{m+1} \in U^\perp$. Wegen $\langle \vec{x}, \vec{b}_j \rangle = \lambda_j$ für $1 \leq j \leq n$ folgt $\lambda_1 = \ldots = \lambda_m = 0$. Also ist $\{\vec{b}_{m+1}, \ldots, \vec{b}_n\}$ eine Basis von $U^\perp$ und $\dim U^\perp = n - m$.

□

Satz 6.4.2. *Es sei U ein endlichdimensionaler Unterraum von V. Für $\vec{v} \in V$ gibt es genau eine Zerlegung $\vec{v} = \vec{u} + \vec{w}$ mit $\vec{u} \in U$ und $\vec{w} \in U^\perp$. Ist $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ eine ONB von U, so ist*

$$\vec{u} = \sum_{j=1}^{n} \langle \vec{v}, \vec{b}_j \rangle \cdot \vec{b}_j. \qquad (*)$$

Es ist $\vec{w} = \vec{0} \Leftrightarrow \vec{v} \in U$ und $\vec{u} = \vec{0} \Leftrightarrow \vec{v} \in U^\perp$.

Beweis. <u>Existenz:</u>
Es sei $\vec{v} \in V$, und wir können $\vec{v} \notin U$ annehmen. Wir setzen $W = \langle U \cup \{\vec{v}\} \rangle$. Nach Satz 6.3.4 gibt es eine Orthonormalbasis $\mathcal{B}$ von W mit $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n, \vec{b}_{n+1}\}$, so dass $\{\vec{b}_1, \ldots, \vec{b}_n\}$ eine ONB von U ist. Weiter gibt es $\lambda_i \in \mathbb{R}$ (bzw. $\mathbb{C}$) mit $\vec{v} = \lambda_1\vec{b}_1 + \ldots + \lambda_n\vec{b}_n + \lambda_{n+1}\vec{b}_{n+1}$. Nun setzen wir $\vec{u} = \lambda_1\vec{b}_1 + \ldots + \lambda_n\vec{b}_n$ und $\vec{w} = \lambda_{n+1}\vec{b}_{n+1}$. Wegen $\lambda_i = \langle \vec{v}, \vec{b}_i \rangle$ ergibt sich die Gleichung $(*)$.

<u>Eindeutigkeit:</u>
Es sei $\vec{v} = \vec{u}_1 + \vec{w}_1 = \vec{u}_2 + \vec{w}_2$ mit $\vec{u}_i \in U$ und $\vec{w}_i \in U^\perp$. Es folgt

$$\underbrace{\vec{u}_1 - \vec{u}_2}_{\in U} = \underbrace{\vec{w}_2 - \vec{w}_1}_{\in U^\perp} \in U \cap U^\perp = \{\vec{0}\}.$$

□

Definition 6.4.3. Es sei U ein endlichdimensionaler Unterraum von V und $\vec{v} \in V$. Unter der orthogonalen Projektion $P_U(\vec{v})$ von $\vec{v}$ auf U versteht man den nach Satz 6.4.2 eindeutig bestimmten Vektor $P_U(\vec{v}) \in U$, für den $\vec{v} = P_U(\vec{v}) + \vec{w}$ mit $\vec{w} \in U^{\perp}$ gilt.

Satz 6.4.3. *Es sei U ein endlichdimensionaler Unterraum von V. Die orthogonale Projektion P_U auf U ist eine Projektion im Sinne von Definition 6.4.1, d.h. es gilt $P_U^2 = P_U$. Es ist* $\operatorname{Kern} P_U = U^{\perp}$ *und $P_U(V) = U$.*

Beweis. Durch Nachrechnen. □

Die Projektion kann nun auch durch die Minimumeigenschaft des Abstandes definiert werden.

Definition 6.4.4. Es sei $\vec{v} \in V$ und $\mathcal{M} \subset V$. Unter dem Abstand $d(\vec{v}, \mathcal{M})$ des Vektors $\vec{v}$ von der Menge $\mathcal{M}$ versteht man

$$d(\vec{v}, \mathcal{M}) = \inf\{||\vec{v} - \vec{x}|| : \vec{x} \in \mathcal{M}\}.$$

Satz 6.4.4. *Es sei $\vec{v} \in V$ und U ein endlich- dimensionaler Unterraum von V. Des weiteren gelte $\vec{v} = P_U(\vec{v}) + \vec{w}$, $\vec{y} \in U$ und $\vec{x} = P_U(\vec{v}) + \vec{y} \in U$. Dann gilt*

i) $d(\vec{v}, U) = ||\vec{v} - P_U(\vec{v})|| = ||\vec{w}||$

ii) Satz des Pythagoras: $||\vec{v} - \vec{x}||^2 = ||\vec{w}||^2 + ||\vec{y}||^2$

iii) $d(\vec{v}, U) = 0 \Leftrightarrow \vec{v} \in U$.

Beweis. Aus der Aussage ii) folgen i) und iii). Daher genügt es, ii) zu beweisen:
Wegen $\vec{y} \in U$ und $\vec{w} \in U^{\perp}$ gilt $\langle \vec{y}, \vec{w} \rangle = 0$. Daher folgt

$$||\vec{v} - \vec{x}||^2 = ||\vec{w} - \vec{y}||^2 = \langle \vec{w} - \vec{y}, \vec{w} - \vec{y} \rangle = ||\vec{w}||^2 + ||\vec{y}||^2.$$

□

Bemerkung 6.4.1. Satz 6.4.4 sagt, dass die Projektion $P_U(\vec{v})$ derjenige Vektor des Unterraums U ist, der $\vec{v}$ "am nächsten liegt".

Kapitel 7

Eigenwerte, Normalformen

7.1 Eigenwerte und Eigenvektoren

Im folgenden sei stets K als Körper und V als Vektorraum über K vorausgesetzt.

Definition 7.1.1. Es sei $\varphi \in L(V, V)$ ein Endomorphismus. Dann heißt $\lambda \in K$ ein Eigenwert (EW) von φ, falls es einen Vektor $\vec{x} \in V$ mit $\vec{x} \neq \vec{0}$ gibt, so dass $\varphi(\vec{x}) = \lambda\vec{x}$ ist. Dieses $\vec{x}$ heißt dann Eigenvektor (EV) von φ zum Eigenwert λ.

Definition 7.1.2. Es sei $\mathcal{A} \in K^{(n,n)}$ eine quadratische Matrix. Dann heißt $\lambda \in K$ ein Eigenwert der Matrix $\mathcal{A}$, falls es einen Vektor $\vec{x} \in K^n$ mit $\vec{x} \neq \vec{0}$ gibt, so dass $\mathcal{A}\vec{x} = \lambda\vec{x}$ ist. Dieses $\vec{x}$ heißt dann Eigenvektor von $\mathcal{A}$ zum Eigenwert λ, d.h. λ ist Eigenwert des Endomorphismus $\varphi(\vec{x}) = \mathcal{A}\vec{x}$.

Beispiel 7.1.1. Die Identität $\varphi = id$ hat den Eigenwert 1 und sonst keinen. Die Nullabbildung $N(\vec{x}) = \vec{0}$ für alle $\vec{x} \in V$ hat als einzigen Eigenwert die Null. Ein Endomorphismus φ hat genau dann den Eigenwert 0, wenn es ein $\vec{x} \neq \vec{0}$ mit $\varphi(\vec{x}) = \vec{0}$ gibt.

Beispiel 7.1.2. Es sei die Projektion $P\colon K^n \to K^n$ mit $P(x_1, \ldots, x_n)^T = (x_1, \ldots, x_r, 0, \ldots, 0)^T$ für $1 < r < n$ gegeben. Sie besitzt den Eigenwert 1, denn jeder Vektor der Form $\vec{x} = (x_1, \ldots, x_r, 0, \ldots, 0)^T$ erfüllt $P(\vec{x}) = \vec{x}$. Andererseits besitzt die Projektion P auch den Eigenwert 0, denn jeder Vektor der Form $\vec{x}' = (0, \ldots, 0, x_{r+1}, \ldots, x_n)^T$ erfüllt $P(\vec{x}') = \vec{0}$.

Satz 7.1.1. *Es sei $\varphi \in L(V, V)$, $\mathcal{B}$ eine Basis von V und $\lambda \in K$. Dann ist λ genau dann ein Eigenwert von φ, wenn λ ein Eigenwert von $\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B})$ ist.*

Bemerkung 7.1.1. Offensichtlich hängt die Aussage auf der linken Seite nicht von der Wahl der Basis $\mathcal{B}$ ab. Damit folgt aus Satz 3.7.5, dass $\mathcal{A} \in K^{(n,n)}$ und $\mathcal{X}^{-1}\mathcal{A}\mathcal{X}$ für alle $\mathcal{X} \in GL(n, K)$ diselben Eigenwerte besitzen. Aus Satz 5.4.1 folgt übrigens, dass sie auch dieselbe Determinante besitzen und sie haben nach Satz 3.5.8 auch denselben Rang.

Dadurch wird die folgende Definition motiviert:

Definition 7.1.3. Matrizen $\mathcal{A}, \mathcal{B} \in K^{(n,n)}$ heißen ähnlich (Schreibweise: $A \approx B$), wenn es ein reguläres $\mathcal{X} \in GL(n, K)$ mit $\mathcal{B} = \mathcal{X}^{-1}\mathcal{A}\mathcal{X}$ gibt.

Daraus folgt:

i) $A \approx B \Leftrightarrow \mathrm{rg}(\mathcal{A}) = \mathrm{rg}(\mathcal{B})$ und $\det \mathcal{A} = \det \mathcal{B}$

ii) Ist $\varphi \in L(V,V)$ und $\mathcal{B}$ eine Basis von V mit $\mathcal{A} = \mathcal{M}(\varphi; \mathcal{B}, \mathcal{B})$, so folgt aus den Sätzen aus Kapitel 3 die Äquivalenz

$$\mathcal{A} \approx \mathcal{B} \Leftrightarrow \quad \text{Es gibt eine Basis } \tilde{\mathcal{B}} \text{ von } V \text{ mit } \mathcal{B} = \mathcal{M}(\varphi; \tilde{\mathcal{B}}, \tilde{\mathcal{B}}).$$

Beweis. (Beweis von Satz 7.1.1)
Es sei $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ eine Basis von V und $\vec{x} = \sum_{i=1}^n \mu_i \vec{b}_i$ sowie $\varphi(\vec{x}) = \sum_{i=1}^n \mu_i' \vec{b}_i$ mit $\mu_i, \mu_i' \in K$. Nach Satz 3.5.2 gilt

$$\begin{pmatrix} \mu_1' \\ \vdots \\ \mu_n' \end{pmatrix} = \mathcal{M}(\varphi; \mathcal{B}, \mathcal{B}) \cdot \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}.$$

Nun ist λ genau dann Eigenwert von φ, wenn $\varphi(\vec{x}) = \lambda \vec{x}$ für ein $\vec{x} \neq \vec{0}$ ist, also für

$$\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} \neq \vec{0} \quad \text{und} \quad \mathcal{M}(\varphi; \mathcal{B}', \mathcal{B}) \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1' \\ \vdots \\ \lambda_m' \end{pmatrix} \stackrel{!}{=} \lambda \cdot \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}.$$

Dies ist genau dann erfüllt, wenn $(\mu_1, \ldots, \mu_n)^T$ Eigenvektor von $\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B})$ zum Eigenwert λ ist. $\square$

Definition 7.1.4. Es sei $\lambda \in K$ ein Eigenwert von $\varphi \in L(V,V)$. Dann heißt

$$U_\lambda = \{\vec{x} \in V : \varphi(\vec{x}) = \lambda \vec{x}\} = \text{Kern}(\varphi - \lambda id)$$

der Eigenraum von λ. Die Elemente von U_λ sind also die Eigenvektoren zum Eigenwert λ und der Nullvektor. Weiter heißt $\dim U_\lambda = \dim \text{Kern}(\varphi - \lambda id)$ die geometrische Ordnung (oder die geometrische Vielfachheit) des Eigenwerts λ. Entsprechend definiert man für Matrizen $\mathcal{A} \in K^{(n,n)}$

$$U_\lambda = \{\vec{x} \in K^n : (\mathcal{A} - \lambda \mathcal{E}_n)\vec{x} = \vec{0}\}$$

und $\dim U_\lambda = \dim \text{Kern}(\mathcal{A} - \lambda \mathcal{E}_n) = n - \text{rg}(\mathcal{A} - \lambda \mathcal{E}_n)$.

Satz 7.1.2. *Es seien $\lambda_1, \ldots, \lambda_r$ paarweise verschiedene Eigenwerte von $\varphi \in L(V,V)$ und $\vec{x}_1, \ldots, \vec{x}_r$ die zugehörigen Eigenvektoren. Dann sind diese $\vec{x}_1, \ldots, \vec{x}_r$ linear unabhängig.*

Bemerkung 7.1.2. Folglich besitzt φ höchstens $n = \dim V$ verschiedene Eigenwerte.

Beweis. (Beweis von Satz 7.1.2)
Es ist $\vec{x}_i$ genau dann Eigenvektor zum Eigenwert λ, wenn $\vec{x}_i \neq 0$ ist und $\varphi(\vec{x}_i) = \lambda \vec{x}_i$ für $i = 1, \ldots, r$ erfüllt ist. Weiter gilt für $1 \leq i, j \leq r$

$$(\varphi - \lambda_i id)(\vec{x}_j) = \varphi(\vec{x}_j) - \lambda_i \vec{x}_j = \lambda_j \vec{x}_j - \lambda_i \vec{x}_j = (\lambda_j - \lambda_i)\vec{x}_j. \tag{$*$}$$

Nun betrachten wir $\alpha_1 \vec{x}_1 + \ldots + \alpha_r \vec{x}_r = \vec{0}$ für $\alpha_1, \ldots, \alpha_r \in K$ und müssen $\alpha_k = 0$ für alle $1 \leq k \leq r$ zeigen. Auf diese Gleichung wenden wir die lineare Abbildung

$$\psi_k = (\varphi - \lambda_1 id) \circ \ldots \circ (\varphi - \lambda_{k-1} id) \circ (\varphi - \lambda_k id) \circ \ldots \circ (\varphi - \lambda_r id)$$

an und erhalten

$$\vec{0} = \psi_k(\vec{0}) = \psi_k\left(\sum_{j=1}^r \alpha_j \vec{x}_j\right) = \sum_{j=1}^r \alpha_j \psi_k(\vec{x}_j) = \sum_{j=1}^r \alpha_j \prod_{\substack{i=1 \\ i \neq k}}^r (\varphi - \lambda_i id)\vec{x}_j \underset{(*)}{=} \sum_{j=1}^r \alpha_j \prod_{\substack{i=1 \\ i \neq k}}^r (\lambda_j - \lambda_i)\vec{x}_j.$$

Das erste Produkt gehört zur Verknüpfung $\circ$ auf $L(V,V)$, das zweite ist ein gewöhnliches Produkt von Elementen aus K. In der letzten Summe verschwinden alle Summanden bis auf den Summanden für $i=k$, also ist

$$\vec{0} = \alpha_k \prod_{\substack{i=1 \\ i\neq k}}^{r} (\lambda_k - \lambda_i)\vec{x}_k.$$

Da die Eigenwerte paarweise verschieden sind, verschwindet das Produkt nicht. Wegen $\vec{x} \neq \vec{0}$ gilt dann $\vec{a}_k = 0$. □

Satz 7.1.3. *Für $\varphi \in L(V,V)$ und $\mathcal{A} \in K^{(n,n)}$ sowie $\lambda \in K$ gilt*

i) Der Skalar λ ist genau dann Eigenwert von φ, wenn $\det(\varphi - \lambda id) = 0$ gilt.

ii) Der Skalar λ ist genau dann Eigenwert von $\mathcal{A}$, wenn $\det(\mathcal{A} - \lambda\mathcal{E}_n) = 0$ gilt.

Beweis. i) Es gilt

$$\begin{aligned} \lambda \text{ ist Eigenwert von } \varphi \quad &\Leftrightarrow \quad \operatorname{Kern}(\varphi - \lambda id) \neq \{\vec{0}\} \Leftrightarrow \varphi - \lambda id \text{ ist nicht injektiv} \\ &\Leftrightarrow \quad \varphi - \lambda id \text{ ist kein Automorphismus} \Leftrightarrow \det(\varphi - \lambda id) = 0. \end{aligned}$$

ii) Es gilt völlig analog

$$\begin{aligned} \lambda \text{ ist Eigenwert von } \mathcal{A} \quad &\Leftrightarrow \quad \exists \vec{x} \neq \vec{0} \colon \mathcal{A}\vec{x} = \lambda\vec{x} \Leftrightarrow \exists \vec{x} \neq \vec{0} \colon (\mathcal{A} - \lambda\mathcal{E}_n)\vec{x} = \vec{0} \\ &\Leftrightarrow \quad \mathcal{A} - \lambda\mathcal{E}_n \text{ ist nicht regulär} \Leftrightarrow \det(\mathcal{A} - \lambda\mathcal{E}_n) = 0. \end{aligned}$$

□

Satz 7.1.4. *Für $\mathcal{A} = (a_{ij}) \in K^{(n,n)}$ ist $\det(\mathcal{A} - \lambda\mathcal{E}_n)$ ein Polynom in λ vom Grad n mit Koeffizienten aus dem Körper K:*

$$P_{\mathcal{A}}(\lambda) = \det(\mathcal{A} - \lambda\mathcal{E}_n) = \alpha_0 + \alpha_1\lambda + \alpha_2\lambda^2 + \ldots + \alpha_n\lambda^n.$$

Es ist zudem $\alpha_n = (-1)^n$, $\alpha_{n-1} = (-1)^{n-1} \cdot (a_{11} + \ldots + a_{nn})$ und $\alpha_0 = \det\mathcal{A}$.

Beweis. Es gilt

$$\mathcal{A} - \lambda\mathcal{E}_n = \begin{pmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{pmatrix} = (b_{ij}).$$

Es folgt

$$\det(\mathcal{A} - \lambda\mathcal{E}_n) = (a_{11} - \lambda) \cdot (a_{22} - \lambda) \cdots (a_{nn} - \lambda) + \sum_{\substack{\sigma \in S_n \\ \sigma \neq id}} \operatorname{sgn}(\sigma) \cdot b_{1,\sigma(1)} \cdot b_{2,\sigma(2)} \cdots b_{n,\sigma(n)}.$$

Jeder Summand ist ein Produkt aus einem Skalar aus K und einem Polynom in λ, welches höchstens den Grad $n-2$ besitzt. Also ist auch die Summe ein Polynom $Q(\lambda)$, welches höchstens den Grad $n-2$ besitzt. Wir multiplizieren das Produkt über die $(a_{ii} - \lambda)$ aus und erhalten

$$\det(\mathcal{A} - \lambda\mathcal{E}_n) = (-1)^n\lambda^n + (-\lambda)^{n-1}(a_{11} + \ldots + a_{nn}) + \tilde{Q}(\lambda)$$

mit einem Polynom $\tilde{Q}(\lambda)$, welches ebenfalls höchstens den Grad $n-2$ besitzt. Damit ist $\det(\mathcal{A} - \lambda\mathcal{E}_n)$ ein Polynom vom Grad n, und es ist $\alpha_n = (-1)^n$ und $\alpha_{n-1} = (-1)^{n-1} \cdot (a_{11} + \ldots + a_{nn})$. Den Wert für α_0 erhält man durch Einsetzen von $\lambda = 0$. Daraus folgt $\alpha_0 = \det\mathcal{A}$. □

Definition 7.1.5. Es sei $\mathcal{A} \in K^{(n,n)}$.
Das Polynom $P_{\mathcal{A}}(\lambda) = \det(\mathcal{A} - \lambda\mathcal{E}_n)$ heißt charakteristisches Polynom von $\mathcal{A}$.
Die Summe $\mathrm{Spur}(\mathcal{A}) = a_{11} + \ldots + a_{nn}$ heißt die Spur von $\mathcal{A}$.

Bemerkung 7.1.3. Satz 7.1.3 besagt, dass die Eigenwerte von $\mathcal{A}$ die Nullstellen des charakteristischen Polynoms sind:

$$\lambda_i \text{ Eigenwert von } \mathcal{A} \Leftrightarrow P_{\mathcal{A}}(\lambda_i) = 0.$$

Dadurch kann man die Eigenwerte einer Matrix rechnerisch bestimmen:

Beispiel 7.1.3. Bestimme die Eigenwerte von

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

Lösung:
Es gilt mit dem Entwicklungssatz von Laplace (Satz 5.3.6)

$$\begin{aligned} \det(\mathcal{A} - \lambda\mathcal{E}_n) &= \det\begin{pmatrix} 1-\lambda & 2 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 1 & 1-\lambda \end{pmatrix} \underset{\text{2. Zeile}}{=} (1-\lambda) \cdot \det\begin{pmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{pmatrix} \\ &= (1-\lambda) \cdot ((1-\lambda)^2 - 1 \cdot 1) = (1-\lambda) \cdot (-2\lambda + \lambda^2) = -\lambda \cdot (\lambda - 1) \cdot (\lambda - 2). \end{aligned}$$

Damit hat $\mathcal{A}$ die Eigenwerte $\lambda_1 = 0$, $\lambda_2 = 1$ und $\lambda_3 = 2$.

7.2 Diagonalisierung

Wir werden im folgenden die Frage der Normalformen von Matrizen behandeln:
Ist eine vorgegebene Matrix $\mathcal{A} \in K^{(n,n)}$ ähnlich (im Sinne von Definition 7.1.3) einer Matrix $\mathcal{B}$ von einer vorgeschriebenen einfachen Gestalt, d.h. gibt es eine Matrix $\mathcal{X} \in GL(n, K)$, so dass

$$\mathcal{B} = \mathcal{X}^{-1}\mathcal{A}\mathcal{X}$$

diese Gestalt hat?
Matrizen besonders einfacher Gestalt sind Diagonalmatrizen.

Definition 7.2.1. Es sei K ein Körper und $n \in \mathbb{N}$. Dann heißt $\mathcal{A} \in K^{(n,n)}$ diagonalisierbar, wenn $\mathcal{A}$ einer Diagonalmatrix

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

mit $\lambda_i \in K$ ähnlich ist.

Satz 7.2.1. *Es sei K ein Körper und $\lambda_i \in K$. Dann ist $\mathcal{A} \in K^{(n,n)}$ genau dann diagonalisierbar, wenn der K^n eine Basis aus Eigenvektoren von $\mathcal{A}$ besitzt. Ist $\mathcal{X}$ die Matrix, deren Spaltenvektoren die Eigenvektoren sind, so gilt*

$$\mathcal{X}^{-1}\mathcal{A}\mathcal{X} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

wobei die λ_i die Eigenwerte von $\mathcal{A}$ sind. Jeder Eigenwert tritt so oft auf, wie seine geometrische Ordnung angibt.

Beweis. "$\Rightarrow$":
Die verschiedenen Eigenwerte seien $\lambda_1, \ldots, \lambda_r$ mit zugehörigen Eigenräumen $U_{\lambda_1}, \ldots, U_{\lambda_r}$. Weiter sei $\dim U_{\lambda_i} = \rho_i$ für $1 \leq i \leq r$. Die Basis aus Eigenvektoren sei $\mathcal{B} = \mathcal{B}_{\lambda_1} \cup \ldots \cup \mathcal{B}_{\lambda_r}$, wobei $\mathcal{B}_{\lambda_i} \subset U_{\lambda_i}$ ist. Es sei $\mathcal{B}_{\lambda_i} = \{\vec{b}_1(\lambda_i), \ldots, \vec{b}_{\tau_i}(\lambda_i)\}$, also $|\mathcal{B}_{\lambda_i}| = \tau_i$. Es ist nun zu zeigen, dass $\rho_i = \tau_i$ für alle $1 \leq i \leq r$ gilt:
Annahme: $\rho_j < \tau_j$:
Dann enthält U_{λ_j} die τ_j linear unabhängigen Vektoren von $\mathcal{B}_{\lambda_j}$ im Widerspruch zu $\dim \mathcal{B}_{\lambda_j} = \rho_j$.
Annahme: $\tau_j < \rho_j$:
Dann gibt es $\vec{c}_j \in U_{\lambda_j}$ mit $\vec{c}_j \notin \langle \mathcal{B}_{\lambda_j} \rangle$. Wir wollen zeigen, dass

$$\mathcal{B} \cup \{\vec{c}_j\} \tag{$*$}$$

linear unabhängig ist. Es sei

$$\mu_1(\lambda_1)\vec{b}_1(\lambda_1) + \ldots + \mu_{\tau_1}(\lambda_1)\vec{b}_{\tau_1}(\lambda_1) + \ldots + \mu_1(\lambda_r)\vec{b}_1(\lambda_r) + \ldots + \mu_{\tau_r}(\lambda_r)\vec{b}_{\tau_r}(\lambda_r) + \nu\vec{c}_j = \vec{0}.$$

Da nach Satz 7.1.2 Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig sind, folgt

$$\mu_1(\lambda_i)\vec{b}_1(\lambda_i) + \ldots + \mu_{\tau_i}(\lambda_i)\vec{b}_{\tau_i}(\lambda_i) = \vec{0}$$

für $i \neq j$ und

$$\mu_1(\lambda_j)\vec{b}_1(\lambda_j) + \ldots + \mu_{\tau_j}(\lambda_j)\vec{b}_{\tau_j}(\lambda_j) + \nu\vec{c}_j = \vec{0}.$$

Wäre $\nu \neq 0$, so folgte $\vec{c}_j \in \langle \mathcal{B}_{\lambda_j} \rangle$, im Widerspruch zur Annahme. Also gilt $\nu = 0$, und wegen der Basiseigenschaft der $\mathcal{B}_{\lambda_j}$ gilt $\mu_k(\lambda_j) = 0$ für alle $1 \leq k \leq \tau_j$. Wegen der Basiseigenschaft der $\mathcal{B}_i$ für $i \neq j$ folgt $\mu_k(\lambda_i) = 0$ für $i \neq j$. Damit ist $(*)$ gezeigt.
Dies steht jedoch im Widerspruch zur Voraussetzung, dass $\mathcal{B}$ eine Basis des K^n darstellt.

Damit ist $\rho_i = \tau_i$ gezeigt.

Nach Definition 3.5.5 ist

$$\mathcal{M}(\varphi_j; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Nach Satz 3.7.5 ist

$$\mathcal{X}^{-1}\mathcal{A}\mathcal{X} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

"$\Leftarrow$":
Es sei

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

und $\varphi_{\mathcal{A}}(\vec{x}) = \mathcal{A}\vec{x}$. Nach Satz 3.7.5 ist

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \in \mathcal{M}(\varphi_{\mathcal{A}}; \mathcal{C}, \mathcal{C})$$

mit $\mathcal{C} = \{\vec{x}_1, \ldots, \vec{x}_n\}$. Es folgt $\varphi_{\mathcal{A}}(\vec{x}_i) = \lambda_i \vec{x}_i$. $\square$

Wir kehren zu Beispiel 3.7.2 zurück, in dem wir die Diagonalisierung einer Matrix angegeben hatten, ohne zu beschreiben, wie sie zu finden war.

Beispiel 7.2.1. Die Matrix

$$\mathcal{A} = \begin{pmatrix} -3 & -4 & -4 \\ -3 & 0 & -3 \\ 8 & 4 & 9 \end{pmatrix}$$

ist zu diagonalisieren.
<u>Berechnung der Eigenwerte:</u>
Das charakteristische Polynom ergibt sich zu

$$\begin{aligned} P_{\mathcal{A}}(\lambda) &= \det \begin{pmatrix} -3-\lambda & -2 & -4 \\ -3 & -\lambda & -3 \\ 8 & 4 & 9-\lambda \end{pmatrix} \\ &\underset{\text{1. Zeile}}{=} (-3-\lambda) \cdot \det \begin{pmatrix} -\lambda & -3 \\ 4 & 9-\lambda \end{pmatrix} + 2 \cdot \det \begin{pmatrix} -3 & -3 \\ 8 & 9-\lambda \end{pmatrix} - 4 \cdot \det \begin{pmatrix} -3 & -\lambda \\ 8 & 4 \end{pmatrix} \\ &= -(\lambda+3) \cdot (\lambda^2 - 9\lambda + 12) + 2 \cdot (3\lambda - 27 + 24) - 4 \cdot (8\lambda - 12) = -\lambda^3 + 6\lambda^2 - 11\lambda + 6. \end{aligned}$$

Durch Probieren findet man $\lambda_1 = 1$ als Lösung der Gleichung $P_{\mathcal{A}}(\lambda) = 0$. Polynomdivision ergibt

$$P_{\mathcal{A}}(\lambda) = -(\lambda - 1) \cdot (\lambda^2 - 5\lambda + 6) = -(\lambda - 1) \cdot (\lambda - 2) \cdot (\lambda - 3).$$

Man erhält also die Eigenwerte $\lambda_1 = 1$, $\lambda_2 = 2$ und $\lambda_3 = 3$. Die zugehörigen Eigenvektoren $\vec{b}_1, \vec{b}_2, \vec{b}_3$ erhält man durch Lösung des LGS

$$(\mathcal{A} - \lambda_i \mathcal{E}_n) \cdot \vec{x} = \vec{0},$$

etwa für $\lambda_1 = 1$

$$(\mathcal{A} - \lambda_1 \mathcal{E}_n) \cdot \vec{x} = \vec{0} \Rightarrow \begin{pmatrix} -4 & -4 & -4 \\ -3 & -1 & -3 \\ 8 & 4 & 8 \end{pmatrix} \vec{x} = \vec{0} \Rightarrow \vec{x} = \vec{b}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Analoges ergibt sich für $\vec{b}_2$ und $\vec{b}_3$, womit sich die schon in Beispiel 3.7.2 angegebenen Vektoren

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

ergeben. Es sei $\mathcal{X}$ die Matrix, die $\vec{b}_1, \vec{b}_2, \vec{b}_3$ als Spaltenvektoren besitzt, also

$$\mathcal{X} = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & -1 \\ -1 & 4 & 2 \end{pmatrix}.$$

Dann ist

$$\mathcal{X}^{-1} \mathcal{A} \mathcal{X} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix},$$

d.h. die Diagonalmatrix mit den Eigenwerten $\lambda_1 = 1$, $\lambda_2 = 2$ und $\lambda_3 = 3$ auf der Diagonalen.

7.3 Normalformen von symmetrischen und unitären Matrizen und der Spektralsatz

Es sei $\mathcal{A} \in \mathbb{C}^{(n,n)}$.
Wir benötigen zuerst die Tatsache, dass das charakteristische Polynom $P_{\mathcal{A}}(\lambda)$ mindestens eine Nullstelle und damit $\mathcal{A}$ mindestens einen Eigenwert besitzt. Dies folgt aus dem sogenannten "Fundamentalsatz der Algebra", den wir hier ohne Beweis angeben:

Satz 7.3.1. *(Fundamentalsatz der Algebra)*
Es sei $n \in \mathbb{N}$, also $n \geq 1$ und

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

mit $a_i \in \mathbb{C}$. Dann hat p mindestens eine komplexe Nullstelle.

Definition 7.3.1. Es sei $(V, \langle\cdot,\cdot\rangle)$ ein Euklidischer Vektorraum. Ein Endomorphismus $\varphi \in L(V,V)$ heißt symmetrisch, wenn $\langle\vec{x}, \varphi(\vec{y})\rangle = \langle\varphi(\vec{x}), \vec{y}\rangle$ für alle $\vec{x}, \vec{y} \in V$ gilt.

Definition 7.3.2. Es sei $(V, \langle\cdot,\cdot\rangle)$ ein unitärer Vektorraum. Ein Endomorphismus $\varphi \in L(V,V)$ heißt hermitesch (oder selbstadjungiert), wenn $\langle\vec{x}, \varphi(\vec{y})\rangle = \langle\varphi(\vec{x}), \vec{y}\rangle$ für alle $\vec{x}, \vec{y} \in V$ gilt.

Satz 7.3.2. *Es sei $(V, \langle\cdot,\cdot\rangle)$ ein endlichdimensionaler Euklidischer Vektorraum und $\mathcal{B}$ eine ONB von V. Ein Endomorphismus $\varphi \in L(V,V)$ ist genau dann symmetrisch, wenn die Darstellungsmatrix $\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B})$ symmetrisch ist.*

Beweis. Der Beweis ergibt sich aus dem Beweis des entsprechenden Satzes für unitäre Vektorräume. □

Satz 7.3.3. *Es sei $(V, \langle\cdot,\cdot\rangle)$ ein endlichdimensionaler unitärer Vektorraum und $\mathcal{B}$ eine Orthonormalbasis von V. Ein Endomorphismus $\varphi \in L(V,V)$ ist genau dann hermitesch, wenn die Darstellungsmatrix $\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B})$ hermitesch ist.*

Beweis. Es sei $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ sowie $\vec{x} = x_1\vec{b}_1 + \ldots + x_n\vec{b}_n$ und $\vec{y} = y_1\vec{b}_1 + \ldots + y_n\vec{b}_n$ mit $x_i, y_i \in \mathbb{C}$. Weiter sei $\mathcal{M} = \mathcal{M}(\varphi; \mathcal{B}, \mathcal{B})$. Nach Satz 3.5.2 ist dann

$$\varphi(\vec{x}) = x'_1\vec{b}_1 + \ldots + x'_n\vec{b}_n \qquad \text{mit} \qquad \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = \mathcal{M} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \text{und}$$

$$\varphi(\vec{y}) = y'_1\vec{b}_1 + \ldots + y'_n\vec{b}_n \qquad \text{mit} \qquad \begin{pmatrix} y'_1 \\ \vdots \\ y'_n \end{pmatrix} = \mathcal{M} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Es ist

$$\langle\vec{x}, \varphi(\vec{y})\rangle = \sum_{i,j=1}^{n} x_i \overline{y'_j} \cdot \langle\vec{b}_i, \vec{b}_j\rangle = (x_1, \ldots, x_n) \cdot \begin{pmatrix} \overline{y'_1} \\ \vdots \\ \overline{y'_n} \end{pmatrix} = (x_1, \ldots, x_n) \cdot \overline{\mathcal{M}} \cdot \begin{pmatrix} \overline{y_1} \\ \vdots \\ \overline{y_n} \end{pmatrix} \tag{1}$$

und

$$\langle\varphi(\vec{x}), \vec{y}\rangle = (x'_1, \ldots, x'_n) \cdot \begin{pmatrix} \overline{y_1} \\ \vdots \\ \overline{y_n} \end{pmatrix} = (x_1, \ldots, x_n) \cdot \mathcal{M}^T \cdot \begin{pmatrix} \overline{y_1} \\ \vdots \\ \overline{y_n} \end{pmatrix}. \tag{2}$$

Da eine Sesquilinearform die Matrix eindeutig bestimmt, folgt aus (1) und (2), dass φ genau dann hemitesch ist, wenn $\overline{\mathcal{M}} = \mathcal{M}^T$ oder $\mathcal{M} = \mathcal{M}^*$ ist. □

Satz 7.3.4. *(Spektralsatz)*

i) Es sei $(V, \langle \cdot, \cdot \rangle)$ ein endlichdimensionaler unitärer Vektorraum und $\varphi \in L(V, V)$ hermitesch. Dann gibt es eine ONB von V, die aus den Eigenvektoren von φ zu reellen Eigenwerten besteht.

ii) Es sei $\mathcal{A} \in \mathbb{C}^{(n,n)}$ eine hermitesche Matrix. Dann gibt es eine unitäre Matrix P, so dass $P^*\mathcal{A}P$ eine reelle Diagonalmatrix ist.

Beweis. i) Wir führen den Beweis durch Induktion nach $n = \dim V$.

Induktionsanfang: $n = 1$:
Setze $\vec{v}_1 = \vec{w} \cdot ||\vec{w}||^{-1}$ für $\vec{w} \in V - \{\vec{0}\}$. Dann ist $\langle \vec{v}_1, \vec{v}_1 \rangle = 1$. Wegen $\dim V = 1$ ist $\varphi(\vec{v}_1) = \lambda_1 \vec{v}_1$ mit $\lambda_1 \in \mathbb{C}$. Es ist

$$\langle \vec{v}_1, \varphi(\vec{v}_1) \rangle = \langle \vec{v}_1, \lambda_1 \vec{v}_1 \rangle = \overline{\lambda_1} = \langle \varphi(\vec{v}_1), \vec{v}_1 \rangle = \langle \lambda_1 \vec{v}_1, \vec{v}_1 \rangle = \lambda_1.$$

Aus $\lambda_1 = \overline{\lambda_1}$ folgt $\lambda_1 \in \mathbb{R}$.

Induktionsschritt: $n \to n + 1$:
Nach Satz 7.3.1 besitzt das charakteristische Polynom $\det(\varphi - \lambda id)$ mindestens eine Nullstelle. Somit hat φ einen Eigenwert λ_1 und einen Eigenvektor $\vec{w}_1$ mit $\varphi(\vec{w}_1) = \lambda_1 \vec{w}_1$. Wir setzen wieder $\vec{v}_1 = \vec{w}_1 \cdot ||\vec{w}_1||^{-1}$, womit $||\vec{v}_1|| = 1$ ist. Nach Satz 6.3.4 kann $\{\vec{v}_1\}$ zu einer ONB $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ von V ergänzt werden.
Die Matrix $\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B})$ hat die Form

$$\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B}) = \left(\begin{array}{c|ccc} \lambda_1 & * & \ldots & * \\ \hline 0 & & & \\ \vdots & & \mathcal{M}' & \\ 0 & & & \end{array} \right).$$

Nach Satz 7.3.3 ist $\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B})$ hermitesch und somit $\lambda_1 \in \mathbb{R}$,

$$\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B}) = \left(\begin{array}{c|ccc} \lambda_1 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & \mathcal{M}' & \\ 0 & & & \end{array} \right),$$

und $\mathcal{M}'$ ist hermitesch.
Nach der Induktionshypothese gibt es eine ONB $\mathcal{B}'$ von $V' = \langle \{\vec{v}_1, \ldots, \vec{v}_n\} \rangle$, so dass

$$\mathcal{M}(\varphi_{V'}; \mathcal{B}', \mathcal{B}') = \begin{pmatrix} \lambda_2 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

mit $\lambda_i \in \mathbb{R}$ ist. Es sei $\mathcal{B}'' = \{\vec{v}_1\} \cup \mathcal{B}'$. Dann ist

$$\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

ii) Wir zeigen ii), indem wir die Erkenntnisse aus Kapitel III nutzen, dass eine lineare Abbildung in einer Matrixschreibweise dargestellt werden kann und folgern das Ergebnis aus i):
Wir setzen $V = \mathbb{C}^n$ mit der ONB $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$. Es sei $\varphi\colon \mathbb{C}^n \to \mathbb{C}^n$, $\vec{x} \to \mathcal{A}\vec{x}$. Nach i) gibt es eine ONB $\mathcal{C} = \{\vec{c}_1, \ldots, \vec{c}_n\}$ von V, so dass

$$\mathcal{M}(\varphi; \mathcal{C}, \mathcal{C}) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

mit $\lambda_i \in \mathbb{R}$ ist. Es sei $\mathcal{P}$ die Matrix, deren Spalten $\vec{c}_1, \ldots \vec{c}_n$ sind. Dann ist $\mathcal{P} \in U(n)$, also $\mathcal{P}^{-1} = \mathcal{P}^*$. Nach Satz 3.7.4 ist

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

□

Wir erhalten als Folgerung:

Satz 7.3.5. *Die Eigenwerte eines hermiteschen Endomorphismus oder eine hermiteschen Matrix sind reell.*

Ganz analog werden nun die entsprechenden Sätze für Euklidische Vektorräume und orthogonale Endomorphismen gezeigt. Wir geben daher nur die Sätze an und verzichten auf ihre Beweise:

Satz 7.3.6. *(Spektralsatz, reeller Fall)*

i) Es sei $(V, \langle\cdot,\cdot\rangle)$ ein endlichdimensionaler Euklidischer Vektorraum und $\varphi \in L(V, V)$ symmetrisch. Dann gibt es eine ONB von V, die aus den Eigenvektoren von φ besteht.

ii) Es sei $\mathcal{A} \in \mathbb{R}^{(n,n)}$ eine symmetrische Matrix. Dann gibt es eine Matrix $\mathcal{P} \in O(n)$, so dass $\mathcal{P}^T\mathcal{A}\mathcal{P}$ eine Diagonalmatrix ist.

Als Folgerung von Satz 7.3.5 ergibt sich:

Satz 7.3.7. *Die Eigenwerte einer reellen symmetrischen Matrix sind reell.*

7.4 Definitheit quadratischer Formen und die Hauptachsentransformation

Definition 7.4.1. Es sei K ein Körper und V ein Vektorraum über K.
Unter einer quadratischen Form Q auf V versteht man die Abbildung

$$Q\colon V \to \mathbb{R},\ \vec{x} \to b(\vec{x}, \vec{x}),$$

wobei b eine symmetrische Bilinearform auf V ist.

Bemerkung 7.4.1. Quadratische Formen gehen also aus symmetrischen Bilinearformen durch Gleichsetzen der Variablen hervor. Es gibt somit eine Matrix $\mathcal{A} \in K^{(n.n)}$, so dass $Q(\vec{x}) = \vec{x}^T\mathcal{A}\vec{x}$ gilt.

Satz 7.4.1. *Es sei K ein Körper, in dem $1+1\neq 0$ gilt. Ist Q eine quadratische Form auf V, so ist die symmetrische Bilinearform von Definition 7.4.1 durch Q eindeutig bestimmt.*
Es gilt $b(\vec{x},\vec{y}) = \frac{1}{2}\cdot(Q(\vec{x}+\vec{y}) - Q(\vec{x}) - Q(\vec{y}))$.

Beweis. Für alle $\vec{x},\vec{y}\in K^n$ gilt $Q(\vec{x}+\vec{y}) = b(\vec{x}+\vec{y},\vec{x}+\vec{y}) = Q(\vec{x})+Q(\vec{y})+2b(\vec{x},\vec{y})$. □

Bemerkung 7.4.2. Wird auf die Forderung der Symmetrie verzichtet, so kann es mehrere Bilinearformen geben, aus denen durch Gleichsetzen dieselbe quadratische Form hervorgeht, wie das folgende Beispiel zeigt.

Beispiel 7.4.1. Es sei $V=\mathbb{R}^2$ und $\vec{x}=(x_1,x_2)^T$ bzw. $\vec{y}=(y_1,y_2)^T$ sowie

$$\begin{aligned} b_1(\vec{x},\vec{y}) &= \vec{x}^T\begin{pmatrix}1&2\\4&5\end{pmatrix}\vec{y} = x_1y_1+2x_1y_2+4x_2y_1+5x_2y_2\\ b_2(\vec{x},\vec{y}) &= \vec{x}^T\begin{pmatrix}1&3\\3&5\end{pmatrix}\vec{y} = x_1y_1+3x_1y_2+3x_2y_1+5x_2y_2. \end{aligned}$$

Dann ist $Q(\vec{x}) = b_1(\vec{x},\vec{x}) = b_2(\vec{x},\vec{x}) = x_1^2+6x_1x_2+5x_2^2$.

Die Eigenschaften der positiven bzw. negativen Definitheit bzw. Semidefinitheit für Bilinearformen (Definition 6.1.9) übertragen sich nun auf quadratische Formen. Eine symmetrische Bilinearform besitzt genau dann eine dieser Eigenschaften, wenn sie auch die zugehörige quadratische Form besitzt.

Definition 7.4.2. (Quadriken)
Es sei $(V,\langle\cdot,\cdot\rangle)$ ein Euklidischer Vektorraum mit $\dim V=n$, weiter $\mathcal{B}$ eine ONB von V, $\mathcal{A}\in\mathbb{R}^{(n,n)}$ symmetrisch, $\vec{b}\in\mathbb{R}^n$ und $c\in\mathbb{R}$. Unter einer Hyperfläche 2. Ordnung (oder Quadrik) versteht man die Menge aller $\vec{v}\in V$, deren Koordinatenvektoren bzgl. $\mathcal{B}$ die Gleichung

$$\vec{x}^T\mathcal{A}\vec{x}+2\vec{b}^T\vec{x}+c=0 \tag{$*$}$$

erfüllen.

Beispiel 7.4.2. Es sei $V=\mathbb{R}^2$, $a\geq b>0$, $c=-1$, $\vec{b}=\vec{0}$ und

$$\mathcal{A}=\begin{pmatrix}\frac{1}{a^2}&0\\0&\frac{1}{b^2}\end{pmatrix}.$$

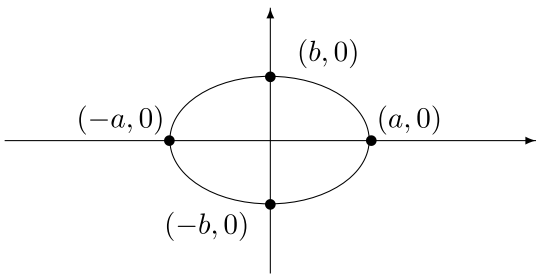
Setzen wir $\vec{x}=(x,y)^T$, so nimmt die Gleichung $(*)$ die Form

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \tag{$**$}$$

an. Die Lösungsmenge von $(**)$

$$E=\left\{\begin{pmatrix}x\\y\end{pmatrix}\in\mathbb{R}^2\colon \frac{x^2}{a^2}+\frac{y^2}{b^2}=1\right\}$$

wird auch als Ellipse mit großer Halbachse a, kleiner Halbachse b und Mittelpunkt $(0,0)$ bezeichnet.

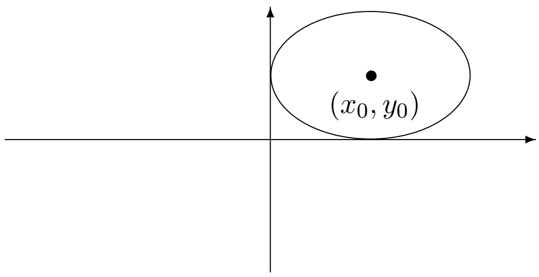


Große und kleine Halbachsen werden auch gemeinsam als Hauptachsen bezeichnet.
Der Kreis um $(0, 0)$ mit Radius r ergibt sich als Spezialfall für $a = b = r$.

Auch alle Mengen, die aus der Ellipse mit der Gleichung $(**)$ durch eine Bewegung hervorgehen, werden als Ellipsen bezeichnet.
Besonders einfache Bewegungen sind Translationen: $T: \vec{x} \to T(\vec{x}) = \vec{x} + \vec{x}_0$. Es sei $\vec{x}_0 = (x_0, y_0)^T$. Es hat $T(E)$ die Gleichung

$$\frac{(x - x_0)^2}{a^2} + \frac{(y - y_0)^2}{b^2} = 1, \qquad (* * *)$$

womit $T(E)$ eine Ellipse mit Hauptachsen a und b und Mittelpunkt (x_0, y_0) ist.



Es ist nun möglich, die Gleichung jeder Quadrik durch eine Hauptachsentransformation in eine zur Darstellung $(* * *)$ ähnliche Form zu bringen. Wir geben zunächst ein Beispiel:

Beispiel 7.4.3. Wir betrachten die Quadrik Q im $\mathbb{R}^2$ mit der Gleichung

$$13x^2 - 10xy + 13y^2 + 2\sqrt{2}x - 34\sqrt{2}y - 22 = 0.$$

Man erhält diese als Spezialfall von $\vec{x}^T \mathcal{A}\vec{x} + 2\vec{b}^T\vec{x} + c = 0$, indem man

$$\mathcal{A} = \begin{pmatrix} 13 & -5 \\ -5 & 13 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} \sqrt{2} \\ -17\sqrt{2} \end{pmatrix}$$

und $c = -22$ setzt. Nach Satz 7.3.6 (Spektralsatz im reellen Fall) gibt es eine orthogonale Matrix $\mathcal{P}$, so dass $\mathcal{P}^T \mathcal{A}\mathcal{P}$ eine Diagonalmatrix ist. Die Spaltenvektoren von $\mathcal{P}$ sind dabei gerade die Eigenvektoren von $\mathcal{A}$. Wir beginnen mit der Bestimmung der Eigenwerte: das charakteristische Polynom ist

$$P_{\mathcal{A}}(\lambda) = \det \begin{pmatrix} 13 - \lambda & -5 \\ -5 & 13 - \lambda \end{pmatrix} = (\lambda - 13)^2 - 25,$$

womit $P_{\mathcal{A}}(\lambda)$ die Nullstellen $\lambda_1 = 8$ und $\lambda_2 = 18$ besitzt. Wir bestimmen die Eigenräume:
$\lambda_1 = 8$:

$$\begin{pmatrix} 5 & -5 \\ -5 & 5 \end{pmatrix} \vec{x} = \vec{0} \Rightarrow U_8 = \left\{ \begin{pmatrix} t \\ t \end{pmatrix} : t \in \mathbb{R} \right\},$$

$\lambda_2 = 18$:

$$\begin{pmatrix} -5 & -5 \\ -5 & -5 \end{pmatrix} \vec{x} = \vec{0} \Rightarrow U_{18} = \left\{ \begin{pmatrix} t \\ -t \end{pmatrix} : t \in \mathbb{R} \right\}.$$

Damit kann

$$\mathcal{P} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

gewählt werden, und es gilt

$$\mathcal{P}^T \mathcal{A}\mathcal{P} = \begin{pmatrix} 8 & 0 \\ 0 & 18 \end{pmatrix}.$$

Der Koordinatenvektor $\vec{x}'$ von $\vec{x}$ bzgl. der ONB

$$\mathcal{B} = \left\{ \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \right\}$$

erfüllt nach Satz 3.7.3 die Bedingung $\vec{x} = \mathcal{P}\vec{x}'$. Die Koordinaten der Punkte der Quadrik Q bzgl. der Basis $\mathcal{B}'$ erfüllen also die Gleichung

$$\vec{x}'^T \begin{pmatrix} 8 & 0 \\ 0 & 18 \end{pmatrix} \vec{x}' + 2\vec{b}^T \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \vec{x}' - 22 = 0$$

oder $8x'^2 + 18y'^2 - 32x' - 36y' - 22 = 0$ bzw.

$$\frac{(x'-2)^2}{9} + \frac{(y'-1)^2}{4} = 1.$$

Somit ist Q eine <u>Ellipse</u> mit großer Halbachse 3 und kleiner Halbachse 2. Die Geraden, auf denen die Hauptachsen liegen, werden von den Eigenvektoren

$$\vec{v}_8 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{und} \quad \vec{v}_{18} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

dargestellt. Die Koordinaten des Mittelpunkts bzgl. der Basis $\mathcal{B}'$ sind $\vec{x}'_m = (2,1)^T$ und bzgl. der Basis $\mathcal{B} = \{\vec{e}_1, \vec{e}_2\}$

$$\vec{x}_m = \mathcal{P}\vec{x}'_m = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{3}{\sqrt{2}} \end{pmatrix}.$$

Da

$$\mathcal{P} = \begin{pmatrix} \cos\frac{\pi}{4} & -\sin\frac{\pi}{4} \\ \sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{pmatrix}$$

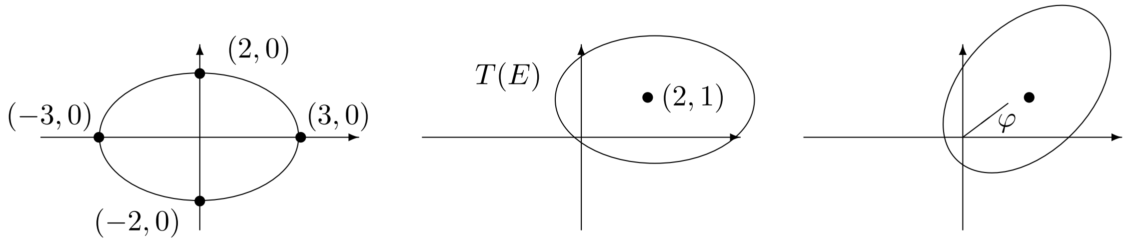
die Drehmatrix zum Winkel $\frac{\pi}{4}$ ist, geht Q aus

$$E = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \colon \frac{x^2}{3^2} + \frac{y^2}{2^2} = 1 \right\}$$

durch die Bewegung $\varphi \circ T$ mit der Translation

$$T \colon \vec{x} \to T(\vec{x}) = \vec{x} + \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

und der Drehung φ und den Winkel $\frac{\pi}{4}$ hervor.



Satz 7.4.2. *(Hauptachsentransformation)*
Es sei $(V, \langle\cdot,\cdot\rangle)$ ein endlichdimensionaler Vektorraum. Zu jeder Quadrik $Q \subset V$ gibt es eine ONB $\mathcal{B}$ von V, so dass die Koordinaten der Punkte von Q bzgl. $\mathcal{B}$ durch eine der Gleichungen folgenden Typs (Normalform) beschrieben werden:

$$\frac{(x_1 - x_1^{(0)})^2}{c_1^2} + \ldots + \frac{(x_d - x_d^{(0)})^2}{c_d^2} - \frac{(x_{d+1} - x_{d+1}^{(0)})^2}{c_{d+1}^2} - \ldots - \frac{(x_r - x_r^{(0)})^2}{c_r^2} = 1 \qquad \text{(I)}$$

$$\frac{(x_1 - x_1^{(0)})^2}{c_1^2} + \ldots + \frac{(x_d - x_d^{(0)})^2}{c_d^2} - \frac{(x_{d+1} - x_{d+1}^{(0)})^2}{c_{d+1}^2} - \ldots - \frac{(x_r - x_r^{(0)})^2}{c_r^2} = -(x_{r+1} - x_{r+1}^{(0)}) \qquad \text{(II)}$$

$$\frac{(x_1 - x_1^{(0)})^2}{c_1^2} + \ldots + \frac{(x_d - x_d^{(0)})^2}{c_d^2} - \frac{(x_{d+1} - x_{d+1}^{(0)})^2}{c_{d+1}^2} - \ldots - \frac{(x_r - x_r^{(0)})^2}{c_r^2} = 0 \qquad \text{(III)}$$

mit $r - d \leq d$. Dabei ist außerdem $1 \leq r \leq n$ bzw. $r \leq n-1$ im Fall (II) und $0 \leq d \leq r$.

Beweis. Es sei $\mathcal{B}_0$ eine ONB von V bzgl. der Q durch die Gleichung

$$\vec{x}^T \mathcal{A} \vec{x} + 2\vec{b}^T \vec{x} + c = 0 \qquad (1)$$

gegeben ist. Nach dem Spektralsatz im reellen Fall (Satz 7.3.6) gibt es eine ONB $\mathcal{C}$ von V, die aus Eigenvektoren von $\mathcal{A}$ besteht. Es sei $\mathcal{C} = \{\vec{c}_1, \ldots, \vec{c}_n\} = \varphi(\mathcal{B}_0)$ mit $\mathcal{M}(\varphi; \mathcal{B}_0, \mathcal{B}_0) = \mathcal{P}$. Dann gilt

$$\mathcal{M}(\varphi; \mathcal{C}, \mathcal{C}) = \mathcal{P}^T \mathcal{A} \mathcal{P} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

wobei λ_i die Eigenwerte von $\mathcal{A}$ sind. Dann wird (1) zu

$$\lambda_1 x_1'^2 + \ldots + \lambda_n x_n'^2 + 2b_1' x_1' + \ldots + 2b_n' x_n' + c' = 0 \qquad (2)$$

mit $b_i' \in \mathbb{R}$ und $c' \in \mathbb{R}$. Die Eigenvektoren in $\mathcal{C}$ seien so umgeordnet, dass $\lambda_1, \ldots, \lambda_r \neq 0$ und $\lambda_{r+1} = \ldots = \lambda_n = 0$ für ein r mit $1 \leq r \leq n$ gelte. Es ist dann $r = \operatorname{rg} \mathcal{A} > 0$. Mit quadratischer Ergänzung können wir (2) zu

$$\lambda_1 \cdot \left(x_1' + \frac{b_1}{\lambda_1}\right)^2 + \ldots + \lambda_r \cdot \left(x_r' + \frac{b_r}{\lambda_r}\right)^2 + 2b_{r+1}' x_{r+1}' + \ldots + 2b_n' x_n' + c'' = 0 \qquad (3)$$

mit $c'' \in \mathbb{R}$ umgeformt werden. Es sei nun $\vec{v} := 2b_{r+1}' \vec{c}_{r+1}' + \ldots + 2b_n \vec{c}_n'$. Dann betrachten wir die zwei Fälle

- $\vec{v} = \vec{0}$:
  Ist $c'' \neq 0$, so kann (3) durch Durchmultiplizieren mit einer geeigneten Konstanten und einer ggfs. notwendigen Umnummerierung der Variablen in die Form (I) gebracht werden, wobei die Striche weggelassen werden. Ist $c'' = 0$, so kann (3) in die Form (III) gebracht werden.

- $\vec{v} \neq \vec{0}$:
  Wegen $\langle v, \vec{c}_j \rangle = 0$ für $1 \leq j \leq r$ kann $\{\vec{c}_1, \ldots, \vec{c}_r\}$ zu einer ONB $\mathcal{B}' = \{\vec{c}_1, \ldots, \vec{c}_r, \vec{c}_{r+1}', \ldots, \vec{c}_n'\}$ von V ergänzt werden, wobei $a\vec{c}_{r+1}' = 2b_{r+1}' \vec{c}_{r+1} + \ldots + 2b_n' \vec{c}_n$ für ein $a \in \mathbb{R}$ ist. Geeignetes Durchmultiplizieren und Umnummerieren der Variablen ergibt dann die Form (II).

□

Beispiel 7.4.4. Für den Fall $n = 2$ erhalten wir die folgenden Klassifikationen der Kurven 2. Ordnung, wobei $x_i^{(0)} = 0$ für alle $1 \leq i \leq n$ sei.

Typ	(r,d)	Gleichung	Bezeichnung
I	$(1,0)$	$-\frac{x_1^2}{c_1^2} = 1$	leere Menge
I	$(1,1)$	$\frac{x_1^2}{c_1^2} = 1$	zwei parallele Geraden
I	$(2,0)$	$-\frac{x_1^2}{c_1^2} - \frac{x_2^2}{c_2^2} = 1$	leere Menge
I	$(2,1)$	$\frac{x_1^2}{c_1^2} - \frac{x_2^2}{c_2^2} = 1$	Hyperbel
I	$(2,2)$	$\frac{x_1^2}{c_1^2} + \frac{x_2^2}{c_2^2} = 1$	Ellipse (Kreis für $c_1 = c_2$)
II	$(1,1)$	$\frac{x_1^2}{c_1^2} = -x_2$	Parabel
III	$(1,1)$	$\frac{x_1^2}{c_1^2} = 0$	Gerade
III	$(2,1)$	$\frac{x_1^2}{c_1^2} - \frac{x_2^2}{c_2^2} = 0$	zwei sich schneidende Geraden
III	$(2,2)$	$\frac{x_1^2}{c_1^2} + \frac{x_2^2}{c_2^2} = 0$	ein Punkt

Beispiel 7.4.5. Auch für $n = 3$ gibt es eine Klassifikation der Quadriken (Flächen 2. Ordnung). Wir geben sie hier nicht vollständig an, sondern beschränken uns auf einige Beispiele:

Typ	(r,d)	Gleichung	Bezeichnung
I	$(2,2)$	$\frac{x_1^2}{c_1^2} + \frac{x_2^2}{c_2^2} = 1$	elliptischer Zylinder (kreisförmig falls $c_1 = c_2$)
I	$(3,3)$	$\frac{x_1^2}{c_1^2} + \frac{x_2^2}{c_2^2} + \frac{x_3^2}{c_3^2} = 1$	Ellipsoid mit Halbachsen c_1, c_2 und c_3 (Sphäre für $c_1 = c_2 = c_3$)
I	$(3,2)$	$\frac{x_1^2}{c_1^2} + \frac{x_2^2}{c_2^2} - \frac{x_3^2}{c_3^2} = 1$	einschaliges Hyperboloid
I	$(3,1)$	$\frac{x_1^2}{c_1^2} - \frac{x_2^2}{c_2^2} - \frac{x_3^2}{c_3^2} = 1$	zweischaliges Hyperbolid
III	$(3,2)$	$\frac{x_1^2}{c_1^2} + \frac{x_2^2}{c_2^2} - \frac{x_3^2}{c_3^2} = 0$	Doppelkegel

Wir suchen nun ein Kriterium, das erlaubt, eine Quadrik, deren Gleichung in allgemeiner Form

$$\vec{x}^T \mathcal{A} \vec{x} + 2\vec{b}^T \vec{x} + c = 0 \tag{$*$}$$

gegeben ist, mit wenig Rechnung- wenigstens grob- zu klassifizieren.
Im Fall $n > 2$ kann man zum Beispiel fragen: Stellt $(*)$ eine Ellipse dar? Die Antwort ist sicherlich nein, wenn die durch $\mathcal{A}$ gegebene Form $Q(\vec{x}) = \vec{x}^T \mathcal{A} \vec{x}$ nicht positiv definit ist.

Definition 7.4.3. Es sei K ein Körper, $n \in \mathbb{N}$, $\mathcal{A} \in K^{(n,n)}$ mit $\mathcal{A} = (a_{ij})_{1 \leq i,j \leq n}$ und $1 \leq k \leq n$. Unter dem k- ten Hauptminor von $\mathcal{A}$ versteht man $\det \mathcal{A}_k$ mit der Teilmatrix $\mathcal{A}_k := (a_{ij})_{1 \leq i,j \leq k}$.

Satz 7.4.3. *Es sei $\mathcal{A} \in \mathbb{R}^{(n,n)}$ symmetrisch.*

i) Die Matrix $\mathcal{A}$ ist genau dann positiv definit (bzw. positiv semidefinit), wenn alle Eigenwerte von $\mathcal{A}$ positiv (bzw. nichtnegativ) sind.

ii) Die Matrix $\mathcal{A}$ ist genau dann negativ definit (bzw. negativ semidefinit), wenn alle Eigenwerte von $\mathcal{A}$ negativ (bzw. nichtpositiv) sind.

Beweis. Es sei $b(\vec{x}, \vec{y}) = \vec{x}^T \mathcal{A} \vec{y}$ mit zugehöriger quadratischer Form $Q(\vec{x}) = \vec{x}^T \mathcal{A} \vec{x}$. Nach Satz 7.3.6 (Spektralsatz im reellen Fall) gibt es ein $\mathcal{P} \in O(n)$, so dass

$$\mathcal{P}^T \mathcal{A} \mathcal{P} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = D$$

mit den Eigenwerten λ_i von $\mathcal{A}$ gilt. Nach Satz 6.1.2 ist D die Matrix der Bilinearform b bzgl. der Basis $\mathcal{B} = \{\mathcal{P}\vec{e}_1, \ldots, \mathcal{P}\vec{e}_n\}$. Sind $(x'_1, \ldots, x'_n)$ die Koordinaten von $\vec{x}$ bzgl. $\mathcal{B}$, so ist $Q(\vec{x}) = \lambda_1 x'^2_1 + \ldots + \lambda_n x'^2_n$. Daraus folgt die Behauptung. □

Während die Anwendung von Satz 7.4.3 die Berechnung der Eigenwerte erfordert, ist dies bei der Anwendung des nächsten Kriteriums nicht notwendig.

Satz 7.4.4. *Es sei $\mathcal{A} \in \mathbb{R}^{(n,n)}$ symmetrisch.*

i) Die Matrix $\mathcal{A}$ ist genau dann positiv definit, wenn alle Hauptminoren von $\mathcal{A}$ positiv sind.

ii) Die Matrix $\mathcal{A}$ ist genau dann negativ definit, wenn die k- ten Hauptminoren von $\mathcal{A}$ für gerades k positiv und für ungerades k negativ sind.

Beweis. i) "$\Rightarrow$":

Es sei $\mathcal{A}$ positiv definit. Weiter sei $\mathcal{A} = (a_{ij})_{1 \le i,j \le n}$ mit den Teilmatrizen $\mathcal{A}_k = (a_{ij})_{1 \le i,j \le k}$. Die zu $\mathcal{A}_k$ gehörende Bilinearform $b_k \colon \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}$, $(\vec{x}, \vec{y}) \to b_k(\vec{x}, \vec{y}) = \vec{x}^T \mathcal{A} \vec{y}$ nimmt dieselben Werte an wie die Einschränkung der Biliearform $b = b_n \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $(\vec{x}, \vec{y}) \to b(\vec{x}, \vec{y}) = \vec{x}^T \mathcal{A} \vec{y}$ auf $\mathbb{R}^{(k)} := \{\vec{x} \in \mathbb{R}^n \colon \vec{x} = (x_1, \ldots, x_k, 0, \ldots, 0)^T\}$, und ist daher ebenfalls positiv definit. Nach Satz 7.3.6 (Spektralsatz im reellen Fall) gibt es $\mathcal{P}_k \in O(k)$, so dass

$$\mathcal{P}_k^T \mathcal{A}_k \mathcal{P}_k = \begin{pmatrix} \lambda_1^{(k)} & & 0 \\ & \ddots & \\ 0 & & \lambda_n^{(k)} \end{pmatrix} = \mathcal{D}_k$$

mit den Eigenwerten $\lambda_i^{(k)}$ von $\mathcal{D}_k$. Dann ist $\det \mathcal{A}_k = \det \mathcal{D}_k > 0$ nach Satz 7.4.3.

"$\Leftarrow$":

Es seien alle Hauptminoren von $\mathcal{A}$ positiv. Wir zeigen durch Induktion nach n, dass $\mathcal{A}$ positiv definit ist.

Induktionsanfang: $n = 1$:

Es gilt $\mathcal{A} = (a_{11})$ mit $a_{11} > 0$ und $\vec{x} = (x)$. Dann ist $\vec{x}^T \mathcal{A} \vec{x} = a_{11} x^2 > 0$, also ist $\mathcal{A}$ positiv definit.

Induktionsschritt: $n - 1 \to n$:

Wir betrachten die nichtquadratische Teilmatrix

$$\mathcal{A}' = (a_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le n-1}} = \begin{pmatrix} a_{11} & \ldots & a_{1,n-1} \\ \vdots & & \vdots \\ a_{n-1,1} & \ldots & a_{n-1,n-1} \\ a_{n1} & \ldots & a_{n,n-1} \end{pmatrix}.$$

Die Matrix $\mathcal{A}'$ enthält $\mathcal{A}_{n-1}$ als Teilmatrix. Es gilt nach Voraussetzung $\det \mathcal{A}_{n-1} > 0$, also $\operatorname{rg} \mathcal{A}_{n-1} = n - 1$ und somit auch $\operatorname{rg} \mathcal{A}' = n - 1$. Die n- te Zeile $\vec{z}_n = (a_{n,1}, \ldots, a_{n,n-1})$ ist daher eine Linearkombination der Zeilen $\vec{z}_j = (a_{j,1}, \ldots, a_{j,n-1})$ mit $1 \le j \le n - 1$, also

$$\vec{z}_n = \sum_{j=1}^{n-1} \mu_j \vec{z}_j \qquad (*)$$

mit $\mu_j \in \mathbb{R}$. Subtraktion des μ_j- fachen der j- ten Zeile von der $n-1$- ten Zeile von $\mathcal{A}'$ kann durch Multiplikation von links mit der "Elementarmatrix"

$$\mathcal{L}_j = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & \ddots & \\ 0 & \dots & -\mu_j & 1 \end{pmatrix},$$

wobei das $-\mu_j$ in der j- ten Spalte steht, erreicht werden. Dazu gilt $\det \mathcal{L}_j = 1$. Aus $(*)$ erhalten wir

$$\mathcal{L}_1 \cdots \mathcal{L}_{n-1}\mathcal{A} = \left(\begin{array}{ccc|c} & & & a_{1n} \\ & \mathcal{A}_{n-1} & & \vdots \\ & & & a_{n-1,n} \\ \hline 0 & \dots & 0 & c \end{array}\right)$$

mit $c \in \mathbb{R}$. Multiplikation von rechts mit $\mathcal{L}_j^T$ hat den entsprechenden Effekt auf die Spalten. Es ist also mit $\mathcal{R} = \mathcal{L}_1 \cdots \mathcal{L}_{n-1}$

$$\mathcal{R}\mathcal{A}\mathcal{R}^T = \left(\begin{array}{ccc|c} & & & 0 \\ & \mathcal{A}_{n-1} & & \vdots \\ & & & 0 \\ \hline 0 & \dots & 0 & c \end{array}\right) = \tilde{\mathcal{A}}.$$

Wegen $\det \tilde{\mathcal{A}} = \det \mathcal{A} \cdot (\det \mathcal{R})^2$ ist $\det \tilde{\mathcal{A}} > 0$. Nach Voraussetzung ist $\det \mathcal{A}_{n-1} > 0$, also $c > 0$. Nach Satz 6.1.2 ist $\tilde{\mathcal{A}}$ die zur Bilinearform b mit $b(\vec{x}, \vec{y}) = \vec{x}^T \mathcal{A} \vec{y}$ gehörige Matrix bezüglich der Basis $\{\mathcal{R}^T \vec{e}_1, \dots, \mathcal{R}^T \vec{e}_n\}$. Damit ist $\tilde{\mathcal{A}}$ genau dann positiv definit, wenn $\mathcal{A}$ positiv definit ist. Nun ist für $\vec{x} = (x_1, \dots, x_n)^T$ und $\vec{x}_{n-1} = (x_1, \dots, x_{n-1})^T$

$$\vec{x}^T \mathcal{A} \vec{x} = \vec{x}_{n-1}^T \mathcal{A}_{n-1} \vec{x}_{n-1} + c x_n^2 > 0$$

für alle $\vec{x} \in \mathbb{R}^n \backslash \{\vec{0}\}$, da nach Induktionshypothese $\mathcal{A}_{n-1}$ positiv definit ist.

ii) Die Matrix $\mathcal{A}$ ist genau dann negativ definit, wenn $-\mathcal{A}$ positiv definit ist. Daraus folgt die Behauptung.

□

Beispiel 7.4.6. Die Lösungsmenge der Gleichung

$$Q(\vec{x}) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2b_1 x + 2b_2 y + c = 0$$

ist keine Ellipse, wenn $a_{11} > 0$ und $a_{11}a_{22} - a_{12}^2 \leq 0$ ist.

Kapitel 8

Matrixpolynome und Jordansche Normalform

8.1 Polynome

Es sei K stets ein Körper.

Definition 8.1.1. Der Polynomring $(K[x], +, \cdot)$, kurz $K[x]$ ist die Menge aller Ausdrücke

$$f = f(x) = a_0 + a_1 x + \ldots + a_n x^n = \sum_{j=0}^{n} a_j x^j$$

mit $a_i \in K$. ist $a_n \neq 0$, so heißt n der Grad von f (Schreibweise: $n = \deg(f)$. So wird $K[x]$ durch die Addition und Multiplikation von Polynomen zu einem Ring. Es sei $f(x) = a_0 + \ldots + a_n x^n$ und $g(x) = b_0 + \ldots + b_m x^m$.

i) $(f+g)(x) = (a_0 + b_0) + (a_1 + b_1)x + \ldots + (a_k + b_k)x^k$ mit $k = \max\{m, n\}$. Vereinbarung dazu ist $a_l = 0$ und $b_l = 0$ für $l > k$.

ii) $(f \cdot g)(x) = c_0 + c_1 x + \ldots + c_{n+m} x^{n+m}$ mit $c_j = \sum_{i=0}^{j} a_i b_{j-i}$.

Das Polynom mit $a_i = 0$ für alle i heißt das Nullpolynom und wird mit 0 bezeichnet. Polynome vom Grad 0 heißen konstante Polynome oder Konstanten.

Folgenden einfachen Satz geben wir ohne Beweis an:

Satz 8.1.1. *Es gilt*

i) $\deg(f+g) \leq \max\{\deg(f), \deg(g)\}$

ii) $\deg(f \cdot g) = \deg(f) + \deg(g)$, *falls* $f, g \neq 0$.

Beweis. ohne Beweis. □

Definition 8.1.2. Es seien $f, g \in K[x]$. Es heißt f Teiler von g (oder f teilt g, Schreibweise: $f|g$), wenn es ein Polynom h mit $g = fh$ gibt. Weiter heißt f echter Teiler von g, wenn $0 < \deg(f) < \deg(g)$ gilt. Das Polynom g heißt irreduzibel oder prim, wenn g keine echten Teiler besitzt.

Wir listen ohne Beweis einige einfache Eigenschaften aus:

Satz 8.1.2. *Es seien $f, g \in K[x]$.*

i) Alle linearen Polynome $g = a_0 + a_1 x$ mit $a_1 \neq 0$ sind irreduzibel.

ii) Aus $fh_1 = fh_2$ und $f \neq 0$ folgt $h_1 = h_2$ (Kürzungsregel).

iii) Aus $f|g$ und $g|h$ folgt $f|h$.

iv) Aus $f|g$ und $f|h$ folgt $f|(g+h)$.

v) Aus $f|g$ und $g|f$ folgt die Existenz eines $c \in K - \{0\}$ mit $g = c \cdot f$.

Beweis. ohne Beweis. □

Satz 8.1.3. *(Division mit Rest)*
Sind $f, g \in K[x]$ und $g \neq 0$, so gibt es $q, r \in K[x]$ mit

$$f = q \cdot g + r \tag{$*$}$$

mit $r = 0$ oder $\deg(r) < \deg(g)$.

Beweis. Für $f = 0$ ist $(*)$ mit $q = r = 0$ erfüllt.
Der Beweis der anderen Fälle geschieht durch Induktion nach $n = \deg(f)$.
Induktionsanfang: $n = 0$:
Für $f = a_0 \neq 0$ können wir $q = 0$ und $r = 0$ wählen, falls $\deg(g) > 0$ ist und $q = a_0 q^{-1}$ und $r = 0$, falls $\deg(g) = 0$.
Induktionsschritt: $n \to n+1$:
Es sei $f = a_0 + \ldots + a_n x^n$ und $g = b_0 + \ldots b_m x^m$ mit $a_n \neq 0$ und $b_m \neq 0$. Für $n < m$ können wir $q = 0$ und $r = f$ wählen. Für $n \geq m$ sei $f_1 = f - a_n b_m^{-1} x^{n-m} g$. Dann ist $\deg(f_1) < \deg(f)$. Nach Induktionshypothese gibt es $q_1, r \in K[x]$ mit $f_1 = q_1 g + r$ und $r = 0$ oder $\deg(r) < \deg(g)$. Dann ist $f = q \cdot g + r$ mit $q = (q_1 + a_n b_m^{-1} x^{n-m}) g$. □

Satz 8.1.4. *Es sei $0 \neq f \in K[x]$ und $x_0 \in K$ eine Nullstelle von f, d.h. $f(x_0) = 0$. Dann existiert ein $q \in K[x]$ mit $f(x) = (x - x_0) \cdot q(x)$.*

Beweis. Satz 8.1.3 (Division mit Rest) gibt $f(x) = (x-a) \cdot q(x) + r(x)$ mit $\deg(r) < 1$, also $r(x) = r$. Einsetzen von $x = x_0$ liefert $r = 0$. □

Satz 8.1.5. *Es sei $f \in K[x]$ und $\deg(f) = n$. Dann hat f höchstens n Nullstellen.*

Beweis. Wenn $x_1, \ldots, x_m$ verschiedene Nullstellen von f sind, so folgt aus Satz 8.1.4, dass auch $f(x) = (x - x_1) \cdots (x - x_m) \cdot q(x)$ für $q \in K[x]$ gilt. Also ist $m \leq m + \deg(q) = \deg(f) = n$. □

Satz 8.1.6. *Es sei $f \in K[x]$ und $\deg(f) \in \{2, 3\}$. Das Polynom f ist genau dann irreduzibel, wenn es keine Nullstellen hat.*

Beweis. "$\Rightarrow$":
Es sei f irreduzibel. Dann kann f keine Nullstellen besitzen, da ansonsten nach Satz 8.1.5 wiederum $f(x) = (x - x_0) \cdot q(x)$ mit $0 < \deg(q) < \deg(f)$ gelten würde.
"$\Leftarrow$":
Das Polynom f habe keine Nullstellen. Wir wollen annehmen, es gebe eine Zerlegung $f = g \cdot h$ mit $0 < \deg(g)$ und $0 < \deg(h)$. Daraus folgt $\deg(g) = 1$ oder $\deg(h) = 1$. Dann hat g oder h und somit auch f eine Nullstelle, ein Widerspruch. □

Definition 8.1.3. (größter gemeinsamer Teiler)
Es seien $f, g \in K[x]$. Das Polynom $d \in K[x]$ heißt gemeinsamer Teiler von f und g, wenn $d|f$ und $d|g$. Es heißt echter gemeinsamer Teiler, wenn zusätzlich $\deg(d) > 0$ gilt.
Die Polynome f und g heißen teilerfremd, wenn sie keine echten gemeinsamen Teiler besitzen.
Das Polynom $d \in K[x]$ heißt größter gemeinsamer Teiler von f und g (Schreibweise: $d = ggT(f, g)$), wenn aus $h|f$ und $h|g$ auch $h|d$ folgt.

Satz 8.1.7. *Es seien $f, g \in K[x] - \{0\}$. Dann existiert $d = ggT(f, g)$ und d ist ein gemeinsamer Teiler von f und g. Sind d_1 und d_2 größte gemeinsame Teiler von f und g, so gibt es $c \in K - \{0\}$ mit $d_2 = cd_1$. Weiter gibt es $h_1, h_2 \in K[x]$ mit $d = h_1 f + h_2 g$.*

Beweis. Es sei $L = \{q_1 f + q_2 g \colon q_1, q_2 \in K[x]\}$ und $d = h_1 f + h_2 g$ mit $h_i \in K[x]$ sei ein Polynom kleinsten Grades aus $L - \{0\}$. Es gelte $h|f$ und $h|g$. Nach Satz 8.1.2 (iv) folgt $h|(h_1 f + h_2 g) = d$. Also gilt $d = ggT(f, g)$.
Ist $\tilde{d} = ggT(f, g)$, so gilt $\tilde{d}|d$ und $d|\tilde{d}$, also nach Satz 8.1.2 (v) auch $\tilde{d} = cd$ mit $c \in K - \{0\}$. Nach Satz 8.1.3 gibt es $q, r \in K[x]$ mit $f = qd + r$ und $\deg(r) < \deg(d)$ oder $r = 0$. Es ist $r = f - qd = (1 - qh_1)f - qh_2 g \in L$. Da d ein Polynom kleinsten Grades aus L ist, folgt $r = 0$, also $d|f$. Ebenso folgt $d|g$, womit d ein gemeinsamer Teiler von f und g ist. □

Satz 8.1.8. *Es seien $f, g, h \in K[x]$ mit $f \neq 0$ und $g \neq 0$. Dann ist $ggT(f, g) = ggT(f, g + hf)$.*

Beweis. Mit $h|f$ und $h|g$ ist $h|f$ und $h|(g+hf)$ äquivalent. Nach Definition 8.1.3 folgt die Behauptung. □

Der größte gemeinsame Teiler kann mittels des Euklidischen Algorithmus bestimmt werden.

Satz 8.1.9. *(Euklidischer Algorithmus)*
Es seien $f, g \in K[x] - \{0\}$. Die Folgen (q_i) und (r_i) von Polynomen $q_i, r_i \in K[x]$ seien durch

$$\begin{aligned}
f &= q_1 \cdot g + r_1, \quad \deg(r_1) < \deg(g),\ r_1 \neq 0 \\
g &= q_2 \cdot r_1 + r_2, \quad \deg(r_2) < \deg(r_1),\ r_2 \neq 0 \\
&\vdots \\
r_{j-2} &= q_j \cdot r_{j-1} + r_j, \quad \deg(r_j) < \deg(r_{j-1}) \\
r_{j-1} &= q_{j+1} \cdot r_j
\end{aligned}$$

definiert. Dann ist $r_j = ggT(f, g)$. Eine Darstellung $r_j = h_1 f + h_2 g$ mit $h_1, h_2 \in K[x]$ kann aus den obigen Gleichungen durch Rückwärtseinsetzen gefunden werden.

Beweis. Nach Satz 8.1.8 folgt aus $f = q_1 g + r_1$

$$ggT(f, g) = ggT(g, f - q_1 g) = ggT(g, r_1).$$

Weiter folgt aus $g = q_2 r_1 + r_2$

$$ggT(f, g) = ggT(g, r_1) = ggT(r_1, g - q_2 r_1) = ggT(r_1, r_2).$$

So fortfahrend erhält man

$$ggT(f, g) = ggT(r_{j-1}, r_j) = r_j.$$

□

Beispiel 8.1.1. Es sei $K = \mathbb{R}$, $f(x) = x^5 + x^3 + 2x^2 + 2$ und $g(x) = x^4 + 2x^3 + 4x^2 + 2x + 3$. Bestimme $d = ggT(f, g)$ und drücke es in der Form $d = h_1 f + h_2 g$ mit $h_1, h_2 \in \mathbb{R}[x]$ aus.

Lösung:
1. Schritt: $f = q_1 g + r_1$ mit $\deg(r_1) < \deg(g)$ oder $r_1 = 0$:

$$
\begin{array}{rrrrrrcl}
(x^5 & & +x^3 & +2x^2 & & +2) & = & (x^4 + 2x^3 + 4x^2 + 2x + 3) \cdot (x - 2) + r_1 \\
-(x^5 & +2x^4 & +4x^3 & +2x^2 & +3x) & & & \\
\hline
& -2x^4 & -3x^3 & & -3x & +2 & & \\
& -(-2x^4 & -4x^3 & -8x^2 & -4x & -6) & & \\
\cline{2-6}
& & x^3 & +8x^2 & +x & +8 & &
\end{array}
$$

Also ist $q_1 = x - 2$ und $r_1 = x^3 + 8x^2 + x + 8$.

2. Schritt: $g = q_2 r_1 + r_2$ mit $\deg(r_2) < \deg(r_1)$ oder $r_2 = 0$:

$$
\begin{array}{rrrrrcl}
(x^4 & +2x^3 & +4x^2 & +2x & +3) & = & (x^3 + 8x^2 + x + 8) \cdot (x - 6) + r_2 \\
-(x^4 & +8x^3 & +x^2 & +8x) & & & \\
\hline
& -6x^3 & +3x^2 & -6x & +3 & & \\
& -(-6x^3 & -48x^2 & -6x & -48) & & \\
\cline{2-5}
& & 51x^2 & & +51 & &
\end{array}
$$

Somit ist $q_2 = x - 6$ und $r_2 = 51(x^2 + 1)$.

3. Schritt: $r_1 = q_3 r_2 + r_3$ mit $\deg(r_3) < \deg(r_2)$ oder $r_3 = 0$:

$$
\begin{array}{rrrrcl}
(x^3 & +8x^2 & +x & +8) & = & (x^2 + 1) \cdot (x + 8) + r_3 \\
-(x^3 & & +x) & & & \\
\cline{1-3}
& 8x^2 & & +8 & & \\
& -(8x^2 & & +8) & & \\
\cline{2-4}
& & & 0 & &
\end{array}
$$

Somit ist $q_3 = \frac{1}{51}(x + 8)$ und $r_3 = 0$.
Damit gilt $ggT(f, g) = d(x) = c \cdot (x^2 + 1)$ mit $c \in \mathbb{R} \backslash \{0\}$.
Die Darstellung $d = h_1 f + h_2 g$ folgt durch Rückwärtseinsetzen: Aus dem zweiten Schritt ergibt sich

$$51(x^2 + 1) = g - (x^3 + 8x^2 + x + 8) \cdot (x - 6)$$

und aus dem ersten Schritt

$$51(x^2 + 1) = g - (f - (x - 2) \cdot g) \cdot (x - 6) = (-x + 6) \cdot f + (x^2 - 8x + 13) \cdot g.$$

Damit ist

$$d = ggT(f, g) = x^2 + 1 = h_1 f + h_2 g$$

mit

$$h_1(x) = -\frac{1}{51}x + \frac{6}{51} \quad \text{und} \quad h_2 = \frac{1}{51}x^2 - \frac{8}{51}x + \frac{13}{51}.$$

Satz 8.1.10. *Es seien $f, g, h \in K[x]$. Weiter gelte $h|(fg)$ und h und f seien teilerfremd. Dann folgt $h|g$.*

Beweis. Nach Satz 8.1.9 gibt es $k_1, k_2 \in K[x]$ mit $k_1 h + k_2 f = 1$. Daraus folgt $h|(k_1 gh + k_2 fg) = g$. □

Satz 8.1.11. *(Eindeutige Zerlegung in Primfaktoren)*
Es sei $f \in K[x]$ und $\deg(f) \geq 1$. Dann gibt es irreduzible Polynome $p-1, \dots, p_m \in K[x]$, so dass

$$f = p_1 \cdots p_m \tag{$*$}$$

gilt. Die p_j sind bis auf die Reihenfolge und bis auf die Multiplikation mit Konstanten eindeutig bestimmt.

Beweis. Existenz:
Der Beweis wird durch Induktion nach $n = \deg(f)$ geführt.
Induktionsanfang $n = 1$:
Das Polynom f ist nach Satz 8.1.2 irreduzibel. Damit gilt $(*)$ mit $m = 1$ und $p_1 = f$.
Induktionsschritt $n \to n+1$:
Die Existenz sei schon für $\deg(f) < n$ beweisen, und es sei $\deg(f) = n$. Ist f irreduzibel, so gilt $(*)$ wieder mit $m = 1$ und $p_1 = f$. Andernfalls ist $f = gh$ mit $\deg(g) < n$ und $\deg(h) < n$. Nach Induktionshypothese ist $g = p_1 \cdots p_k$ und $h = q_1 \cdots q_l$ mit irreduziblen Polynomen p_i bzw. q_i. Also gilt $f = p_1 \cdots p_k \cdot q_1 \cdots q_l$.

Eindeutigkeit:
Wir führen den Beweis wieder durch Induktion nach $n = \deg(f)$. Induktionsanfang $n = 1$:
In $(*)$ ist $m = 1$ und $p_1 = c \cdot f$ mit $c \in K - \{0\}$.
Induktionsschritt $n \to n+1$:
Die Existenz sei schon für $\deg(f) < n$ beweisen, und es sei $\deg(f) = n$. Es sei $f = p_1 \cdots p_k = q_1 \cdots q_l$ mit irreduziblen Polynomen $p_i, q_i \in K[x]$.
Es gilt $p_1 | (q_1 \cdots q_l)$. Nach Satz 8.1.10 muss dann $p_1 | q_i$ für mindestens ein i mit $1 \leq i \leq l$ gelten. Die Nummerierung sei so, dass $p_1 | q_1$ gelte. da q_1 irreduzibel ist, folgt $q_1 = c \cdot p_1$ mit $c \in K - \{0\}$. Also gilt $p_1 \cdot p_k = c \cdot q_1 \cdots q_l$.
Die Behauptung folgt durch Anwenden der Induktionshypothese auf das Polynom $h = p_2 \cdots p_k$. □

Definition 8.1.4. (kleinestes gemeinsames Vielfaches)
Es seien $f, g \in K[x]$.

i) Das Polynom $l \in K[x]$ heißt gemeinsames Vielfaches von f und g, wenn $f|l$ und $g|l$.

ii) Das Polynom $l \in K[x]$ heißt kleinstes gemeinsames Vielfaches von f und g (Schreibweise: $l = kgV(f, g)$), falls für jedes gemeinsame Vielfache h von f und g auch $l|h$ gilt.

Satz 8.1.12. *Es seien $f, g \in K[x]$ und $f = p_1^{\alpha_1} \cdots p_m^{\alpha_m}$ sowie $g = p_1^{\beta_1} \cdots p_m^{\beta_m}$ mit paarweise teilerfremden irreduziblen Polynomen p_i und $\alpha_i \geq 0$, $\beta_i \geq 0$ und $1 \leq i \leq m$. Dann gilt $f|g \Leftrightarrow \alpha_i \leq \beta_i$ für $1 \leq i \leq m$.*

Beweis. Es gilt $f|g \Leftrightarrow g = h \cdot f$ mit $h \in K[x]$, was mit $p_1^{\alpha_1} \cdots p_m^{\alpha_m} = p_1^{\beta_1} \cdots p_m^{\beta_m} \cdot h$ gleichbedeutend ist. Die Behauptung folgt aus der Eindeutigkeit der Primfaktorzerlegung (Satz 8.1.11). □

Satz 8.1.13. *Es seien $f, g \in K[x]$ und $f = p_1^{\alpha_1} \cdots p_m^{\alpha_m}$ sowie $g = p_1^{\beta_1} \cdots p_m^{\beta_m}$ mit paarweise teilerfremden irreduziblen Polynomen p_i und $\alpha_i \geq 0$ und $\beta_i \geq 0$. Dann gilt*

i) $ggT(f, g) = \prod_{i=1}^{m} p_i^{\min\{\alpha_i, \beta_i\}}$

ii) Das $kgV(f, g)$ existiert und ist bis auf einen konstanten Faktor eindeutig bestimmt.
Es ist $kgV(f, g) = \prod_{i=1}^{m} p_i^{\max\{\alpha_i, \beta_i\}}$.

Beweis. i) Es sei $h = \prod_{i=1}^{m} p_i^{\min\{\alpha_i,\beta_i\}}$, und es gelte $k|f$ und $k|g$. Nach Satz 8.1.12 ist dann $k = p_1^{\gamma_1} \cdots p_m^{\gamma_m}$ mit $\gamma_i \leq \alpha_i$ und $\gamma_i \leq \beta_i$. Die Behauptung folgt dann aus Definition 8.1.3.

ii) Der Beweis verläuft analog.

□

Die Begriffe des ggT und des kgV können auch für mehr als zwei Argumente definiert werden.

Definition 8.1.5. Es seien $f_1, \ldots, f_k \in K[x]$.

i) Es ist $ggT(f_1, \ldots, f_k) = ggT(ggT(f_1, \ldots, f_{k-1}), f_k)$

ii) Es ist $kgV(f_1, \ldots, f_k) = kgV(kgV(f_1, \ldots, f_{k-1}), f_k)$

Satz 8.1.14. *Es seien $f_1, \ldots, f_k \in K[x]$ und $f_i = p_1^{\alpha_{1,i}} \cdots p_m^{\alpha_{m,i}}$ mit $1 \leq i \leq k$, $\alpha_{j,i} \geq 0$ und $1 \leq j \leq m$ mit paarweise verschiedenen teilerfremden irreduziblen Polynomen $p_1, \ldots, p_m$. Dann ist*

i) $ggT(f_1, \ldots, f_k) = p_1^{\gamma_1} \cdots p_m^{\gamma_m}$ *mit* $\gamma_j = \min\{\alpha_{j,1}, \ldots, \alpha_{j,k}\}$

ii) $kgV(f_1, \ldots, f_k) = p_1^{\delta_1} \cdots p_m^{\delta_m}$ *mit* $\delta_j = \max\{\alpha_{j,1}, \ldots, \alpha_{j,k}\}$

Beweis. ohne Beweis. □

8.2 Minimalpolynome

In diesem Abschnitt sei K stets ein Körper und V ein endlichdimensionaler Vektorraum über K mit $\dim V = n$.

Definition 8.2.1. (Matrixpolynome und Polynome von Endomorphismen)
Es sei $f(x) = a_0 + \ldots + a_m x^m \in K[x]$.

i) Es sei $\varphi \in L(V, V)$. Dann setzen wir $f(\varphi) = a_0 + a_1\varphi + \ldots + a_m\varphi^m$.

ii) Es sei $\mathcal{A} \in K^{(n,n)}$. Dann setzen wir $f(\mathcal{A}) = a_0 + a_1\mathcal{A} + \ldots + a_m\mathcal{A}^m$.

Satz 8.2.1. *Es sei $f(x) = a_0 + \ldots + a_m x^m \in K[x]$, $\dim V < \infty$ und $\mathcal{B}$ eine Basis von V. Weiter sei $\varphi \in L(V, V)$ und $\mathcal{A} = \mathcal{M}(\varphi; \mathcal{B}, \mathcal{B})$. Dann ist $\mathcal{M}(f(\varphi); \mathcal{B}, \mathcal{B}) = f(\mathcal{A})$.*

Beweis. Dies folgt aus Satz 3.5.7, nach dem die Abbildung $\Phi\colon L(V, V) \to K^{(n,n)}$, $\varphi \to \mathcal{M}(\varphi; \mathcal{B}, \mathcal{B})$ ein Ringisomorphismus, also insbesondere relationstreu ist. □

Satz 8.2.2. *i) Es sei $\varphi \in L(V, V)$ und $f, g \in K[x]$.
Dann ist $(f + g)(\varphi) = f(\varphi) + g(\varphi)$ und $(f \cdot g)(\varphi) = f(\varphi) \cdot g(\varphi)$.*

*ii) Es sei $\mathcal{A} \in K^{(n,n)}$ und $f, g \in K[x]$.
Dann ist $(f + g)(\mathcal{A}) = f(\mathcal{A}) + g(\mathcal{A})$ und $(f \cdot g)(\mathcal{A}) = f(\mathcal{A}) \cdot g(\mathcal{A})$.*

Beweis. Die Behauptung folgt wegen $\varphi^k \circ \varphi^l = \varphi^{k+l}$ und $\mathcal{A}^k \circ \mathcal{A}^l = \mathcal{A}^{k+l}$. □

Satz 8.2.3. *Es sei $\varphi \in L(V,V)$. Zu jedem $\vec{x} \in V - \{\vec{0}\}$ gibt es genau ein Polynom $f \in K[x]$ mit den folgenden Eigenschaften:*

(1) $f(x) = a_0 + a_1 x + \ldots + a_{m-1}x^{m-1} + x^m$ *mit* $m \geq 1$

(2) $f(\varphi)(\vec{x}) = \vec{0}$

(3) Ist $g \in K[x]$ *mit* $g(\varphi)(\vec{x}) = \vec{0}$, *so gilt* $f|g$.

Beweis. Existenz:
Es sei $m \geq 1$ die kleinste natürliche Zahl, für die die Vektoren $\vec{x}, \varphi(\vec{x}), \ldots, \varphi^m(\vec{x})$ linear abhängig sind. Dann gibt es $a_0, \ldots, a_m \in K$ mit $a_0 \cdot id(\vec{x}) + \ldots + a_{m-1}\varphi^{m-1} + a_m\varphi^m(\vec{x}) = \vec{0}$. es ist dann $a_m \neq 0$, und wir können durch Durchdividieren $a_m = 1$ erreichen. Somit hat $f(x) := a_0 + \ldots + a_{m-1}x^{m-1} + x^m$ die Eigenschaften (1) und (2).
Zu (3): Division mit Rest (Satz 8.1.3 ergibt $g = q \cdot f + r$ mit $r = 0$ oder $\deg(r) < m$. es folgt $r(\varphi)(\vec{x}) = \vec{0}$. Wäre $r = b_0 + \ldots + b_s x^s \neq 0$, so wäre $b_0\vec{x} + \ldots + b_s\varphi^s(\vec{x}) = \vec{0}$ im Widerspruch zur linearen Unabhängigkeit der Vektoren $\vec{x}, \varphi(\vec{x}), \ldots, \varphi^s(\vec{x})$. Also ist $r = 0$ und $f|g$.
Eindeutigkeit:
Die Polynome f_1 und f_2 mögen die Eigenschaften (1), (2) und (3) besitzen. Dann gilt $f_1|f_2$ und $f_2|f_1$. Nach Satz 8.1.2 (v) gibt es ein $c \in K - \{0\}$, so dass $f_1 = cf_2$ gilt. Aus(1) folgt $c = 1$. □

Definition 8.2.2. Das Polynom von Satz 8.2.3 heißt Minimalpolynom von $\vec{x}$ bzgl. φ.

Satz 8.2.4. *Es sei $\varphi \in L(V,V)$. Es gibt genau ein Polynom $m \in K[x]$ mit den folgenden Eigenschaften:*

(1) $m(x) = a_0 + a_1 x + \ldots + a_{s-1}x^{s-1} + x^s$ *mit* $s \geq 1$

(2) $m(\varphi)(\vec{x}) = \vec{0}$ *für alle* $\vec{x} \in V$

(3) Ist $g \in K[x]$ *mit* $g(\varphi)(\vec{x}) = \vec{0}$ *für alle* $\vec{x} \in V$, *so gilt* $m|g$.

Beweis. Existenz:
Es sei $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ eine Basis von V, und es seien f_i die Minimalpolynome von $\vec{b}_i$ bzgl. φ mit $1 \leq i \leq n$. Es sei $f = (f_1, \ldots, f_n)$. Dann ist $f(\varphi)(\vec{b}_i) = \vec{0}$ für $1 \leq i \leq n$ und damit auch $f(\varphi)(\vec{x}) = \vec{0}$ für alle $\vec{x} \in V$. Somit hat f die Eigenschaften (1) und (2). Es sei m ein Polynom kleinsten Grades mit diesen Eigenschaften. Die Eigenschaft (3) und die Eindeutigkeit für m folgen dann wie im Beweis von Satz 8.2.3. □

Wir kommen zum entsprechenden Satz für Matrizen:

Satz 8.2.5. *Es sei $\mathcal{A} \in K^{(n,n)}$. Es gibt genau ein Polynom $m \in K[x]$ mit den folgenden Eigenschaften:*

(1) $m(x) = a_0 + a_1 x + \ldots + a_{s-1}x^{s-1} + x^s$ *mit* $s \geq 1$

(2) $m(\mathcal{A}) = 0^{(n,n)}$

(3) Ist $g \in K[x]$ *mit* $g(\mathcal{A}) = \vec{0}^{(n,n)}$, *so gilt* $m|g$.

Das Polynom m hat die Eigenschaften (1)-(3) für den Endomorphismus $\varphi_{\mathcal{A}} \colon \vec{x} \to \mathcal{A}\vec{x}$.

Beweis. Nach Satz 8.2.1 hat m genau dann die Eigenschaften (1)-(3), wenn es die Eigenschaften (1)-(3) von Satz 8.2.4 für $\varphi_{\mathcal{A}}$ hat. Die Behauptung folgt aus Satz 8.2.4. □

Definition 8.2.3. i) Es sei $\varphi \in L(V, V)$. Das Polynom m von Satz 8.2.4 heißt Minimalpolynom von φ.

ii) Es sei $\mathcal{A} \in K^{(n,n)}$. Das Polynom m von Satz 8.2.4 heißt Minimalpolynom von $\mathcal{A}$.

8.3 Direkte Summen, Zerlegung in invariante Unterräume

In diesem Abschnitt sei K stets ein Körper und V ein endlichdimensionaler Vektorraum über K mit $\dim V < \infty$.

Definition 8.3.1. Es seien $U_1, \ldots, U_r$ Unterräume von V.

i) Die Unterräume $U_1, \ldots, U_r$ heißen unabhängig, wenn aus $\vec{u}_1 + \ldots + \vec{u}_r = \vec{0}$ mit $\vec{u}_i \in U_i$ dann $\vec{u}_i = \vec{0}$ für $1 \leq i \leq r$ folgt.

ii) Man nennt V direkte Summe der Unterräume U_i (Schreibweise: $V = U_1 \oplus \ldots \oplus U_r$), wenn $V = U_1 + \ldots + U_r$ gilt und $U_1, \ldots, U_r$ unabhängig sind.

Satz 8.3.1. *i) Zwei Unterräume U_1 und U_2 von V sind genau dann unabhängig, wenn die Bedingung $U_1 \cap U_2 = \{\vec{0}\}$ gilt.*

ii) Es gilt genau dann $V = U_1 \oplus U_2$, wenn $V = U_1 + U_2$ und $U_1 \cap U_2 = \{\vec{0}\}$ gilt.

Beweis. i) "$\Rightarrow$":
Es sei $\vec{u} \in U_1 \cap U_2$. Wegen $\vec{u} + (-\vec{u}) = \vec{0}$ folgt $\vec{u} = \vec{0}$.
"$\Leftarrow$":
Es sei $\vec{u}_1 + \vec{u}_2 = \vec{0}$ mit $\vec{u}_1 \in U_1$ und $\vec{u}_2 \in U_2$. Dann ist $\vec{u}_1 = -\vec{u}_2 \in U_1 \cap U_2$, also $\vec{u}_1 = \vec{u}_2 = \vec{0}$.

ii) Dies folgt aus (i) und aus Definition 8.3.1.

□

Satz 8.3.2. *(Verfeinerung einer direkten Summenzerlegung)*
Es sei $V = U_1 \oplus \ldots \oplus U_r$ und $U_i = U_{i,1} \oplus \ldots \oplus U_{i,s_i}$ mit $1 \leq i \leq r$. Dann ist

$$V = U_{1,1} \oplus \ldots \oplus U_{1,s_1} \oplus \ldots \oplus U_{r,1} \oplus \ldots \oplus U_{r,s_r}.$$

Beweis. Für $\vec{v} \in V$ gibt es $\vec{u}_i \in U_i$ mit

$$\vec{v} = \vec{u}_1 + \ldots + \vec{u}_r. \tag{1}$$

Weiter gibt es

$$\vec{u}_{i,j} \in U_{i,j} \quad \text{mit} \quad \vec{u}_i = \vec{u}_{i,1} + \ldots \vec{u}_{i,s_i}. \tag{2}$$

Aus (1) und (2) folgt

$$\vec{v} = \vec{u}_{1,1} + \ldots + \vec{u}_{1,s_1} + \ldots + \vec{u}_{r,1} + \ldots + \vec{u}_{r,s_r},$$

also $V = U_{1,1} + \ldots + U_{r,s_r}$.
Es sei $\vec{u}_{1,1} + \ldots + \vec{u}_{1,s_1} + \ldots + \vec{u}_{r,1} + \ldots + \vec{u}_{r,s_r} = \vec{0}$ für $\vec{u}_{i,j} \in U_{i,j}$. Da die U_i unabhängig sind, folgt $\vec{w}_i = \vec{u}_{1,1} + \ldots + \vec{u}_{i,s_i} = \vec{0}$ für $1 \leq i \leq r$. Wegen der Unabhängigkeit der $U_{i,j}$ folgt $\vec{u}_{i,j} = \vec{0}$ für alle i und j. □

Satz 8.3.3. *i) Es sei $V = U_1 \oplus \ldots \oplus U_r$, und U_i habe die Basis $\mathcal{B}_i$ für $1 \leq i \leq r$. Dann ist $\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_r$ eine Basis von V.*

ii) Es gilt $\dim V = \dim U_1 + \ldots + \dim U_r$.

Beweis. i) Es sei $\mathcal{B}_i = \{\vec{b}_{i,1}, \ldots, \vec{b}_{i,s_i}\}$ und $\vec{v} \in V$. Dann gibt es $\vec{u}_i \in U_i$ mit $\vec{v} = \vec{u}_1 + \ldots + \vec{u}_r$ und $\lambda_{i,j} \in K$ mit $\vec{v}_i = \lambda_{i,1}\vec{b}_{i,1} + \ldots + \lambda_{i,s_i}\vec{b}_{i,s_i}$. Also gilt

$$\vec{v} = \lambda_{1,1}\vec{b}_{1,1} + \ldots + \lambda_{1,s_1}\vec{b}_{1,s_1} + \ldots + \lambda_{r,1}\vec{b}_{r,1} + \ldots + \lambda_{r,s_r}\vec{b}_{r,s_r},$$

womit $V = \langle \mathcal{B} \rangle$ gilt. Es sei $\lambda_{1,1}\vec{b}_{1,1} + \ldots + \lambda_{r,s_r}\vec{b}_{r,s_r} = \vec{0}$. Wegen der Unabhängigkeit der U_i folgt $\vec{w}_i = \lambda_{i,1}\vec{b}_{i,1} + \ldots + \lambda_{i,s_i}\vec{b}_{i,s_i} = \vec{0}$ für $1 \leq i \leq r$ und damit $\lambda_{1,1} = \ldots = \lambda_{r,s_r} = 0$, also $V = \langle\langle \mathcal{B} \rangle\rangle$.

ii) Dies folgt sofort aus Satz 2.4.6.

□

Definition 8.3.2. Es sei $\varphi \in L(V, V)$.

i) Ein Unterraum U von V heißt <u>φ- invarianter Unterraum</u> von V, falls $\varphi(U) \subset U$ gilt.

ii) Wir sagen, V <u>zerfällt bzgl. des Endomorphismus φ</u>, wenn $r \geq 2$ ist und es φ- invariante Unterräume $U_1, \ldots, U_r$ von V mit $U_i \neq \{\vec{0}\}$ und $V = U_1 \oplus \ldots \oplus U_r$ gibt.

Satz 8.3.4. *Es sei $\varphi \in L(V, V)$ und $V = U_1 \oplus \ldots \oplus U_r$ mit φ- invarianten Unterräumen U_i, wobei U_i die Basis $\mathcal{B}_i$ habe. Es sei $\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_r$. Dann besitzt die Darstellungsmatrix $\mathcal{M}(\varphi, \mathcal{B}, \mathcal{B})$ folgende <u>Kästchenform</u>*

$$\mathcal{M}(\varphi, \mathcal{B}, \mathcal{B}) = \begin{pmatrix} \mathcal{M}_1 & & & 0 \\ & \mathcal{M}_2 & & \\ & & \ddots & \\ 0 & & & \mathcal{M}_r \end{pmatrix}$$

mit entlang der Diagonalen angeordneten quadratischen Teilmatrizen $\mathcal{M}_i$ mit $\mathcal{M}_i = \mathcal{M}(\varphi_{|U_i}, \mathcal{B}_i, \mathcal{B}_i)$.

Beweis. Durch Nachrechnen. □

Satz 8.3.5. *Es sei $\varphi \in L(V, V)$ und $f_1, f_2 \in K[x]$ mit $f = f_1 f_2$ und $ggT(f_1, f_2) = 1$. Weiter sei $U_i = \operatorname{Kern} f_i(\varphi)$ mit $i = 1, 2$. Ist $f(\varphi)(\vec{x}) = \vec{0}$ für alle $\vec{x} \in V$, so ist $V = U_1 \oplus U_2$ und U_1 und U_2 sind φ- invariante Unterräume.*

Beweis. Nach Satz 8.1.7 gibt es $h_1, h_2 \in K[x]$ mit $h_1 f_1 + h_2 f_2 = 1$ und damit

$$h_1(\varphi) f_1(\varphi)(\vec{x}) + h_2(\varphi) f_2(\varphi)(\vec{x}) = \vec{x} \tag{1}$$

für alle $\vec{x} \in V$. Es ist weiter

$$f_2(\varphi) h_1(\varphi) f_1(\varphi)(\vec{x}) = h_1(\varphi) f(\varphi)(\vec{x}) = \vec{0},$$

also

$$h_1(\varphi) f_1(\varphi)(\vec{x}) \in \operatorname{Kern}(f_2(\varphi)) = U_2 \tag{2}$$

und analog

$$h_2(\varphi) f_2(\varphi)(\vec{x}) \in U_1. \tag{3}$$

Aus (1), (2) und (3) folgt

$$U_1 + U_2 = V. \tag{4}$$

Ist $\vec{x} \in U_1 \cap U_2$, so ist $h_1(\varphi)f_1(\varphi)(\vec{x}) = \vec{0}$ sowie $h_2(\varphi)f_2(\varphi)(\vec{x}) = \vec{0}$, also ist nach (1) auch $\vec{x} = \vec{0}$. Damit ist

$$U_1 \cap U_2 = \{\vec{0}\}. \tag{5}$$

Aus (4) und (5) folgt $V = U_1 \oplus U_2$.
Ist $\vec{x} \in U_i$, so folgt $f_i(\varphi)(\varphi(\vec{x})) = \varphi(f_i(\varphi)(\vec{x})) = \varphi(\vec{0}) = \vec{0}$, d.h. $\varphi(\vec{x}) \in U_i$. Daher sind U_1 und U_2 auch φ- invariant. $\square$

Satz 8.3.6. *Es sei $\varphi \in L(V, V)$ mit Minimalpolynom m, für das $m = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$ mit $r \geq 2$ und paarweise verschiedenen teilerfremden irreduziblen Polynomen p_i gelte. Dann ist $V = U_1 \oplus \ldots \oplus U_r$ mit $U_i = \operatorname{Kern} p_i^{\alpha_i}$, wobei $p_i^{\alpha_i}$ die Minimalpolynome der Restriktionen $\varphi_{|U_i}$ sind.*

Beweis. Wir führen den Beweis durch Induktion nach r:
Es sei $f_1 = p_1^{\alpha_1}$ und $f_2 = p_2^{\alpha_2} \cdots p_r^{\alpha_r}$. Nach Satz 8.3.5 ist $V = U_1 \oplus W_2$ mit $U_1 = \operatorname{Kern}(f_1(\varphi)$ und $W_2 = \operatorname{Kern}(f_2(\varphi)$. Es seien m_1 bzw. m_2 die Minimalpolynome von $\varphi_{|U_1}$ bzw. $\varphi_{|W_2}$.
Es sei $\vec{x} \in V$. Dann gibt es $\vec{u}_1 \in U_1$ und $\vec{w}_2 \in W_2$ mit $\vec{x} = \vec{u}_1 + \vec{w}_2$. Es ist dann

$$(m_1 \cdot m_2)(\varphi)(\vec{x}) = m_2(\varphi)m_1(\varphi)\vec{u}_1 + m_1(\varphi)m_2(\varphi)\vec{w}_2 = \vec{0}.$$

Nach Satz 8.2.4 gilt

$$(f_1 f_2)|(m_1 m_2). \tag{1}$$

Andererseits gilt wegen $f_1(\varphi)(U_1) = \{\vec{0}\}$ und $f_2(\varphi)(W_2) = \{\vec{0}\}$ auch

$$m_1|f_1 \quad \text{und} \quad m_2|f_2. \tag{2}$$

Aus (1) und (2) folgt $m_1 = f_1$ und $m_2 = f_2$.
Wir wenden dann auf W_2 und $\varphi_{|W_2}$ die Induktionshypothese an und erhalten die Behauptung. $\square$

8.4 Allgemeine Normalform

Die in diesem Abschnitt besprochene Normalform kann für Vektorräume über einem beliebigen Körper K erhalten werden. Im nächsten Abschnitt diskutieren wir dann die Jordansche Normalform über dem Körper $\mathbb{C}$ der komplexen Zahlen.

Nach den Sätzen 8.3.4 und 8.3.6 gibt es zu $\varphi \in L(V, V)$ stets eine Basis $\mathcal{B}$, so dass die Darstellungsmatrix $\mathcal{M}(\varphi, \mathcal{B}, \mathcal{B})$ eine Kästchenform

$$\mathcal{M}(\varphi, \mathcal{B}, \mathcal{B}) = \begin{pmatrix} \mathcal{M}_1 & & & 0 \\ & \mathcal{M}_2 & & \\ & & \ddots & \\ 0 & & & \mathcal{M}_r \end{pmatrix}$$

hat, so dass die Minimalpolynome der Teilmatrizen $\mathcal{M}_i$ Potenzen irreduzibler Polynome sind. Wir erhalten die allgemeine Normalform von $\mathcal{M}(\varphi, \mathcal{B}, \mathcal{B})$, indem wir weitere Basen einführen, so dass die $\mathcal{M}_i$ Normalformen erhalten.

Wir geben ohne Beweis folgenden

Satz 8.4.1. *Es sei* $\mathcal{A} \in K^{(n,n)}$, $p \in K[x]$, $\deg(p) = k$ *und* p^m *das Minimalpolynom von* $\mathcal{A}$. *Es sei* $\varphi_{\mathcal{A}} \colon \vec{x} \to \mathcal{A}\vec{x}$ *der zu* $\mathcal{A}$ *gehörige Endomorphismus. Dann gibt es eine Basis* $\mathcal{B}$ *des* K^n, *so dass*

$$\mathcal{M}(\varphi_{\mathcal{A}}, \mathcal{B}, \mathcal{B}) = \begin{pmatrix} \mathcal{M}_{11} & & & & & & \\ & \ddots & & & & & \\ & & \mathcal{M}_{1r_1} & & & & \\ & & & \ddots & & & \\ & & & & \mathcal{M}_{m1} & & \\ & & & & & \ddots & \\ & & & & & & \mathcal{M}_{mr_m} \end{pmatrix}$$

mit

$$\mathcal{M}_{ij} = \begin{pmatrix} 0 & 0 & \dots & 0 & a_{i,0} \\ 1 & 0 & \dots & & a_{i,1} \\ 0 & 1 & & & \vdots \\ & & \ddots & & \vdots \\ & & & 1 & a_{i,ik-1} \end{pmatrix},$$

wobei

$$p(x)^i = -a_{i,0} - a_{i,1}x - \ldots - a_{i,ik-2}x^{ik-2} - a_{i,ik-1}x^{ik-1} + x^{ik}$$

ist. Für $i < m$ *kann* $r_i = 0$ *sein. Dann treten die Kästchen* $\mathcal{M}_{ij}$ *nicht auf. Das Kästchen* $\mathcal{M}_{m1}$ *tritt immer auf. Das Minimalpolynom von* $\mathcal{M}_{ij}$ *ist* p^i.

Beweis. ohne Beweis. □

Satz 8.4.2. *Es sei*

$$\mathcal{A} = \begin{pmatrix} 0 & 0 & \dots & 0 & a_0 \\ 1 & 0 & \dots & & a_1 \\ 0 & 1 & & & \vdots \\ & & \ddots & & \vdots \\ & & & 1 & a_{k-1} \end{pmatrix} \in K^{(k,k)}.$$

Dann ist $P_{\mathcal{A}}(\lambda) = (-1)^k \cdot (-a_0 - \ldots - a_{k-1}\lambda^{k-1} + \lambda^k)$.

Beweis. Wir zeigen dies mittels Induktion nach k:
Induktionsanfang: $k = 2$:
Es gilt

$$\det \begin{pmatrix} -\lambda & a_0 \\ 1 & a_1 - \lambda \end{pmatrix} = (-1)^2 \cdot (\lambda^2 - a_1\lambda - a_0).$$

Induktionsschritt: $k \to k+1$:
Nun gilt mittels Entwicklung nach der ersten Zeile und Einsetzen der Induktionshypothese

$$
\begin{aligned}
P_{\mathcal{A}}(\lambda) &= \det \begin{pmatrix} -\lambda & 0 & \dots & 0 & a_0 \\ 1 & -\lambda & \dots & & a_1 \\ 0 & 1 & & & \vdots \\ & & \ddots & & \vdots \\ & & & 1 & a_{k-1}-\lambda \end{pmatrix} \\
&= -\lambda \cdot \det \begin{pmatrix} -\lambda & \dots & a_1 & \\ 1 & & \vdots & \\ & \ddots & & \vdots \\ & & 1 & a_{k-1}-\lambda \end{pmatrix} + (-1)^{k+1} a_0 \\
&= (-\lambda) \cdot (-1)^{k-1} \cdot (-a_0 - \ldots - a_{k-2}\lambda^{k-2} + \lambda^{k-1}) + (-1)^{k+1} a_0 \\
&= (-1)^k \cdot (-a_0 - \ldots - a_{k-1}\lambda^{k-1} + \lambda^k).
\end{aligned}
$$

□

Satz 8.4.3. *(Cayley- Hamilton)*
Es sei $\mathcal{A} \in K^{(n,n)}$ mit dem charakteristischen Polynom $P_{\mathcal{A}}(\lambda)$.

i) Es ist $P_{\mathcal{A}}(\mathcal{A}) = 0^{(n,n)}$.

ii) Für das Minimalpolynom m von $\mathcal{A}$ gilt $m | P_{\mathcal{A}}$.

Beweis. Es sei $\varphi_{\mathcal{A}}\colon \vec{x} \to \mathcal{A}\vec{x}$. Nach den Sätzen 8.3.6, 8.4.1 und 8.4.2 gibt es eine Basis $\mathcal{B}$, so dass

$$
\mathcal{M} = \mathcal{M}(\varphi_{\mathcal{A}}, \mathcal{B}, \mathcal{B}) = \begin{pmatrix} \mathcal{M}_1 & & 0 \\ & \ddots & \\ 0 & & \mathcal{M}_r \end{pmatrix}
$$

gilt, wobei für die Minimalpolynome m_i der Teilmatrizen $\mathcal{M}_i$ von $K^{(s_i,s_i)}$ dann $m_i = \pm P_{\mathcal{M}_i}$ gilt. Es ist dann

$$
P_{\mathcal{A}}(\lambda) = P_{\mathcal{M}}(\lambda) = \det \begin{pmatrix} \mathcal{M}_1 - \lambda\mathcal{E}_{s_1} & & & 0 \\ & \mathcal{M}_2 - \lambda\mathcal{E}_{s_2} & & \\ & & \ddots & \\ 0 & & & \mathcal{M}_r - \lambda\mathcal{E}_{s_r} \end{pmatrix}
$$

$$
= \prod_{i=1}^{r} \det(\mathcal{M}_i - \lambda\mathcal{E}_{s_i}) = \pm \prod_{i=1}^{r} m_i. \tag{1}
$$

Das Minimalpolynom m von $\mathcal{M}$ ist

$$
m = kgV(m_1, \ldots, m_r). \tag{2}
$$

Aus (1) und (2) folgt $m | P_{\mathcal{A}}$ und somit $P_{\mathcal{A}}(\mathcal{A}) = 0^{(n,n)}$. □

8.5 Jordansche Normalform

In diesem Abschnitt betrachten wir nur den Körper $\mathbb{C}$ der komplexen Zahlen.

Satz 8.5.1. *Die einzigen irreduziblen Polynome von $\mathbb{C}[x]$ sind die linearen Polynome.*

Beweis. Es sei $f \in \mathbb{C}[x]$ und $\deg(f) \geq 2$. Nach dem Fundamentalsatz der Algebra (Satz 7.3.1) hat f eine Nullstelle $x_0 \in \mathbb{C}$. Damit gilt $f(x) = (x - x_0) \cdot g(x)$ mit $\deg(g) \geq 1$. $\square$

Satz 8.5.2. *Es sei $\mathcal{A} \in \mathbb{C}^{(n,n)}$. Eine zu $\mathcal{A}$ ähnliche Matrix J heißt Jordansche Normalform von $\mathcal{A}$, wenn J eine Kästchenform der folgenden Art hat:*

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_r \end{pmatrix}$$

mit Jordankästchen

$$J_i = \begin{pmatrix} \lambda_i & & & 0 & \\ 1 & \lambda_i & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & 0 & & 1 & \lambda_i \end{pmatrix} \in \mathbb{C}^{(s_i,s_i)}$$

mit $s_i \in \mathbb{N}$.

Bemerkung 8.5.1. Im Fall $s_i = 1$ ist $J = \big(\lambda_i\big)$.
Nach den Sätzen 8.3.4 und 8.3.6 lässt sich alles auf den Fall zurückführen, dass für das Minimalpolynom von $\mathcal{A}$ nun $m(x) = (x - \lambda)^s$ mit $\lambda \in \mathbb{C}$ und $s \in \mathbb{N}$ gilt.

Hier haben wir den folgenden

Satz 8.5.3. *Es sei V ein n- dimensionaler Vektorraum über $\mathbb{C}$ und $\varphi \in L(V,V)$. Das Minimalpolynom von φ sei $m(x) = (x-\lambda)^{s_0}$. Dann gibt es $s_1 \geq s_2 \geq \ldots \geq s_r \geq 1$ mit $s_1 = s_0$ und $s_1 + \ldots + s_r = n$ sowie Vektoren $\vec{x}_i \in V$ mit Minimalpolynomen $m_i(x) = (x - \lambda)^{s_i}$, so dass*

$$\mathcal{B} = \bigcup_{i=1}^{r} \big\{(\varphi - \lambda)^j(\vec{x}_i) \colon 0 \leq j < s_i\big\}$$

eine Basis von V bildet. Die Darstellungsmatrix bzgl. $\mathcal{B}$ ist

$$\mathcal{M} = \mathcal{M}(\varphi, \mathcal{B}, \mathcal{B}) = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{pmatrix}$$

mit Jordankästchen

$$J_i = \begin{pmatrix} \lambda_i & & & 0 & \\ 1 & \lambda_i & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & 0 & & 1 & \lambda_i \end{pmatrix} \in \mathbb{C}^{(s_i,s_i)}$$

mit $1 \leq i \leq r$.

Beweis. ohne Beweis. $\square$

Beispiel 8.5.1. Es sei

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

Finde die Jordansche Normalform von $\mathcal{A}$.

<u>Lösung:</u>

<u>1. Schritt:</u>

Bestimmung des Minimalpolynoms: nach Satz 8.4.3 (Satz von Cayley- Hamilton) gilt für das Minimalpolynom m von $\mathcal{A}$ die Aussage $m|P_{\mathcal{A}}$. Das charakteristische Polynom $P_{\mathcal{A}}$ ergibt sich zu

$$P_{\mathcal{A}} = \det \begin{pmatrix} 1-\lambda & 0 & 1 \\ -2 & 3-\lambda & 0 \\ -1 & 1 & 1-\lambda \end{pmatrix} = (1-\lambda) \cdot (2-\lambda)^2.$$

Für jeden Eigenwert λ_i von $\mathcal{A}$ gilt $(\lambda - \lambda_i)|m$, da $(\lambda - \lambda_i)$ Minimalpolynom jedes Eigenvektors zum Eigenwert λ_i ist. Also ist $m_0(x) = (x-1)(x-2)$ oder $m_1(x) = (x-1)(x-2)^2$ das Minimalpolynom von $\mathcal{A}$. Rechnung ergibt $(\mathcal{A} - \mathcal{E})(\mathcal{A} - 2\mathcal{E}) \neq 0^{(3,3)}$, also ist $m_1(x)$ das Minimalpolynom.

<u>2. Schritt:</u>

Zerlegung von $V = \mathbb{C}^3$ in $\varphi_{\mathcal{A}}$- invariante Unterräume nach Satz 8.3.6:

Es ist $m_1(x) = p_1(x)^2 \cdot p_2(x)$ mit $p_1(x) = x - 2$ und $p_2(x) = x - 1$. Nach Satz 8.3.6 ist $V = U_1 \oplus U_2$ mit $U_1 = \operatorname{Kern} p_1(\varphi_{\mathcal{A}})^2$ und $U_2 = \operatorname{Kern} p_2(\varphi_{\mathcal{A}})$. Der Unterraum U_1 ergibt sich als Lösung des LGS

$$(\mathcal{A} - 2\mathcal{E})^2 \vec{x} = \vec{0} \Rightarrow \begin{pmatrix} 0 & 1 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} \vec{x} = \vec{0}.$$

Damit ist $U_1 = \langle\langle \vec{e}_1, (0,2,1)^T \rangle\rangle$.

Wir wenden nun Satz 8.5.2 auf $\varphi_{\mathcal{A}_{|U_1}}$ mit $\vec{x}_1 = \vec{e}_1 = \vec{b}_{1,1}$ und $(\varphi - 2id)\vec{e}_1 = (-1,-2,-1)^T = \vec{b}_{1,2}$ an und erhalten die Basis $\mathcal{B}_1 = \{\vec{b}_{1,1}, \vec{b}_{1,2}\}$ von U_1. Es ist

$$\mathcal{M}(\varphi_{\mathcal{A}_{|U_1}}, \mathcal{B}_1, \mathcal{B}_1) = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}. \tag{1}$$

Der Unterraum U_2 ergibt sich nun als Lösung des LGS

$$(\mathcal{A} - \mathcal{E})\vec{x} = \vec{0} \Rightarrow \begin{pmatrix} 0 & 0 & 1 \\ -2 & 2 & 0 \\ -1 & 1 & 0 \end{pmatrix} \vec{x} = \vec{0},$$

also $U_2 = \langle \vec{b}_2 \rangle$ mit $\vec{b}_2 = (1,1,0)^T$ und erhalten die Basis $\mathcal{B}_2 = \{\vec{b}_2\}$. Es ist

$$\mathcal{M}(\varphi_{\mathcal{A}_{|U_2}}, \mathcal{B}_2, \mathcal{B}_2) = \begin{pmatrix} 1 \end{pmatrix}. \tag{2}$$

Es sei $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$. Aus (1) und (2) erhalten wir dann

$$\mathcal{M}(\varphi_{\mathcal{A}}, \mathcal{B}, \mathcal{B}) = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Nach Satz 5.2.2 ist dann

$$\mathcal{X}^{-1} \mathcal{A} \mathcal{X} = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

mit

$$\mathcal{X} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & -2 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Definition 8.5.1. Es sei $\mathcal{A} \in \mathbb{C}^{(n,n)}$ mit $P_{\mathcal{A}}(\lambda) = (-1)^n(\lambda - \lambda_i)^{s_1} \cdots (\lambda - \lambda_r)^{s_r}$, wobei die λ_i paarweise verschieden seien. Dann heißt s_i die <u>algebraische Ordnung</u> des Eigenwerts λ_i.

Satz 8.5.4. *Es sei $\mathcal{A} \in \mathbb{C}^{(n,n)}$.*

i) Die geometrische Ordnung eines Eigenwertes ist kleiner gleich seiner algebraischen Ordnung.

ii) Die Matrix $\mathcal{A}$ ist genau dann diagonalisierbar, wenn für alle Eigenwerte die geometrische Ordnung gleich der algebraischen Ordnung ist.

Beweis. ohne Beweis. □

Kapitel 9

Matrixnormen und Systeme linearer Differentialgleichungen

9.1 Vektornormen und Matrixnormen

In diesem Kapitel sei $K = \mathbb{R}$ oder $K = \mathbb{C}$.
Wir verallgemeinern zunächst den Begriff der Norm von Definition 6.2.1:

Definition 9.1.1. Es sei V ein Vektorraum über K. Unter einer Vektornorm auf V versteht man eine Abbildung $||\cdot||\colon V \to \mathbb{R}$, so dass

(I) $||\vec{x}|| \geq 0$ für alle $\vec{x} \in V$ und $||\vec{x}|| = 0 \Leftrightarrow \vec{x} = \vec{0}$

(II) $||\lambda\vec{x}|| = |\lambda| \cdot ||\vec{x}||$ für alle $\lambda \in K$ und alle $\vec{x} \in V$

(III) $||\vec{x} + \vec{y}|| \leq ||\vec{x}|| + ||\vec{y}||$ für alle $\vec{x}, \vec{y} \in V$.

Beispiel 9.1.1. i) Die zu einem Euklidischen oder unitären Vektorraum gehörige Norm nach Definition 6.2.1.

ii) Für $1 \leq p < \infty$:

$$||\vec{x}||_p = \left(\sum_{i=1}^{n} |x_i^p|\right)^{1/p}$$

mit $\vec{x} = (x_1, \ldots, x_n)^T$. Insbesondere ist $||\vec{x}||_1 = \sum_{i=1}^{n} |x_i|$.

iii) $||\vec{x}||_\infty = \max_{1\leq i\leq n} |x_i|$.

Definition 9.1.2. Unter einer Matrixnorm $||\cdot||$ versteht man eine Vektornorm auf dem Vektorraum $K^{(n,n)}$ im Sinne von Definition 9.1.1.

Definition 9.1.3. Es sei $||\cdot||$ eine Matrixnorm und $||\cdot||_V$ eine Vektornorm. Die Norm $||\cdot||$ heißt

i) submultiplikativ, falls $||\mathcal{A}\mathcal{B}|| \leq ||\mathcal{A}|| \cdot ||\mathcal{B}||$ für alle $\mathcal{A}, \mathcal{B} \in K^{(n,n)}$ gilt.

ii) konsistent mit $||\cdot||_V$, falls $||\mathcal{A}\vec{x}||_V \leq ||\mathcal{A}|| \cdot ||\vec{x}||_V$ für alle $\mathcal{A} \in K^{(n,n)}$ und alle $\vec{x} \in K^n$ gilt.

iii) Es heißt

$$||\mathcal{A}|| := \max_{||\vec{x}||_V=1} ||\mathcal{A}\vec{x}||_V = \max_{\vec{x}\neq\vec{0}} \frac{||\mathcal{A}\vec{x}||_V}{||\vec{x}||_V}$$

die von $||\cdot||$ induzierte Norm.

Satz 9.1.1. *Es sei* $||\cdot||_V$ *eine Vektornorm und* $||\cdot||$ *die von* $||\cdot||_V$ *induzierte Matrixnorm. Dann gilt*

i) Die Norm $||\cdot||$ *ist konsistent mit* $||\cdot||_V$.

ii) Die Norm $||\cdot||$ *ist submultiplikativ.*

Beweis. i) Dies folgt unmittelbar aus Definition 9.1.3 (iii).

ii) Es sei $\mathcal{A}, \mathcal{B} \in K^{(n,n)}$. Nach Definition 9.1.3 (iii) gilt

$$||\mathcal{AB}|| \leq \max_{||\vec{x}||_V=1} ||\mathcal{AB}\vec{x}||_V \underset{(i)}{\leq} \max_{||\vec{x}||_V=1} ||\mathcal{A}|| \cdot ||\mathcal{B}\vec{x}||_V = ||\mathcal{A}|| \cdot \max_{||\vec{x}||_V=1} ||\mathcal{B}\vec{x}||_V = ||\mathcal{A}|| \cdot ||\mathcal{B}||.$$

□

Definition 9.1.4. Es sei $\mathcal{A} = (a_{ij})_{1\leq i,j\leq n} \in K^{(n,n)}$. Unter der Zeilensummenmaximumsnorm von $\mathcal{A}$ versteht man

$$||\mathcal{A}|| = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}|.$$

Satz 9.1.2. *Die von der Vektornorm* $||\vec{x}||_\infty$ *induzierte Matrixnorm ist die Zeilensummenmaximumsnorm.*

Beweis. ohne Beweis. □

9.2 Unendliche Folgen und Reihen von Matrizen

Definition 9.2.1. Es sei $n \in \mathbb{N}$, $l \in \mathbb{N}_0$, $||\cdot||$ eine Matrixnorm auf $K^{(n.n)}$ und $\mathcal{A}_l \in K^{(n,n)}$. Die Matrixfolge $(\mathcal{A}_l)_{l=0}^{\infty}$ heißt konvergent gegen $\mathcal{A} \in K^{(n,n)}$ (bzgl. $||\cdot||$) (Schreibweise: $\lim_{l\to\infty} \mathcal{A}_l = \mathcal{A}$), wenn $\lim_{l\to\infty} ||\mathcal{A}_l - \mathcal{A}|| = 0$ ist.
Es sei $\vec{x}_l \in K^n$. Dann heißt $(\vec{x}_l)_{l=0}^{\infty}$ konvergent gegen $\vec{x} \in K^n$ (Schreibweise: $\lim_{l\to\infty} \vec{x}_l = \vec{x}$), wenn $\lim_{l\to\infty} ||\vec{x}_l - \vec{x}|| = 0$ ist. Hier bedeutet $||\cdot||$ die Norm im Sinne von Definition 6.2.1.

Bemerkung 9.2.1. Die Eigenschaft der Konvergenz kann von der gewählten Norm abhängen. Die Begriffe sind äquivalent, wenn die Normen äquivalent sind.

Definition 9.2.2. Es seien $||\cdot||_1$ und $||\cdot||_2$ Matrixnormen auf dem $K^{(n,n)}$. Dann heißen $||\cdot||_1$ und $||\cdot||_2$ äquivalent, wenn es Konstanten $c_1, c_2 > 0$ gibt, so dass

$$c_1||\mathcal{A}||_1 \leq ||\mathcal{A}||_2 \leq c_2||\mathcal{A}||_1$$

für alle $\mathcal{A} \in K^{(n,n)}$ gilt. Eine entsprechende Definition gilt für Vektornormen.

Definition 9.2.3. Es sei $n \in \mathbb{N}$ und $\mathcal{A} = (a_{ij})_{1\leq i,j\leq n} \in K^{(n,n)}$. Die Schurnorm $||\cdot||_S$ ist durch

$$||\mathcal{A}||_S = \left(\sum_{i,j=1}^{n} |a_{ij}|^2\right)^{1/2}$$

definiert.

Satz 9.2.1. *Die Schurnorm ist submultiplikativ und konsistent mit* $||\cdot||_2$.

Beweis. Die Submultiplikativität folgt aus der Cauchy- Schwarzschen- Ungleichung, die Konsistenz mit $||\cdot||_2$ wegen

$$\sum_{i,j=1}^{n} |a_{ij}x_j|^2 \leq \left(\sum_{i,j=1}^{n} |a_{ij}|^2\right) \cdot \left(\sum_{j=1}^{n} |x_j|^2\right).$$

□

Bemerkung 9.2.2. Wird $\mathcal{A}$ als Vektor des $\mathbb{R}^{n^2}$ betrachtet, so ist die Euklidische Norm gerade die Schurnorm. Die Konvergenzdefinition 9.2.1 stimmt mit derjenigen aus der Analysis überein.

Satz 9.2.2. *i) Die Matrixnorm $||\cdot||$ auf dem $K^{(n,n)}$ sei äquivalent zur Schurnorm. Es seien $\mathcal{A}_l = (a_{ij}^{(l)})_{1\leq i,j\leq n}$ und $\mathcal{A} = (a_{ij})_{1\leq i,j\leq n}$. Dann gilt*

$$\lim_{l\to\infty} \mathcal{A}_l = \mathcal{A} \Leftrightarrow \lim_{l\to\infty} a_{ij}^{(l)} = a_{ij}$$

für alle $1 \leq i,j \leq n$.

ii) Es seien $\vec{x}_l, \vec{x} \in K^n$ mit $\vec{x}_l = (x_1^{(l)}, \ldots, x_n^{(l)})^T$ und $\vec{x} = (x_1, \ldots, x_n)^T$. Dann gilt

$$\lim_{l\to\infty} \vec{x}_l = \vec{x} \Leftrightarrow \lim_{l\to\infty} x_i^{(l)} = x_i$$

für alle $1 \leq i \leq n$.

Beweis. i) Für $\mathcal{C} = (c_{ij})_{1\leq i,j\leq n}$ ist

$$\max_{1\leq i,j\leq n} |c_{ij}| \leq ||\mathcal{C}||_S \leq n^2 \cdot \max_{1\leq i,j\leq n} |c_{ij}|. \tag{1}$$

Ist $||\cdot||$ zu $||\cdot||_S$ äquivalent, so gibt es wegen (1) positive Konstanten C_1, C_2 mit

$$C_1 \max_{1\leq i,j\leq n} |c_{ij}| \leq ||\mathcal{C}||_S \leq C_2 \cdot \max_{1\leq i,j\leq n} |c_{ij}|. \tag{2}$$

Die Behauptung folgt nun aus (2) und aus der Konvergenzdefinition der Analysis, wenn wir $c_{ij} = a_{ij} - a_{ij}^{(l)}$ setzen.

ii) Dies ist aus der Vorlesung "Analysis" bekannt.

□

Im Rest von Kapitel 9 sei stets vorausgesetzt, dass $||\cdot||$ eine zur Schurnorm äquivalente Matrixnorm ist. Die vorkommenden Folgen und Funktionen haben dann genau dann Eigenschaften wie Stetigkeit, Differenzierbarkeit, etc. bzgl. einer dieser Matrixnormen, wenn sie sie bzgl. aller Matrixnormen haben. Mittels Satz 9.2.2 können die Behauptungen für die Matrizen $\mathcal{A} = (a_{ij})$ auf die entsprechenden aus der Analysis bekannten Behauptungen für die "Komponenten" a_{ij} zurückgeführt werden. Daher geben wir die restlichen Sätze in diesem Abschnitt ohne Beweis.

Bemerkung 9.2.3. Man kann zeigen, dass auf einem endlichdimensionalen Vektorraum über $\mathbb{R}$ oder $\mathbb{C}$ alle Vektornormen äquivalen sind (Übungsaufgabe). Wenn dies bekannt ist, kann auf die Forderung der Äquivalenz zur Schurnorm verzichtet werden.

Definition 9.2.4. Es sei $\mathcal{A} \in K^{(n,n)}$ und $l \in \mathbb{N}_0$. Die Folge $|\mathcal{A}_l|_{l=0}^{\infty}$ heißt Cauchyfolge, wenn für alle $\epsilon > 0$ ein $l_0 = l_0(\epsilon)$ mit $||\mathcal{A}_{l_1} - \mathcal{A}_{l_2}|| < \epsilon$ für alle $l_1, l_2 \geq l_0$ existiert.

Satz 9.2.3. *(Cauchykriterium)*
Es sei $\mathcal{A}_l \in K^{(n,n)}$ und $l \in \mathbb{N}_0$. Dann ist $(\mathcal{A}_l)_{l=0}^{\infty}$ genau dann konvergent, wenn es eine Cauchyfolge ist.

Beweis. ohne Beweis. □

Definition 9.2.5. Es sei $\mathcal{M} \subset \mathbb{R}$, $\mathcal{M} \neq \emptyset$ und t_0 sei ein Häufungspunkt von $\mathcal{M}$. Es sei $\mathcal{A}_0 \in K^{(n,n)}$. Weiter sei $\mathcal{A}\colon \mathcal{M} \to K^{(n,n)}$, $t \to \mathcal{A}(t)$ eine matrixwertige Funktion. Wir sagen $\lim\limits_{t\to t_0} \mathcal{A}(t) = \mathcal{A}_0$, wenn für alle $\epsilon > 0$ ein $\delta = \delta(\epsilon) > 0$ existiert, für das $||\mathcal{A}(t) - \mathcal{A}_0|| < \epsilon$ für alle $t \in \mathcal{M}$ mit $|t - t_0| < \delta$ gilt.

Satz 9.2.4. *(Folgenkriterium für Grenzwerte)*
Es sei $\mathcal{M} \subset \mathbb{R}$, $\mathcal{M} \neq \emptyset$ und t_0 sei ein Häufungspunkt von $\mathcal{M}$. Es sei weiter $\mathcal{A}_0 \in K^{(n,n)}$ und $\mathcal{A}\colon \mathcal{M} \to K^{(n,n)}$, $t \to \mathcal{A}(t)$. Es gilt genau dann $\lim\limits_{t\to t_0} \mathcal{A}(t) = \mathcal{A}_0$, wenn für jede Folge $(t_l)_{l=0}^{\infty}$ mit $t_l \in \mathcal{M}$ und $\lim\limits_{l\to\infty} t_l = t_0$ dann $\lim\limits_{l\to\infty} \mathcal{A}(t_l) = \mathcal{A}(t_0)$ gilt.

Beweis. ohne Beweis. □

Aus dem Folgenkriterium ergibt sich

Satz 9.2.5. *Es sei $\mathcal{M} \subset \mathbb{R}$, $\mathcal{M} \neq \emptyset$ und t_0 sei ein Häufungspunkt von $\mathcal{M}$.*
Es sei weiter $\mathcal{A}\colon \mathcal{M} \to K^{(n,n)}$, $t \to \mathcal{A}(t) = (a_{ij}(t))_{1\le i,j\le n}$ und $\mathcal{A}_0 = (a_{ij}^{(0)})_{1\le i,j\le n}$. Es gilt genau dann $\lim\limits_{t\to t_0} \mathcal{A}(t) = \mathcal{A}_0$, wenn $\lim\limits_{t\to t_0} a_{ij}(t) = a_{ij}^{(0)}$ für alle $1 \le i,j \le n$ gilt.

Beweis. ohne Beweis. □

Definition 9.2.6. Es sei $\mathcal{M} \subset \mathbb{R}$, $\mathcal{M} \neq \emptyset$ und $t_0 \in \mathcal{M}$. Es heißt $\mathcal{A}\colon \mathcal{M} \to K^{(n,n)}$, $t \to \mathcal{A}(t)$ stetig in t_0, wenn für alle $\epsilon > 0$ ein $\delta = \delta(\epsilon) > 0$ mit $||\mathcal{A}(t) - \mathcal{A}(t_0)|| < \epsilon$ für alle $t \in \mathcal{M}$ mit $|t - t_0| < \delta$ existiert.

Satz 9.2.6. *(Folgenkriterium für Stetigkeit)*
Es sei $\mathcal{M} \subset \mathbb{R}$, $\mathcal{M} \neq \emptyset$ und $t_0 \in \mathcal{M}$ und $\mathcal{A}\colon \mathcal{M} \to K^{(n,n)}$, $t \to \mathcal{A}(t)$. Dann ist $\mathcal{A}$ genau dann stetig in t_0, wenn für jede Folge $(t_l)_{l=0}^{\infty}$ mit $t_l \in \mathcal{M}$ und $\lim\limits_{l\to\infty} t_l = t_0$ dann $\lim\limits_{l\to\infty} \mathcal{A}(t_l) = \mathcal{A}(t_0)$ gilt.

Beweis. ohne Beweis. □

Aus dem Folgenkriterium ergibt sich

Satz 9.2.7. *Es sei $\mathcal{M} \subset \mathbb{R}$, $\mathcal{M} \neq \emptyset$ und $t_0 \in \mathcal{M}$ sowie $\mathcal{A}\colon \mathcal{M} \to K^{(n,n)}$, $t \to \mathcal{A}(t) = (a_{ij}(t))_{1\le i,j\le n}$. Dann ist $\mathcal{A}$ genau dann in t_0 stetig, wenn a_{ij} in t_0 für $1 \le i,j \le n$ stetig ist.*

Beweis. ohne Beweis. □

Definition 9.2.7. Es sei $\mathcal{M} \subset \mathbb{R}$, $\mathcal{M} \neq \emptyset$ und $\mathcal{A}_l\colon \mathcal{M} \to K^{(n,n)}$, $t \to \mathcal{A}_l(t)$ mit $l \in \mathbb{N}_0$. Weiter sei $\mathcal{A}(t) \in K^{(n,n)}$. Die Folge matrixwertiger Funktionen $(\mathcal{A}_l(t))_{l=0}^{\infty}$ heißt auf $\mathcal{M}$ gleichmäßig konvergent gegen $\mathcal{A}(t)$, wenn für alle $\epsilon > 0$ ein $l_0 = l_0(\epsilon) > 0$ mit $||\mathcal{A}_l(t) - \mathcal{A}(t)|| < \epsilon$ für alle $l \ge l_0$ und $t \in \mathcal{M}$ existiert.

Satz 9.2.8. *(Cauchykriterium für gleichmäßige Konvergenz)*
Es sei $\mathcal{M} \subset \mathbb{R}$, $\mathcal{M} \neq \emptyset$ und $\mathcal{A}_l\colon \mathcal{M} \to K^{(n,n)}$, $t \to \mathcal{A}_l(t)$ mit $l \in \mathbb{N}_0$. Die Folge $(\mathcal{A}_l(t))_{l=0}^{\infty}$ ist auf $\mathcal{M}$ genau dann gleichmäßig konvergent, wenn für alle $\epsilon > 0$ ein $l_0 = l_0(\epsilon) > 0$ mit $||\mathcal{A}_{l_1}(t) - \mathcal{A}_{l_2}(t)|| < \epsilon$ für alle $l_1, l_2 \ge l_0$ existiert.

Beweis. ohne Beweis. $\square$

Satz 9.2.9. *Es sei $\mathcal{M} \subset \mathbb{R}$, $\mathcal{M} \neq \emptyset$ und $\mathcal{A}_l \colon \mathcal{M} \to K^{(n,n)}$, $t \to \mathcal{A}_l(t) = (a_{ij}^{(l)}(t))_{1\le i,j\le n}$ und $l \in \mathbb{N}$. Dann ist $(\mathcal{A}_l(t))_{l=0}^{\infty}$ genau dann gleichmäßig konvergent auf $\mathcal{M}$, wenn $(a_{ij}^{(l)}(t))_{l=0}^{\infty}$ gleichmäßig konvergent auf $\mathcal{M}$ für alle $1 \le i,j \le n$ ist.*

Beweis. ohne Beweis. $\square$

Satz 9.2.10. *Es sei $\mathcal{M} \subset \mathbb{R}$, $\mathcal{M} \neq \emptyset$ und $\mathcal{A}_l \colon \mathcal{M} \to K^{(n,n)}$, $t \to \mathcal{A}_l(t)$ sowie $\mathcal{A} \colon \mathcal{M} \to K^{(n,n)}$, $t \to \mathcal{A}(t)$. Die $\mathcal{A}_l$ seien auf $\mathcal{M}$ stetig und $\mathcal{A}_l(t)$ konvergiere auf $\mathcal{M}$ gleichmäßig gegen $\mathcal{A}(t)$. Dann ist $\mathcal{A}(t)$ auf $\mathcal{M}$ stetig.*

Beweis. ohne Beweis. $\square$

Definition 9.2.8. i) Es sei $I \subset \mathbb{R}$ ein Intervall und $\mathcal{A} \colon I \to K^{(n,n)}$, $t \to \mathcal{A}(t)$ eine matrixwertige Funktion auf I. Dann heißt $\mathcal{A}$ in $t_0 \in I$ <u>differenzierbar</u> mit <u>Ableitung</u> $\mathcal{A}'(t_0)$, falls

$$\lim_{h\to 0} \frac{\mathcal{A}(t_0+h)-\mathcal{A}(t_0)}{h} = \mathcal{A}'(t_0)$$

gilt.

ii) Eine vektorwertige Funktion $\vec{x} \colon I \to K^{(n,n)}$, $t \to \vec{x}(t)$ heißt in $t_0 \in I$ <u>differenzierbar</u> mit <u>Ableitung</u> $\vec{x}'(t_0)$, falls

$$\lim_{h\to 0} \frac{\vec{x}(t_0+h)-\vec{x}(t_0)}{h} = \vec{x}'(t_00)$$

gilt.

Satz 9.2.11. *Es sei $I \subset \mathbb{R}$ ein Intervall und $\mathcal{A} \colon I \to K^{(n,n)}$, $t \to \mathcal{A}(t) = (a_{ij}(t))_{1\le i,j\le n}$.*

i) Dann ist $\mathcal{A}$ in t_0 genau dann differenzierbar mit Ableitung $\mathcal{A}'(t)_0$, wenn die a_{ij} für $1 \le i,j \le n$ in t_0 differenzierbar mit Ableitungen $a'_{ij}(t_0)$ sind. Es ist dann $\mathcal{A}'(t_0) = (a'_{ij}(t_0))_{1\le i,j\le n}$.

ii) Es ist $\vec{x}(t) = (x_1(t), \ldots, x_n(t))^T$ in t_0 genau dann differenzierbar mit Ableitung $\vec{x}'(t)_0$, wenn die x_i für $1 \le i \le n$ in t_0 differenzierbar mit Ableitungen $x'_i(t_0)$ sind.
Es ist dann $\vec{x}'(t_0) = (x'_1(t_0), \ldots, x'_n(t_0))^T$.

Beweis. ohne Beweis. $\square$

Satz 9.2.12. *(Differentiationsregeln)*
Es sei $I \subset \mathbb{R}$ ein Intervall und $\mathcal{A}, \mathcal{B} \colon I \to K^{(n,n)}$, $\vec{x} \colon I \to K^n$ sowie $f \colon I \to K$ seien in t_0 differenzierbar. Dann sind auch $\mathcal{A}+\mathcal{B}$, $\mathcal{A}\mathcal{B}$, $\mathcal{A}\vec{x}$ und $f\mathcal{A}$ differenzierbar, und es gilt

i) $(\mathcal{A}+\mathcal{B})'(t_0) = \mathcal{A}'(t_0) + \mathcal{B}'(t_0)$

ii) $(\mathcal{A}\mathcal{B})'(t_0) = \mathcal{A}'(t_0)\mathcal{B}(t_0) + \mathcal{A}(t_0)\mathcal{B}'(t_0)$

iii) $(\mathcal{A}\vec{x})'(t_0) = \mathcal{A}'(t_0)\vec{x}(t_0) + \mathcal{A}(t_0)\vec{x}'(t_0)$

iv) $(f\mathcal{A})'(t_0) = f'(t_0)\mathcal{A}(t_0) + f(t_0)\mathcal{A}'(t_0)$

Beweis. ohne Beweis. $\square$

Satz 9.2.13. *Es sei $I = [a, b]$ mit $a < b \in \mathbb{R}$ ein kompaktes Intervall.*

i) Es seien $\mathcal{A}_l \colon I \to K^{(n,n)}$, $t \to \mathcal{A}_l(t)$ mit $l \in \mathbb{N}_0$. Die $\mathcal{A}_l$ seien differenzierbar, und die Folge $(\mathcal{A}_l)_{l=0}^{\infty}$ konvergiere für ein $t_0 \in I$ und die Folge $(\mathcal{A}'_l)_{l=0}^{\infty}$ konvergiere auf I gleichmäßig. Dann konvergiert $(\mathcal{A}_l)_{l=0}^{\infty}$ auf I gleichmäßig gegen eine auf I differenzierbare Funktion $\mathcal{A}$, und es gilt

$$\mathcal{A}'(t) = \lim_{l\to\infty} \mathcal{A}'_l(t)$$

für alle $t \in I$.

ii) Entsprechendes gilt für Funktionen $\vec{x}_l \colon I \to K^n$, $t \to \vec{x}_l(t)$.

Beweis. ohne Beweis. □

Definition 9.2.9. Es sei $I = [a, b]$ mit $a < b \in \mathbb{R}$ ein kompaktes Intervall.

i) Es sei $\mathcal{A} \colon I \to K^{(n,n)}$, $t \to \mathcal{A}(t) = (a_{ij}(t))_{1\le i,j\le n}$, und die Integrale $\int_a^b a_{ij}(t)\,dt$ mögen für $1 \le i, j \le n$ existieren. Dann setzen wir $\int_a^b \mathcal{A}(t)\,dt = \left(\int_a^b a_{ij}(t)\,dt\right)_{1\le i,j\le n}$.

ii) Es sei $\vec{x} \colon I \to K^n$, $t \to \vec{x}(t) = (x_1(t), \ldots, x_n(t))^T_{1\le i,j\le n}$, und die Integrale $\int_a^b x_i(t)\,dt$ mögen für $1 \le i \le n$ existieren. Dann setzen wir $\int_a^b \vec{x}(t)\,dt = \left(\int_a^b x_1(t)\,dt, \ldots, \int_a^b x_n(t)\,dt\right)^T$.

Satz 9.2.14. *Es sei $I = [a, b]$ mit $a < b \in \mathbb{R}$ ein kompaktes Intervall.*

i) Es seien $\mathcal{A}_l \colon I \to K^{(n,n)}$, $t \to \mathcal{A}_l(t)$ mit $l \in \mathbb{N}_0$. Die $\mathcal{A}_l$ seien integrierbar, und die Folge $(\mathcal{A}_l)_{l=0}^{\infty}$ konvergiere auf I gegen $\mathcal{A}$ gleichmäßig. Dann ist auch $\mathcal{A}$ auf I integrierbar, und es gilt

$$\int_a^b \mathcal{A}(t)\,dt = \lim_{l\to\infty} \int_a^b \mathcal{A}_l(t)\,dt.$$

ii) Entsprechendes gilt für Funktionen $\vec{x}_l \colon I \to K^n$, $t \to \vec{x}_l(t)$.

Beweis. ohne Beweis. □

Definition 9.2.10. Es sei $\mathcal{A}_l \in K^{(n,n)}$ mit $l \in \mathbb{N}_0$. Unter der unendlichen Reihe $\sum_{l=0}^{\infty} \mathcal{A}_l$ versteht man die Folge der Partialsummen $\left(\sum_{l=0}^{N} \mathcal{A}_l\right)_{N=0}^{\infty}$.

9.3 Matrixexponentialfunktion

Satz 9.3.1. *Es sei $\mathcal{A} \in K^{(n,n)}$. Dann konvergiert die unenedliche Reihe*

$$e^{\mathcal{A}} := \sum_{l=0}^{\infty} \frac{\mathcal{A}^l}{l!}.$$

Es ist $||e^{\mathcal{A}}|| \le e^{||\mathcal{A}||}$.

Beweis. Aus der Vorlesung Analysis ist die Konvergenz der unendlichen Reihe $e^{||\mathcal{A}||} := \sum_{l=0}^{\infty} \frac{||\mathcal{A}||^l}{l!}$ bekannt.
Es sei $\epsilon > 0$. Nach dem Cauchykriterium (Satz 9.2.3) gibt es $l_0 \in \mathbb{N}$, so dass

$$\sum_{l=l_0}^{l_0+m} \frac{||\mathcal{A}||_S^l}{l!} < \epsilon$$

für alle $m \in \mathbb{N}_0$ gilt. Wegen der Submultiplikativität von $||\cdot||_S$ und der Dreiecksungleichung folgt dann

$$\left|\left|\sum_{l=l_0}^{l_0+m} \frac{\mathcal{A}_S^l}{l!}\right|\right| < \epsilon$$

für alle $m \in \mathbb{N}_0$. Also ist die Folge $\left(\sum_{l=0}^{N} \frac{\mathcal{A}^l}{l!}\right)_{N=0}^{\infty}$ eine Cauchyfolge und damit nach dem Cauchykriterium konvergent. Wegen

$$\left|\left|\sum_{l=0}^{N} \frac{\mathcal{A}_S^l}{l!}\right|\right|_S \leq \sum_{l=0}^{N} \frac{||\mathcal{A}||^l}{l!} \leq e^{||\mathcal{A}||_S}$$

folgt $||e^{\mathcal{A}}|| \leq e^{||\mathcal{A}||}$. □

Definition 9.3.1. Die Matrix $e^{\mathcal{A}}$ in Satz 9.3.1 heißt Matrixexponentialfunktion von $\mathcal{A}$ (Schreibweise: $e^{\mathcal{A}}$ oder $\exp(\mathcal{A})$).

Satz 9.3.2. *Es seien $\mathcal{A}, \mathcal{B} \in K^{(n,n)}$ mit $\mathcal{AB} = \mathcal{BA}$. Dann ist $e^{\mathcal{A}+\mathcal{B}} = e^{\mathcal{A}} \cdot e^{\mathcal{B}}$.*

Beweis. Es sei $\epsilon > 0$ gegeben. Dann gibt es $N = N(\epsilon)$, so dass

$$\sum_{k=N+1}^{\infty} \frac{||\mathcal{A}||_S^k}{k!} < \epsilon \quad \text{und} \quad \sum_{l=N+1}^{\infty} \frac{||\mathcal{B}||_S^l}{l!} < \epsilon. \tag{1}$$

Es ist

$$\sum_{k=0}^{2N} \frac{\mathcal{A}^k}{k!} \sum_{l=0}^{2N} \frac{\mathcal{A}^l}{l!} = \sum_{m=0}^{2N} \frac{1}{m!} \sum_{k=0}^{m} \frac{m!}{k!(m-k)!} \mathcal{A}^k \mathcal{B}^{m-k} + \Sigma' \tag{2}$$

mit

$$\Sigma' = \sum_{\substack{1 \leq k,l \leq 2N \\ 2N+1 \leq k+l \leq 4N}} \frac{\mathcal{A}^k}{k!} \frac{\mathcal{B}^l}{l!}.$$

Wegen (1) ist

$$||\Sigma'||_S < \epsilon \cdot \left(e^{||\mathcal{A}||_S} + e^{||\mathcal{B}||_S}\right). \tag{3}$$

Wegen $\mathcal{AB} = \mathcal{BA}$ gilt der Binomische Lehrsatz

$$(\mathcal{A} + \mathcal{B})^m = \sum_{k=0}^{m} \frac{m!}{k!(m-k)!} \mathcal{A}^k \mathcal{B}^{m-k}. \tag{4}$$

Aus (2), (3) und (4) folgt

$$e^{\mathcal{A}} \cdot e^{\mathcal{B}} = \lim_{N\to\infty} \left(\sum_{k=0}^{N} \frac{\mathcal{A}^k}{k!}\right) \cdot \left(\sum_{l=0}^{N} \frac{\mathcal{B}^l}{l!}\right) = \lim_{N\to\infty} \sum_{m=0}^{N} \frac{1}{m!} (\mathcal{A} + \mathcal{B})^m = e^{\mathcal{A}+\mathcal{B}}.$$

□

Satz 9.3.3. *Für $\mathcal{A} \in K^{(n,n)}$ ist $e^{\mathcal{A}} \in GL(n,K)$. Es ist $\left(e^{\mathcal{A}}\right)^{-1} = e^{-\mathcal{A}}$.*

Beweis. Dies folgt sofort aus Satz 9.3.2 wegen $\exp(0^{(n,n)}) = \mathcal{E}_n$. □

Satz 9.3.4. *Es sei $\mathcal{A} \in K^{(n,n)}$ und $\mathcal{X} \in GL(n,K)$. Dann ist $\mathcal{X}^{-1}\exp(\mathcal{A})\mathcal{X} = \exp(\mathcal{X}^{-1}\mathcal{A}\mathcal{X})$.*

Beweis. Es ist $(\mathcal{X}^{-1}\mathcal{A}\mathcal{X})^l = \mathcal{X}^{-1}\mathcal{A}^l\mathcal{X}$. Daher gilt

$$\mathcal{X}^{-1}\exp(\mathcal{A})\mathcal{X} = \lim_{N\to\infty} \mathcal{X}^{-1}\left(\sum_{l=0}^{N}\frac{\mathcal{A}^l}{l!}\right)\mathcal{X} = \lim_{N\to\infty}\sum_{l=0}^{N}\frac{1}{l!}(\mathcal{X}^{-1}\mathcal{A}\mathcal{X})) = \exp(\mathcal{X}^{-1}\mathcal{A}\mathcal{X}).$$

□

Satz 9.3.4 ermöglicht in vielen Fällen eine einfache Berechnung der Matrixexponentialfunktion. Man wählt $\mathcal{X}$ so, dass $\mathcal{X}^{-1}\mathcal{A}\mathcal{X}$ einfache Gestalt hat: Diagonalgestalt oder Jordansche Normalform. Wie der folgende Satz zeigt, ist in diesen Fällen die Berechnung der Matrixexponentialfunktion sehr einfach.

Satz 9.3.5. *i) Es sei $\mathcal{A} \in K^{(n,n)}$, und $\mathcal{A}$ habe die Kästchengestalt*

$$\mathcal{A} = \begin{pmatrix} \mathcal{M}_1 & & 0 \\ & \ddots & \\ 0 & & \mathcal{M}_r \end{pmatrix}$$

mit quadratischen Teilmatrizen $\mathcal{M}_1, \ldots, \mathcal{M}_r$. Dann ist

$$e^{\mathcal{A}} = \begin{pmatrix} e^{\mathcal{M}_1} & & 0 \\ & \ddots & \\ 0 & & e^{\mathcal{M}_r} \end{pmatrix}.$$

ii) Für eine Diagonalmatrix

$$D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \quad \textit{ist} \quad e^D = \begin{pmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n} \end{pmatrix}$$

mit $\lambda_i \in K$.

iii) Es sei

$$N = \begin{pmatrix} 0 & & & & \\ \alpha & 0 & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & \alpha & 0 \end{pmatrix} \in K^{(n,n)}$$

mit $\alpha \in K$. Dann ist

$$e^N = \begin{pmatrix} 1 & & & & \\ \alpha & 1 & & & \\ \frac{\alpha^2}{2!} & \ddots & \ddots & & \\ \vdots & \ddots & \ddots & \ddots & \\ \frac{\alpha^{n-1}}{(n-1)!} & \cdots & \frac{\alpha^2}{2!} & \alpha & 1 \end{pmatrix}.$$

Beweis. Durch Nachrechnen. □

9.4 Systeme von linearen Differentialgleichungen

Eine Differentialgleichung (DGL) stellt eine Beziehung zwischen einer Funktion und ihrer ersten Ableitung oder auch höheren Ableitungen her. Ein System von Differentialgleichungen stellt solche Beziehung zwischen mehreren Funktionen und ihren Ableitungen her.
Eine der bekanntesten Differentialgleichungen ist

$$y' = y \tag{1}$$

mit der Lösung

$$y(t) = ce^t, \tag{2}$$

wobei C eine beliebige Konstante ist. Die Lösung (2) gilt auf dem Lösungsintervall $I = (-\infty, \infty)$. Es kann vorkommen, dass das Lösungsintervall nur ein endliches Intervall ist. Die Konstante C in (2) ist eindeutig bestimmt, wenn (1) durch eine Anfangsbedingung zu einem Anfangswertproblem (AWP) erweitert wird. Das AWP

$$y' = y \quad \text{und} \quad y(0) = c_0 \text{ (Anfangsbedingung)} \tag{3}$$

hat die eindeutig bestimmte Lösung $y(t) = c_0e^t$. Das Problem (3) kann auf viele Weisen verallgemeinert werden. Wir betrachten hier Systeme linearer DGL mit konstanten Koeffizienten.

Definition 9.4.1. Ein System linearer DGL erster Ordnung mit konstanten Koeffizienten hat die Form

$$\begin{aligned} y_1' &= a_{11}y_1 + \ldots + a_{1n}y_n + f_1 \\ y_2' &= a_{21}y_1 + \ldots + a_{2n}y_n + f_2 \\ &\vdots \\ y_n' &= a_{n1}y_1 + \ldots + a_{nn}y_n + f_n \end{aligned}$$

oder in Matrixschreibweise

$$\vec{y}\,' = \mathcal{A}\vec{y} + \vec{f} \tag{4}$$

mit den Vektoren $\vec{y}\,' = (y_1', \ldots, y_n')^T$, $\vec{y} = (y_1, \ldots, y_n)^T$ sowie $\vec{f} = (f_1, \ldots, f_n)^T$ und der Matrix $\mathcal{A} = (a_{ij})_{1\leq i,j\leq n}$. Dabei seien die $a_{ij} \in K$, und $f\colon I \to K^n$ sei stetig.
Ferner seien die Anfangsbedingungen vorgeschrieben: $\vec{y}(t_0) = \vec{y}^{(0)}$ für ein $t_0 \in I$.
Das System (4) heißt homogen, falls $\vec{f} = \vec{0}$, ansonsten inhomogen.

Satz 9.4.1. *Es sei $I \subset \mathbb{R}$ ein Intervall, und $\mathcal{A}\colon I \to K^{(n,n)}$, $t \to \mathcal{A}(t)$ sei auf I stetig differenzierbar. Dann ist auch $f\colon t \to e^{\mathcal{A}(t)}$ auf I differenzierbar mit der Ableitung $f'(t) = \mathcal{A}'(t)e^{\mathcal{A}(t)}$.*

Beweis. Es sei

$$\mathcal{F}(t, N) = \sum_{k=0}^{N} \frac{\mathcal{A}(t)^k}{k!}$$

mit $N \in \mathbb{N}_0$. Dann ist

$$\mathcal{F}'(t, N) = \sum_{k=1}^{N} \mathcal{A}'(t)\frac{\mathcal{A}(t)^{k-1}}{(k-1)!}$$

nach Satz 9.2.12.

Es sei $J \subset I$ ein kompaktes Teilintervall von I und $M_1 = \sup\{||\mathcal{A}'(t)||_S \exp(||\mathcal{A}(t)||_S) \colon t \in J\}$ bzw. $M_2 = \sup\{||\mathcal{A}(t)||_S \colon t \in J\}$. Dann ist

$$\left|\left|\sum_{k=N+1}^{\infty} \mathcal{A}'(t)\frac{\mathcal{A}(t)^{k-1}}{(k-1)!}\right|\right|_S \leq \frac{M_2^k}{(k-1)!}M_1.$$

Damit erfüllt die Reihe $\sum_{k=1}^{\infty} \mathcal{A}'(t)\frac{\mathcal{A}(t)^{k-1}}{(k-1)!}$ das Cauchykriterium für gleichmäßige Konvergenz (Satz 9.2.8). Somit konvergiert $\mathcal{F}'(t,N)$ gleichmäßig gegen

$$\sum_{k=1}^{\infty} \mathcal{A}'(t)\frac{\mathcal{A}(t)^{k-1}}{(k-1)!} = \mathcal{A}'(t)e^{\mathcal{A}(t)}.$$

Wegen $\lim_{N\to\infty} \mathcal{F}(t,N) = e^{\mathcal{A}(t)}$ für alle $t \in J$ folgt nach Satz 9.2.13

$$\frac{d}{dt}e^{\mathcal{A}(t)} = \mathcal{A}'(t)e^{\mathcal{A}(t)}$$

für alle $t \in J$. Da J ein beliebiges endliches Teilintervall von I ist, folgt die Behauptung. □

Satz 9.4.2. *Das homogene Anfangswertproblem*

$$\vec{y}\,' = \mathcal{A}\vec{y}, \quad \vec{y}(0) = \vec{y}^{(0)} \tag{1}$$

besitzt genau eine Lösung auf $(-\infty, \infty)$*, nämlich*

$$\vec{y}(t) = e^{\mathcal{A}t}\vec{y}^{(0)}. \tag{2}$$

Beweis. Existenz:
Die Lösung (2) löst das Problem nach Satz 9.4.1.

Eindeutigkeit:
Es sei $\vec{z}(t)$ eine Lösung von (1). Dann gilt

$$\frac{d}{dt}\left(e^{-\mathcal{A}t}\vec{z}(t)\right) = e^{-\mathcal{A}t}\vec{z}\,'(t) - \mathcal{A}e^{-\mathcal{A}t}\vec{z}(t) = e^{-\mathcal{A}t}\vec{z}\,'(t) - e^{-\mathcal{A}t}\mathcal{A}\vec{z}(t) = e^{-\mathcal{A}t}\vec{z}\,'(t) - e^{-\mathcal{A}t}\vec{z}\,'(t) = 0.$$

Deshalb ist $e^{-\mathcal{A}t}\vec{z}(t)$ konstant, d.h. $\vec{z}(t) = e^{\mathcal{A}t}\vec{c}$ mit $\vec{c} \in K^n$. Aus $\vec{z}(0) = \vec{y}^{(0)}$ ergibt sich $\vec{c} = \vec{y}^{(0)}$. □

Satz 9.4.3. *Es sei* $f \colon [0,a] \to K^n$ *stetig mit* $a > 0$ *und* $\mathcal{A} \in K^{(n,n)}$*. Dann besitzt das Anfangswertproblem*

$$\vec{y}\,'(t) = \mathcal{A}\vec{y}(t) + \vec{f}(t), \quad \vec{y}(0) = \vec{y}^{(1)} \tag{1}$$

genau eine Lösung $\vec{y}$*, nämlich*

$$\vec{y}(t) = e^{\mathcal{A}t}\vec{y}^{(0)} + \int_0^t e^{\mathcal{A}(t-s)}\vec{f}(s)\,ds. \tag{2}$$

Beweis. Existenz:
Nach dem Hauptsatz der Differential- und Integralrechnung ist

$$\frac{d}{dt}\int_0^t e^{\mathcal{A}(t-s)}\vec{f}(s)\,ds = e^{\mathcal{A}(t-s)}\vec{f}(s)\big|_{s=t} = \vec{f}(t).$$

Nach Satz 9.4.2 folgt, dass (2) eine Lösung von (1) ist.

Eindeutigkeit:
Es seien $\vec{y}_1(t)$ und $\vec{y}_2(t)$ Lösungen von (1). Dann löst $\vec{u}(t) = \vec{y}_1(t) - \vec{y}_2(t)$ das homogene Problem $\vec{u}\,'(t) = \mathcal{A}\vec{u}(t)$ mit $\vec{u}(0) = \vec{0}$. Nach Satz 9.4.2 ist $\vec{u}\,'(t) = 0$. □

Beispiel 9.4.1. Man finde die Lösung des AWP

$$\vec{y}\,' = \begin{pmatrix} 9 & -7 & 0 & 2 \\ 7 & -5 & 0 & 2 \\ 4 & -4 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \cdot \vec{y} \quad \text{und} \quad \vec{y}(0) = \begin{pmatrix} 3 \\ 2 \\ -1 \\ 0 \end{pmatrix}.$$

Lösung:
Wir bestimmen die Jordansche Normalform der Koeffizientenmatrix.
1. Schritt:
Wir berechnen das charakteristische Polynom der Koeffizientenmatrix

$$\mathcal{A} = \begin{pmatrix} 9 & -7 & 0 & 2 \\ 7 & -5 & 0 & 2 \\ 4 & -4 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Es ist

$$\det(\mathcal{A} - \lambda\mathcal{E}_4) = (2-\lambda)^2 \cdot ((9-\lambda)\cdot(-5-\lambda) + 49) = (\lambda - 2)^4.$$

2. Schritt:
Berechnung des Minimalpolynoms:
Es ist

$$\mathcal{A} - 2\mathcal{E}_4 = \begin{pmatrix} 7 & -7 & 0 & 2 \\ 7 & -7 & 0 & 2 \\ 4 & -4 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$
$$(\mathcal{A} - 2\mathcal{E}_4)^2 = 0^{(4,4)},$$

weswegen $m(x) = (x-2)^2$ ist. Die Matrix $\mathcal{A}$ ist somit nicht diagonalisierbar, sonst wäre $m(x) = x-2$. Die Jordansche Normalform hat somit die Form

$$J = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \quad \text{oder} \quad J = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix}.$$

Der Eigenraum U_2 ergibt sich als Lösung des LGS $(\mathcal{A} - 2\mathcal{E}_4)\vec{x} = \vec{0}$. Wegen $\mathrm{rg}(\mathcal{A} - 2\mathcal{E}_4) = 2$ ist $\dim U_2 = 2$. Also liegt der zweite Fall mit

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

vor.
3. Schritt:
Wir suchen eine Basis $\mathcal{B}$, so dass die zu $\varphi_{\mathcal{A}}\colon \vec{x} \to \mathcal{A}\vec{x}$ gehörige Darstellungsmatrix $\mathcal{M}(\varphi, \mathcal{B}, \mathcal{B})$ Jordansche Normalform hat. Nach Satz 8.5.2 erhält man eine solche Basis wie folgt: es sei $\vec{x}_1$ ein Vektor mit

Minimalpolynom $m(x) = (x-2)^2$. Es sei $\vec{x}_2 = (\mathcal{A} - 2\mathcal{E}_4)\vec{x}_1$. Es sei $\vec{x}_3$ ein Vektor mit Minimalpolynom $m(x) = (x-2)^2$ und $\vec{x}_3 \notin \langle \vec{x}_1, \vec{x}_2 \rangle$. Wir setzen $\vec{x}_4 = (\mathcal{A} - 2\mathcal{E}_4)\vec{x}_3$. Dann kann $\mathcal{B} = \{\vec{x}_1, \vec{x}_2, \vec{x}_3, \vec{x}_4\}$ gewählt werden. Wir beginnen mit $\vec{x}_1 = \vec{e}_1$ und erhalten $\vec{x}_2 = (\mathcal{A} - 2\mathcal{E}_4)\vec{x}_1 = (7, 7, 4, 0)^T$. Es ist $\vec{e}_4 \notin \langle \vec{x}_1, \vec{x}_2 \rangle$. Also ist $\vec{x}_3 = \vec{e}_4$ und $\vec{x}_4 = (\mathcal{A} - 2\mathcal{E}_4)\vec{x}_3 = (2, 2, 1, 0)^T$.
Damit gilt $J = \mathcal{X}^{-1}\mathcal{A}\mathcal{X}$ mit der Matrix $\mathcal{X}$, deren Spalten die Vektoren der Basis $\mathcal{B}$ sind. Also gilt

$$\mathcal{X}^{-1}\mathcal{A}\mathcal{X} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

mit

$$\mathcal{X} = \begin{pmatrix} 1 & 7 & 0 & 2 \\ 0 & 7 & 0 & 2 \\ 0 & 4 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{und} \quad \mathcal{X}^{-1} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 4 & -7 & 0 \end{pmatrix}.$$

<u>4. Schritt:</u>
Berechnung von $\exp(\mathcal{A}t)$:
Nach Satz 9.3.4 ist

$$\exp(\mathcal{A}t) = \exp(\mathcal{X}(Jt)\mathcal{X}^{-1}) = \mathcal{X}\exp(Jt)\mathcal{X}^{-1}.$$

Nach Satz 9.3.5 ist

$$\exp(Jt) = \begin{pmatrix} e^{\mathcal{M}t} & \\ & e^{\mathcal{M}t} \end{pmatrix} \quad \text{mit} \quad \mathcal{M} = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}.$$

Es ist

$$e^{\mathcal{M}t} = \exp\left(\begin{pmatrix} 2t & 0 \\ 0 & 2t \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ t & 0 \end{pmatrix}\right) = \begin{pmatrix} e^{2t} & 0 \\ 0 & e^{2t} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} = \begin{pmatrix} e^{2t} & 0 \\ te^{2t} & 2^{2t} \end{pmatrix}.$$

Also ist

$$\exp(Jt) = \begin{pmatrix} e^{2t} & 0 & 0 & 0 \\ te^{2t} & e^{2t} & 0 & 0 \\ 0 & 0 & e^{2t} & 0 \\ 0 & 0 & te^{2t} & e^{2t} \end{pmatrix}$$

und

$$\begin{aligned} \exp(\mathcal{A}t) &= e^{2t}\begin{pmatrix} 1 & 7 & 0 & 2 \\ 0 & 7 & 0 & 2 \\ 0 & 4 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ t & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & t & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 4 & -7 & 0 \end{pmatrix} \\ &= e^{2t}\begin{pmatrix} 1+7t & -7t & 0 & 2t \\ 14t & -(7t-1) & 0 & 2t \\ 4t & -4t & 1 & t \\ 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

Die Lösung ist damit

$$\vec{y}(t) = \exp(\mathcal{A}t)\vec{y}(0) = e^{2t}\begin{pmatrix} 3+7t \\ 2+28t \\ -1+4t \\ 0 \end{pmatrix}.$$

Kapitel 10

Lineare Optimierung

10.1 Ein Beispiel zur Einführung

Eine Firma will ein neues Getränk auf den Markt bringen. Es soll aus schon vorhandenen Flüssigkeiten gemischt werden, und zwar stehen die Flüssigkeiten A, B und C zur Verfügung. Eine Flasche soll 1000g des Getränks und maximal 230 g Zucker enthalten. Der Zuckergehalt pro Gramm Flüssigkeit (in Gramm) ist wie folgt:

A	$0,25$
B	$0,15$
C	$0,05$

Die Kosten der Flüssigkeiten pro Gramm (in 10^{-3} Euro) sind ebenfalls wie folgt:

A	1
B	2
C	4

Lösung:
Der Anteil der Flüssigkeiten A und B (in Gramm) sei x bzw. y.
Zunächst wollen wir den zulässigen Bereich, die durch lineare Ungleichungen definierte Menge X, bestimmen.
Der Anteil der Flüssigkeiten A und B muss nichtnegativ sein, also haben wir

$$x \geq 0 \tag{1}$$

bzw.

$$y \geq 0. \tag{2}$$

Ferner kann der Anteil von A und B nicht größer sein, als die Gesamtmenge (1000g) des Getränks, also gilt

$$x + y \leq 1000. \tag{3}$$

Die letzte Ungleichung wird durch die vorgeschriebene obere Grenze des Zuckergehaltes geliefert:

$$0,25x + 0,15y + 0,05(1000 - x - y) \leq 230$$

oder

$$2x + y \leq 1800. \tag{4}$$

Der durch die Ungleichungen (1) bis (4) beschriebener zulässiger Bereich X ist ein zweidimensionales Polyeder.
Es ist nun der Punkt von X zu finden, in dem die Kostenfunktion

$$k(x, y) = x + 2y + 4(1000 - x - y) = -3x - 2y + 4000$$

ihr Minimum annimmt. Man kann dies auch als Maximumproblem formulieren, indem man die Zielfunktion $z(x, y) = 3x + 2y$ einführt. Es ist der Punkt von X zu finden, in dem z sein Maximum annimmt. Die Menge X hat die Ecken $E_1 = (0, 0)$, $E_2 = (900, 0)$, $E_3 = (800, 200)$ und $E_4 = (0, 1000)$. Wir werden später sehen, dass die Zielfunktion ihr Maximum stets in einer der Ecken annimmt. In unserem Fall ergibt die Ecke E_3 den maximalen Wert für z (und damit minimale Kosten).
Man sollte also 800g von Flüssigkeit A, 200g von Flüssigkeit B und nichts von Flüssigkeit C verwenden.

Bevor wir diese Überlegungen verallgemeinern, benötigen wir Definitionen der schon aufgetretenen Begriffe "Polyeder", "Ecke" und weiterer Begriffe.

10.2 Konvexe Mengen und Funktionen

Definition 10.2.1. i) Es seien $\vec{x}, \vec{y} \in \mathbb{R}^n$. Dann ist die Punktmenge

$$\mathcal{S} = \{\vec{s} \colon t\vec{x} + (1-t)\vec{y},\ 0 \le t \le 1\}$$

die Strecke zwischen $\vec{x}$ und $\vec{y}$. Weiter heißt

$$\overset{\circ}{\mathcal{S}} = \{\vec{s} \colon t\vec{x} + (1-t)\vec{y},\ 0 < t < 1\}$$

die offene Strecke oder das Innere der Strecke zwischen $\vec{x}$ und $\vec{y}$.

ii) Eine Teilmenge $K \subset \mathbb{R}^n$ heißt konvex, wenn für je zwei Punkte $\vec{x}, \vec{y} \in K$ auch die Strecke zwischen $\vec{x}$ und $\vec{y}$ in K liegt.

iii) Es sei $b \in \mathbb{R}$, $\vec{a} = (a_1, \ldots, a_n)^T \in \mathbb{R}^n - \{\vec{0}\}$. Dann heißt $H = \{\vec{x} \in \mathbb{R}^n : \vec{a}^T\vec{x} \le b\}$ ein Halbraum von $\mathbb{R}^n$.

Definition 10.2.2. i) Es sei $\vec{x}_0 \in \mathbb{R}^n$. Die Menge $U_\epsilon(\vec{x}_0) = \{\vec{x} \in \mathbb{R}^n : ||\vec{x} - \vec{x}_0|| < \epsilon\}$ nennen wir eine ϵ- Umgebung von $\vec{x}_0$.

ii) Es sei $\mathcal{M} \subset \mathbb{R}^n$. Dann heißt $\vec{x}_0 \in \mathcal{M}$ innerer Punkt von $\mathcal{M}$, falls es ein $\epsilon > 0$ mit $U_\epsilon(\vec{x}_0) \subset \mathcal{M}$ gibt. Die Menge aller inneren Punkte von $\mathcal{M}$ wird mit $\overset{\circ}{\mathcal{M}}$ bezeichnet und heißt auch das Innere von $\mathcal{M}$.

iii) Die Menge $\mathcal{M}$ heißt offen, wenn $\mathcal{M} = \overset{\circ}{\mathcal{M}}$ gilt. Weiter heißt K abgeschlossen, wenn $\mathbb{R}^n - K$ offen ist. Der Abschluss $\overline{\mathcal{M}}$ von $\mathcal{M}$ ist der Durchschnitt aller abgeschlossenen Mengen, die $\mathcal{M}$ enthalten.

iv) Es sei $\mathcal{M} \subset \mathbb{R}^n$. Ein Punkt $\vec{x}_0 \in \mathbb{R}^n$ heißt Randpunkt von $\mathcal{M}$, wenn für alle $\epsilon > 0$ die Umgebung $U_\epsilon(\vec{x}_0)$ sowohl Punkte von $\mathcal{M}$ als auch von $\mathbb{R}^n - \mathcal{M}$ enthält. Die Menge aller Randpunkte von $\mathcal{M}$ heißt Rand von $\mathcal{M}$.

Definition 10.2.3. i) Ein konvexes Polyeder der Dimension n ist der Durchschnitt endlich vieler Halbräume des $\mathbb{R}^n$, die mindestens einen inneren Punkt gemeinsam haben.

ii) Ein Polyeder P heißt beschränkt, wenn es eine positive reelle Zahl $S > 0$ gibt, so dass für alle $\vec{x} = (x_1, \dots, x_n)^T \in P$ dann $|x_i| < S$ gilt.

iii) Eine konvexe Untermenge K eines Polyeders P heißt extrem, wenn für je zwei Punkte $\vec{x}, \vec{y} \in P$ die offene Strecke genau dann K schneidet, wenn beide Punkte in K liegen.

iv) Ist P ein konvexes Polyeder und H ein Halbraum mit Randhyperebene E, so dass $P \subset H$ und $P \cap E \neq \emptyset$ gilt, so heißt $P \cap E$ eine Randseite von P.

v) Es sei K konvex. Dann heißt $\vec{x} \in K$ Extrempunkt oder Ecke von K, wenn $\vec{x}$ nicht im Inneren einer in K enthaltenen Strecke liegt.

Definition 10.2.4. Es sei $\mathcal{M} \subset \mathbb{R}^n$. Der Durchschnitt aller konvexen Mengen, die $\mathcal{M}$ enthalten, heißt konvexe Hülle von $\mathcal{M}$.

Satz 10.2.1. *Es sei $P \subset \mathbb{R}^n$ ein konvexes Polyeder der Dimension n.*

i) Die Menge P ist eine abgeschlossene Teilmenge des $\mathbb{R}^n$ Das Innere $\overset{\circ}{P}$ von P ist konvex, und P ist der Abschluss von $\overset{\circ}{P}$.

ii) Ist K eine extreme Untermenge von P und gilt $K \neq P$, so sind alle Punkte von K Randpunkt von P.

Beweis. ohne Beweis. □

Definition 10.2.5. Es sei $K \subset \mathbb{R}^n$ konvex. Eine Funktion $f\colon K \to \mathbb{R}$ heißt konvex, wenn für alle $\vec{x}, \vec{y} \in K$ und $0 < t < 1$ auch

$$f(t\vec{x} + (1-t)\vec{y}) \leq t \cdot f(\vec{x}) + (1-t) \cdot f(\vec{y})$$

gilt.

Satz 10.2.2. *Es sei $K \subset \mathbb{R}^n$ konvex und $f\colon K \to \mathbb{R}$ ebenfalls konvex. Jedes lokale Minimum von f ist ein globales Minimum. Nimmt f sein Supremum auf K an, dann auf dem Rand von K.*

Beweis. ohne Beweis. □

Satz 10.2.3. *Es sei P ein beschränktes konvexes Polyeder und $f\colon P \to \mathbb{R}$ konvex. Dann nimmt f sein Maximum in einem Extrempunkt von P an.*

Beweis. ohne Beweis. □

Satz 10.2.4. *Es sei $P \subset \mathbb{R}^n$ ein beschränktes konvexes Polyeder der Dimension n und $z\colon P \to \mathbb{R}$ linear. Dann gilt*

i) Die Abbildung z nimmt sein Maximum in (mindestens) einen Extrempunkt an.

ii) Es sei $\vec{x}^{(0)}$ ein Extrempunkt, in dem z sein Maximum nicht annimmt. Dann gibt es einen weiteren Extrempunkt $\vec{x}^{(1)}$, so dass die Strecke zwischen $\vec{x}^{(0)}$ und $\vec{x}^{(1)}$ extrem bzgl. P ist und dass $z(\vec{x}^{(1)}) \geq z(\vec{x}^{(0)})$ gilt.

iii) Zu $\vec{x}^{(0)}$ gibt es eine Folge $\vec{x}^{(1)}, \vec{x}^{(2)}, \dots, \vec{x}^{(k)}$ von Extrempunkten, so dass $\vec{x}^{(i)}$ und $\vec{x}^{(i+1)}$ jeweils extreme Strecken beranden und dass

$$z(\vec{x}^{(0)}) = z(\vec{x}^{(1)}) = \ldots = z(\vec{x}^{(k-1)}) < z(\vec{x}^{(k)}).$$

Beweis. ohne Beweis. □

10.3 Das Simplexverfahren

Wir haben in Abschnitt 10.1 bereits ein Beispiel eines linearen Optimierungsproblems kennengelernt. Für die allgemeine Behandlung führen wir zunächst eine Normierung durch:
Durch Einführung von "Schlupfvariablen" $x_{n+1}, \ldots, x_{n+M}$ wird eine Ungleichung

$$\sum_{j=1}^{n} a_{ij}x_j \leq b_i$$

mit $b_i \geq 0$ durch

$$\sum_{j=1}^{n} a_{ij}x_j + x_{n+i} = b_i$$

mit $x_{n+i} \geq 0$ sowie eine Ungleichung

$$\sum_{j=1}^{n} a_{ij}x_j \geq b_i$$

mit $b_i \geq 0$ durch

$$\sum_{j=1}^{n} a_{ij}x_j - x_{n+i} = b_i$$

mit $x_{n+i} \geq 0$ ersetzt.

Definition 10.3.1. Unter der normierten Form eines linearen Optimierungsproblems versteht man

$$\begin{aligned} \text{Max.: } z &= \vec{c}^T\vec{x} \\ \mathcal{A}\vec{x} &= \vec{b} \\ \vec{x} &\geq \vec{0}, \end{aligned}$$

wobei $\mathcal{A} \in \mathbb{R}^{(n,n)}$ und $\vec{c}, \vec{x}, \vec{b} \in \mathbb{R}^n$ sind. Die Spalten von $\mathcal{A}$ seien $\vec{a}_j \in \mathbb{R}^m$ mit $1 \leq j \leq n$.
Es sei $\operatorname{rg}\mathcal{A} = m$. Ist $\vec{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$, so versteht man unter $\vec{x} \geq \vec{0}$, dass $x_i \geq 0$ für alle $1 \leq i \leq n$ gelte.

Definition 10.3.2. Eine quadratische Teilmatrix $\mathcal{B} \in \mathbb{R}^{(m,m)}$ von $\mathcal{A}$ heißt Basismatrix von $\mathcal{A}$, wenn $\operatorname{rg}\mathcal{B} = m$ ist, ihre Spalten also linear unabhängig sind.
Es sei $\mathcal{B} = (\vec{a}_1, \ldots, \vec{a}_m)$. Dann sind die Spalten $\vec{a}_1, \ldots, \vec{a}_n$ Linearkombinationen von $\vec{a}_1, \ldots, \vec{a}_m$. Wir verwenden die Bezeichung

$$\vec{a}_j = \sum_{i=1}^{m} y_{ij}\vec{a}_i$$

mit $j = 1, \ldots, n$ oder $\vec{a}_j = \mathcal{B}\vec{y}_j$ mit $\vec{y}_j = (y_{1j}, \ldots, \vec{y}_{mj})^T$. Dabei ist $y_{ij} = \delta_{ij}$ für $1 \leq j \leq m$.
Somit besteht $\mathcal{A}$ aus den zwei Teilmatrizen $\mathcal{A} = (\mathcal{B}, N)$ mit $N = (\vec{a}_{m+1}, \ldots, \vec{a}_n)$.
Für $\vec{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ sei $\vec{x}_\mathcal{B} = (x_1, \ldots, x_m)^T$ und $\vec{x}_N = (x_{m+1}, \ldots, x_n)^T$.
Eine Lösung von $\mathcal{A}\vec{x} = \vec{b}$ heißt Basislösung, wenn $\mathcal{B}\vec{x}_\mathcal{B} = \vec{b}$ und $\vec{x}_N = \vec{0}$ gilt. Die Komponenten von $\vec{x}_\mathcal{B}$ heißen Basisvariable.
Entsprechend wird $\vec{c}_\mathcal{B}^T$ in $\vec{c}^T = (\vec{c}_\mathcal{B}^T, \vec{c}_N^T)$ zerlegt, wobei $\vec{c}_\mathcal{B}^T = (c_{\mathcal{B}_1}, \ldots, c_{\mathcal{B}_m})$ gilt.
Für die Basislösung $\vec{x}$ gilt $z = \vec{c}^T\vec{x} = \vec{c}_\mathcal{B}^T\vec{x}_\mathcal{B}$.
Wir setzen $z_j = \vec{c}_\mathcal{B}^T\vec{y}_j$.
Eine Basislösung (bzw. die zugehörige Ecke) heißt entartet, falls $z_j - c_j < 0$ und falls

$$\min_i \left\{ \frac{x_i}{y_{ij}} : y_{ij} > 0 \right\} = 0$$

ist.
In der Folge nehmen wir an, dass keine entarteten Ecken existieren.

Wir kommen nun zur Beschreibung des Simplexverfahrens:

Wir bestimmen rekursiv eine Folge von Basislösungen, die durch ihre Basismatrizen $\mathcal{B}^{(0)}, \mathcal{B}^{(1)}, \ldots$ bestimmt sind.

Start $\mathcal{B}^{(0)}$:
Bestimme eine beliebige Basislösung. Durch Umnummerierung der Variablen kann erreicht werden, dass $\vec{x} = (\vec{x}_{\mathcal{B}}, \vec{0})$ mit $\vec{x}_{\mathcal{B}} = (x_1, \ldots, x_m)$ gilt.
Schritt $\mathcal{B}^{(p)} \to \mathcal{B}^{(p+1)}$:
Prüfe, ob $\mathcal{B}^{(p)}$ das untenstehende Abbruchkriterium erfüllt. Wenn nicht, dann ist die neue Basismatrix wie folgt zu konstruieren:
Bestimme ein Paar (r, j) mit

$$1 \le r \le m \quad \text{und} \quad m \le j \le n, \tag{$*$}$$

für das

$$z_j - c_j < 0 \tag{$**$}$$

und

$$y_{rj} > 0 \quad \text{bzw.} \quad x_r \neq 0 \tag{$***$}$$

gilt. Der Wert $\frac{x_r}{y_{rj}}(z_j - c_j)$ ist unter den Bedingungen $(*)$, $(**)$ und $(***)$ minimal.
Die neue Basismatrix $\mathcal{B}^{(p+1)}$ entsteht durch Ersetzen der Spalte $\vec{a}_r$ von $\mathcal{B}^{(p)}$ durch $\vec{a}_j$.

Abbruchkriterium:
Fall 1:
Gilt $z_j - c_j \ge 0$ für $m < j \le n$ für die zu $\mathcal{B}^{(p)}$ gehörige Basislösung, so nimmt die Zielfunktion z für diese Basislösung ihr Maximum an.
Fall 2:
Gilt $z_j - c_j < 0$ für ein j mit $m < j \le n$ und ist $y_{ij} \le 0$ für alle $i \le m$, so besitzt das lineare Optimierungsproblem keine Lösung. Die Zielfunktion ist nicht nach oben beschränkt.

Beispiel 10.3.1. Ein Landwirt möchte 50ha Land bebauen, und zwar mit Weizen, Rüben oder Mais. Die Anbaufläche für Rüben soll 20ha nicht überschreiten. Es kann maximal Arbeit für 1300 Stunden verrichtet werden. Arbeitsstunden und Gewinn pro Hektar sind für die drei Produkte in folgender Tabelle zusammengefasst:

	Arbeitszeit in Stunden/ ha	Gewinn in Euro/ ha
Weizen	20	5000
Rüben	40	8000
Mais	30	6000

Welche Kombination bringt den höchsten Gewinn?

Lösung:
Es seien x_1, x_2 bzw x_3 die Anzahl in Hektar der Anbauflächen für Weizen, Rüben bzw. Mais.
Mit $\vec{x} = (x_1, x_2, x_3, x_4, x_5)^T$, $\vec{b} = (50, 130, 20)^T$, $\vec{c} = (5, 8, 6, 0, 0)^T$ und

$$\mathcal{A} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 4 & 3 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

erhalten wir dann das normierte Problem

$$\begin{aligned} \text{Max.: } z &= \vec{c}^T \vec{x} \\ \mathcal{A}\vec{x} &= \vec{b} \\ \vec{x} &\ge \vec{0}. \end{aligned}$$

Für $\mathcal{B}^{(0)}$ setzen wir

$$\mathcal{B}^{(0)} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

mit der Basislösung $\vec{x} = (50, 0, 0, 30, 20)^T$.
Für den Schritt von $\mathcal{B}^{(0)}$ auf $\mathcal{B}^{(1)}$ erhalten wir $\vec{y}_4$ und $\vec{y}_5$ als Lösungen von

$$\mathcal{B}^{(0)}\vec{y}_4 = \vec{a}_4 = \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix} \quad \text{bzw.} \quad \mathcal{B}^{(0)}\vec{y}_5 = \vec{a}_5 = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix},$$

also $\vec{y}_4 = (1, 2, 0)^T$ und $\vec{y}_5 = (1, 1, -1)^T$. Es ist $\vec{c}_{\mathcal{B}}^T = (5, 0, 0)^T$ und

$$z_4 = (5, 0, 0) \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = 5 \quad \text{bzw.} \quad z_5 = (5, 0, 0) \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = 5.$$

Also ist $z_4 - c_4 = 5 - 8 < 0$ und $z_5 - c_5 = 5 - 6 < 0$.
Für die Werte r mit $x_r \neq 0$ ergibt sich

$$\begin{array}{lll} j = 4: & \dfrac{x_1}{y_{4,1}} = \dfrac{50}{1} = 50, & \dfrac{x_2}{y_{4,2}} = \dfrac{30}{2} = 15 \\ j = 5: & \dfrac{x_1}{y_{5,1}} = \dfrac{50}{1} = 50, & \dfrac{x_2}{y_{5,2}} = \dfrac{30}{1} = 30. \end{array}$$

Also entsteht $\mathcal{B}^{(1)}$ durch Ersetzen der Spalte $\vec{a}_2$ von $\mathcal{B}^{(0)}$ durch $\vec{a}_4$. Somit ist

$$\mathcal{B}^{(1)} = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 4 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Vor dem Schritt von $\mathcal{B}^{(1)}$ auf $\mathcal{B}^{(2)}$ überprüfen wir das Abbruchkriterium:
Zu $\mathcal{B}^{(1)}$ gehört die Basislösung $\vec{x}_1 = (35, 15, 0, 0, 5)^T$. Lösungen von

$$\mathcal{B}^{(1)}\vec{y}_4 = \vec{a}_4 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{bzw.} \quad \mathcal{B}^{(1)}\vec{y}_5 = \vec{a}_5 = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}$$

sind $\vec{y}_4 = (-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2})^T$ und $\vec{y}_5 = (\frac{1}{2}, \frac{1}{2}, -\frac{1}{2})^T$. Es ist $\vec{c}_{\mathcal{B}}^T = (5, 8, 0)^T$ und

$$z_4 = (5, 8, 0) \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix} = \frac{3}{2} \quad \text{bzw.} \quad z_5 = (5, 8, 0) \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix} = \frac{13}{2}$$

mit $c_4 = 0$ und $c_5 = 0$. Damit ist $z_4 - c_4 > 0$ und $z_5 - c_5 > 0$.
Damit ist Fall 1 des Abbruchkriteriums erfüllt. Also nimmt z für die zu $\mathcal{B}^{(1)}$ gehörige Basislösung ihr Maximum an.
Der Landwirt sollte Weizen auf 35ha, Rüben auf 15ha und keinen Mais anbauen.