Skript zur Vorlesung

Lineare Algebra für die Sekundarstufe I (L2/L5)

Sommersemester 2017

Martin Möller und Jürgen Wolfart

Frankfurt am Main, 10. Juli 2017

Inhaltsverzeichnis

Einleitung

Ein primäres Ziel der Linearen Algebra ist das Lösen linearer Gleichungssysteme. Zu einem gegebenen solchen Gleichungssystem will man zuerst wissen, ob es lösbar ist und wenn ja, wie man eine Lösung findet. Aber auch die Frage, ob es mehr als eine Lösung gibt, ist oft relevant. Allgemeiner formuliert, ist das erste zentrale Ziel die Struktur der Lösungsmenge eines linearen Gleichungssystems. Dabei spielt der Begriff des Vektorraums eine zentrale Rolle.

Den Begriff des Vektorraums haben sicherlich viele Leser bereits kennengelernt, oftmals in Form des 3-dimensionalen „Anschauungsraums", „der Welt, in der wir leben". Doch: was bedeutet eigentlich 3-dimensional? Und auch wenn dieses Beispiel sicher wichtig ist, gibt es viele Beispiele von Vektorräumen, in denen Vektoren nicht (gut) mit „Pfeilen" veranschaulicht werden können.

Deswegen werden wir Begriffe wie „Vektorraum"oder „Dimension"mit einer präzisen Definition einführen. Dem Leser sei nahegelegt, diese Definition auswendig (!) zu beherrschen. Die Anschauung ist gut und wichtig, aber ebenso wichtig ist es, auch das abstrakte Konzept beschreiben zu können, für das man gerade ein Beispiel in Händen hält.

In diesem Skript sind einige Beispiele und Beweise aufgenommen, die in einer Vorlesung für Lehramtskandidaten der Studiengänge L2/L5 nicht diskutiert wurden, zum Beispiel der formale Beweis der Korrektheit des Gauss-Algorithmus. Solche Abschnitte sind mit dem Symbol (***) gekennzeichnet.

Quellen und Literatur: Bücher zur Linearen Algebra sind zahlreich. Empfehlen kann man z.B. „Lineare Algebra" von G. Fischer, Vieweg Verlag.

Dieses Skript entstand parallel zu einer Vorlesung von Martin Möller im Sommersemester 2017 an der Johann Wolfgang Goethe-Universität Frankfurt am Main. Es basiert auf Vorlesungsausarbeitungen zu mehreren Vorlesungen von Jürgen Wolfart in früheren Semestern, sowie auf dem des Vorlesungsskript „Lineare Algebra" (für Bachelor und L3) von Martin Möller. An der LATEXnischen Umsetzung haben Cornelia Salzmann und Jonathan Zachhuber, sowie Cristina Sarti und Claudia Baden in der Vorgängerversion, maßgeblichen Anteil.

1 Lineare Gleichungssysteme

Lineare Gleichungssysteme sind eine Kollektion von Gleichungen in einer Reihe von Unbestimmten $x, y, \ldots$ (oder auch $x_1, x_2, \ldots$), die in den Gleichungen nur mit reellen Konstanten multipliziert auftreten, also zum Beispiel

$$5x - 7y \quad = \quad 3 \ , \tag{1.1}$$

oder

$$\left\{ \begin{array}{lcl} x - 3y + 5z & = & 0 \\ 35x + 10y - 5z & = & 0 \end{array} \right\}, \tag{1.2}$$

oder auch

$$\left\{ \begin{array}{lcr} x - 3y + 5z & = & 1 \\ 35x + 10y + 5z & = & -1 \\ 36x + 7y & = & 5 \end{array} \right\}, \tag{1.3}$$

Gefragt ist in allen Fällen nach der Lösungsmenge, also nach denjenigen $\{(x, y) \in \mathbb{R}^2\}$ bzw. denjenigen $\{(x, y, z) \in \mathbb{R}^2\}$ die alle Gleichungen gleichzeitig erfüllen. Gleichungssysteme, bei denen die Zahlen auf der rechten Seite alle Null sind, werden *homogene Gleichungssysteme* genannt. Diese sind etwas einfacher zu lösen und auch in in ihrer geometrischen Anschauung direkt zu erkennen.

1.1 Geometrische Vorstellung

Wir beginnen in der Ebene $\mathbb{R}^2$, also mit 2 Variablen. Die Gleichung (1.1) beschreibt eine Gerade im $\mathbb{R}^2$. Das kann man z.B. durch Auflösen der Gleichung nach y einsehen, d.h. man parametrisiert die Gerade durch $y = -\frac{3}{7} + \frac{5}{7}x$ wie in Abbildung 1.1 dargestellt.

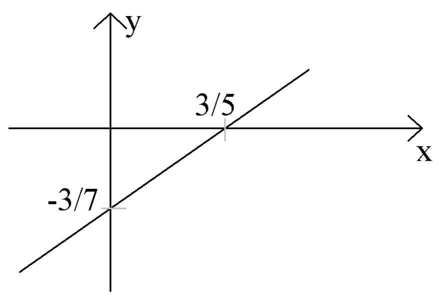

Abbildung 1.1: Gerade, die (1.1) löst.

Diese Art der Parametrisierung ist nicht immer möglich, wie z.B. in den folgenden Spezialfällen sichtbar wird.

$5x + 0y = 3$: ist ein Parallele zu y-Achse.

$0x - 0y = 3$: Keine Lösung, leere Menge.

$0x + 0y = 0$: Lösungsmenge $\mathbb{R}^2$.

Wir betrachten nun Lösungsmengen von linearen Gleichungen und Gleichungssystemen im Raum $\mathbb{R}^3$, wie zum Beispiel

$$\left\{ \begin{array}{rcr} x - 3y + 5z & = & 1 \\ 35x + 10y - 5z & = & -1 \end{array} \right\}. \tag{1.4}$$

Eine Gleichung $ax + by + cz = d \quad (a, b, c, d \in \mathbb{R})$ beschreibt eine Ebene im $\mathbb{R}^3$–ausgenommen im Fall $a = b = c = 0$. In diesem Fall ist die Lösungsmenge die leere Menge für $d \neq 0$ und der ganze Raum $\mathbb{R}^3$ für $d = 0$. In der Regel sollten Gleichungssystem mit der Variablenzahl wie in (1.2) und (1.4) räumliche Geraden als Lösungsmengen haben, das Gleichungssystem (1.3) einen Punkt $(x_0, y_0, z_0) \in \mathbb{R}^3$ als Lösungsmenge.

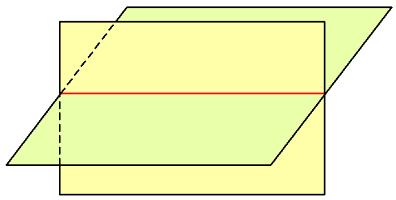

Abbildung 1.2

Das ist aber nicht immer der Fall, denn geometrisch kann es passieren, dass

i) sich drei Ebenen in einer gemeinsamen Gerade schneiden (vgl. Abbildung 1.3), wie z.B. im Fall (1.4). Hier ist die dritte Gleichung die Summe der beiden ersten ,

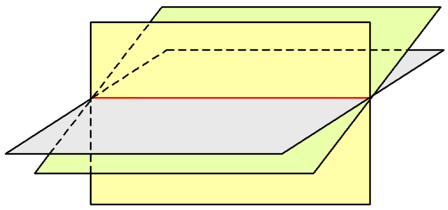

Abbildung 1.3

ii) sich je zwei Ebenen in parallelen Geraden im Raum schneiden wie im Fall des Gleichungssystems

$$\left\{\begin{array}{lcr} x-3y+5z & = & 1 \\ 35x+10y-5z & = & -1 \\ 36x+7y & = & 0 \end{array}\right\}, \tag{1.5}$$

dessen Lösungsmenge in der Abbildung 1.4)

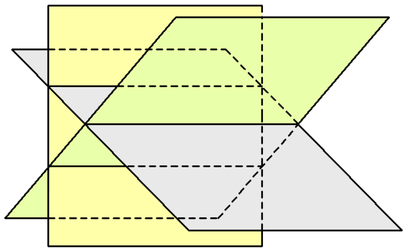

Abbildung 1.4

skizziert ist.

Bei homogenen linearen Gleichungssystemen ist das Nulltupel stets eine Lösung, die Lösungsmenge also insbesondere nicht leer. Geometrisch ist die Lösungsmenge also nur der Nullpunkt (Ursprung) selbst, eine Gerade durch den Nullpunkt oder eine Ebene, die durch den Ursprung geht.

1.2 Ad-hoc-Lösungsverfahren

Für kleine Zahl vor Variablen ist es möglich mit folgenden Verfahren Lösungen zu bestimmen. Sofern die Lösung eindeutig oder die Lösungsmenge leer ist, funktionieren diese Verfahren zuverlässig. In anderen Fällen ist das Verfahren fehleranfällig und es ist ratsam die systematischen Verfahren (Gauss-Algorithmus, siehe unten) zu verwenden.

Als Beispiel für das *Einsetzungsverfahren* betrachten wir

$$\left\{\begin{array}{lcr} 5x-7y & = & 15 \\ x+8y & = & 0 \end{array}\right\}. \tag{1.6}$$

Die zweite Gleichung nach x aufgelöst ist $x = -8y$. Die eingesetzt in $5x - 7y = 15$ ergibt $40y - 7y = 15$ und damit ist

$$y = -\frac{15}{47} \quad \text{und schließlich} \quad x = \frac{120}{47}\,.$$

Man schreibt in diesem Fall $\mathbb{L} = \{(\frac{120}{47}, -\frac{15}{47})\}$.

Als Beispiel für das *Additionsverfahren* multiplizieren wir die zweite Gleichung von (1.6) mit -5 und erhalten das System

$$\left\{ \begin{array}{rcl} 5x - 7y & = & 15 \\ -5x - 40y & = & 0 \end{array} \right\} .$$

Durch Addition der Zeilen folgt $-47y = 15$, also $y = -\frac{15}{47}$ und schließlich $x = \frac{120}{47}$. Letztlich ist dieses Verfahren aber nur eine wenig systematische Version des Folgenden.

1.3 Der Gauss-Algorithmus

Wir beginnen mit einem Beispiel für ein lineares Gleichungssystem und für unsere Lösungsstrategie. Gegeben seien die Gleichungen

$$\begin{aligned} x + 2y + 3z &= 0 \\ x + 3y + 4z &= 1. \end{aligned}$$

Gesucht sind diejenigen $(x, y, z) \in \mathbb{R}^3$, die beide Gleichungen erfüllen. Erfüllt (x, y, z) beide Gleichungen, so erfüllt (x, y, z) auch die erste und die Differenz der beiden Gleichungen. Genauer ist die Lösungsmenge von

$$\begin{aligned} x + 2y + 3z &= 0 \\ y + z &= 1 \end{aligned}$$

genau die des ursprünglichen Systems, denn die zweite Gleichung oben können wir als Summe dieser beiden Gleichungen zurückgewinnen.

Wir addieren das (-2)-fache der zweiten Gleichung auf die erste und erhalten

$$\begin{aligned} x + z &= -2 \\ y + z &= 1. \end{aligned}$$

Wählen wir $z \in \mathbb{R}$ beliebig, so ist die Lösung notwendigerweise $y = 1 - z$ und $x = -2 - z$. Anders ausgedrückt: Eine Lösung ist $(x, y, z) = (-2, 1, 0)$ und jede weitere erhält man durch Addition von $(-\lambda, -\lambda, \lambda)$ für ein $\lambda \in \mathbb{R}$ zu dieser ersten Lösung. Diesen Lösungsprozess formalisieren wir nun.

Definition 1.1 *Ein* lineares Gleichungssystem (LGS) *(über dem Körper der reellen Zahlen $\mathbb{R}$) ist eine Menge von linearen Gleichungen*

$$\left\{ \begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2 \\ \vdots & & & & & & & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m \end{array} \right\} \tag{1.7}$$

mit $a_{ij} \in \mathbb{R}$ und $b_i \in \mathbb{R}$. Ein Tupel $(x_1, \dots, x_n) \in \mathbb{R}^n$, das alle diese Gleichungen erfüllt, heißt Lösung *des linearen Gleichungssystems. Die Gleichungen*

$$\sum_{j=1}^{n} a_{kj} x_j = 0 \quad \textit{für} \quad k = 1, \dots, m$$

nennen wir das zugehörige homogene Gleichungssystem*.*

Statt über den reellen Zahlen kann man auch versuchen, lineare Gleichungssyteme, deren Koeffizienten alle rationalen Zahlen sind, nur in den rationalen Zahlen $\mathbb{Q}$ zu lösen. Das funktioniert auch, im Verfahren gibt es keine Unterschiede. Wir schreiben daher manchmal auch „ein lineares Gleichungssystem über dem Körper K“ und für K kann man dann wahlweise $\mathbb{R}$ oder $\mathbb{Q}$ lesen.

Wir können lineare Gleichungssysteme mittels Matrizen beschreiben. Sei dazu

$$A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \in \mathbb{R}^{m \times n} \quad \text{und} \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^{m \times 1}.$$

Für den Moment sind Matrizen eine formale Organisation der Koeffizienten, die Schreibarbeit beim Lösen des linearen Gleichungssystems erspart. Wir werden im Abschnitt über Matrizenmultiplikation einen weiteren Grund für diese Schreibweise kennenlernen. Das ganze lineare Gleichungssystem können wir also durch die *erweiterte Matrix* $(A \mid b) \in K^{m \times (n+1)}$ beschreiben. Dabei deutete die Trennlinie an, dass es sich ursprünglich um eine Matrix und einen Vektor mit gleich vielen Zeilen handelte.

Elementare Zeilenumformungen

Im folgenden Abschnitt beschreiben wir, welche Modifikationen wir im einleitenden Beispiel durchgeführt haben, um die Lösungsmenge eines Gleichungssystems nicht zu verändern, wohl aber die Gestalt des linearen Gleichungssystems so zu verbessern, dass wir die Lösung (fast) sofort ablesen können. Wir verwenden drei elementare Operationen.

$M_i(\lambda)$: Multiplikation der i-ten Zeile mit $\lambda \in K \setminus \{0\}$.
V_{ij} : Vertauschen der i-ten- und j-ten-Zeile.
$E_{ij}(\lambda)$: Addieren des λ-fachen der j-ten Zeile auf die i-te Zeile, $\lambda \in K \setminus \{0\}$, wobei die j-te Zeile unverändert bleibt.

Im Abschnitt über Matrizen werden wir zeigen:

Lemma 1.2 *Die elementaren Zeilenumformungen ändern die Lösungsmenge eines Gleichungssystems nicht.*

Wir können nun lineare Gleichungssysteme umformen, aber wo wollen wir eigentlich hin? Wir betrachten zunächst ausschließlich *homogene LGS*, d.h. LGS mit $b = 0$. Zum allgemeinen Fall kehren wir weiter unten zurück. Die Lösungsmenge des LGS $A \cdot x = 0$ mit

$$A = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & \vdots & \vdots & & & \vdots \\ \vdots & \ddots & \ddots & \vdots & \vdots & & & \vdots \\ 0 & \cdots & 0 & 1 & 0 & \cdots & \cdots & 0 \end{pmatrix},$$

sehen wir sofort: Es muss $x_1 = x_2 = \ldots = x_k = 0$ sein und die restlichen Einträge sind beliebig. Die Lösungsmenge $\mathbb{L}_{A,0}$ ist also

$$\begin{aligned} \mathbb{L}_{A,0} &= \{x \in K^n : x_1 = x_2 = \ldots = x_n = 0\} \\ &= \{\lambda_{k+1} \cdot e_{k+1} + \lambda_{k+2} \cdot e_{k+2} + \ldots + \lambda_n \cdot e_n, \quad \lambda_{k+1}, \ldots, \lambda_n \in K\}, \end{aligned}$$

wobei $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)^t$ den i-ten Einheitsvektor bezeichnet. Auf so eine einfache Gestalt können wir nicht jedes (homogene) LGS bringen. Ist z.B.

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

so können wir die 5 durch Addition von (-5) mal der zweiten Zeile zur ersten zu einer 0 machen, aber mehr Nullen können wir nicht erreichen. Der Kompromiß zwischen Wünschenswertem und Möglichem wird in folgender Definition festgehalten.

Definition 1.3 *Sei K ein Körper und $m, n \in \mathbb{N}$. Eine Matrix A ist in* Zeilenstufenform*, wenn sie folgende Gestalt besitzt:*

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * & 0 & \cdots & 0 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & \cdots & 0 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & \cdots & 0 & * & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \ddots & 0 & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} \in K^{m \times n}$$

d.h. wenn folgende Eigenschaften erfüllt sind:

i) Ist die i-te Zeile eine Nullzeile von A, so auch die $(i + k) - te$ für alle $k \in \mathbb{N}$. („Nullzeilen stehen ganz unten.“)

ii) Der erste Eintrag ungleich Null jeder Zeile von A ist gleich 1. Diesen Eintrag nennt man Pivotelement *(dieser Zeile).*

iii) Ist $j > i$ und der k-te Eintrag der i-ten Zeile das Pivotelement und der ℓ-te Eintrag der j-ten Zeile das Pivotelement, so ist $\ell > k$. („Die Pivotelemente in den darunterliegenden Zeilen stehen weiter rechts.“)

iv) Ist der (i, j)-te Eintrag ein Pivotelement, so ist $a_{kj} = 0$ für $k < i$. („Oberhalb der Pivotelemente stehen nur Nullen.“)

Eine Matrix nennen wir in spezieller Zeilenstufenform, wenn grob gesprochen „alle Pivotelemente maximal weit vorne stehen“. Wir kürzen die Anzahl der Pivotelemente einer Matrix in Zeilenstufenform mit r ab und machen diesen Begriff präzise.

Definition 1.4 *Sei K ein Körper und $m, n \in \mathbb{N}$. Eine Matrix A ist in* spezieller Zeilenstufenform, *falls sie in Zeilenstufenform ist und die Pivotelemente an den Stellen a_{ii} für $i = 1, \ldots, r$ stehen.*

Bei einer Matrix in spezieller Zeilenstufenform steht also oben links ein $r \times r$-Block mit Einsen auf der Diagonale und Nullen daneben. Eine Matrix in spezieller Zeilenstufenform hat also die Einträge

$$\begin{pmatrix}
1 & 0 & \cdots & 0 & 0 & c_{1\,r+1} & c_{1\,r+2} & \cdots & c_{1\,n} \\
0 & 1 & \cdots & 0 & 0 & c_{2\,r+1} & c_{2\,r+2} & \cdots & c_{2\,n} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & \cdots & 1 & 0 & c_{r-1\,r+1} & c_{r-1\,r+2} & \cdots & c_{r-1\,n} \\
0 & 0 & \cdots & 0 & 1 & c_{r\,r+1} & c_{r\,r+2} & \cdots & c_{r\,n} \\
0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0
\end{pmatrix} \in K^{m \times n} \tag{1.8}$$

Der Zweck der Zeilenstufenform und der speziellen Zeilenstufenform wird durch die folgenden Sätze offenbar.

Satz 1.5 (Gauß-Algorithmus) *Jede Matrix läßt sich nach endlich vielen Schritten in Zeilenstufenform bringen.*

Der Beweis hiervon ist auch gleichzeitig der Algorithmus. Er wird am Ende des Abschnitts der Vollständigkeit halber gegeben. Vermutlich ist es intuitiver, sich von den Beispielen leiten zu lassen.

Satz 1.6 *Die Lösungsmenge* $\mathbb{L}_{A,0}$ *eines homogonen linearen Gleichungssystems in spezieller Zeilenstufenform* (1.8) *ist gegeben durch*

$$\mathbb{L}_{A,0} = \left\{ \lambda_1 \begin{pmatrix} c_{1\,r+1} \\ c_{2\,r+1} \\ \vdots \\ c_{r-1\,r+1} \\ c_{r\,r+1} \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} c_{1\,r+2} \\ c_{2\,r+2} \\ \vdots \\ c_{r-1\,r+2} \\ c_{r\,r+2} \\ 0 \\ -1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \lambda_{n-r} \begin{pmatrix} c_{1\,n} \\ c_{2\,n} \\ \vdots \\ c_{r-1\,n} \\ c_{r\,n} \\ 0 \\ 0 \\ \vdots \\ -1 \end{pmatrix} \;,\; \lambda_1, \ldots, \lambda_{n-r} \in \mathbb{R} \right\}$$

Mit anderen Worten gesagt: Man erhält die Lösungsmenge $\mathbb{L}_{(A,0)}$ eines homogonen linearen Gleichungssystems in spezieller Zeilenstufenform (1.8), indem man n-Tupel bildet (wir werden diese später Spaltenvektoren nennen), ingesamt $n-r$ Stück bestehend aus den ersten r Einträgen der Nicht-Pivot-Spalten von (1.8) ergänzt um eine -1 und aufgefüllt mit Nullen, wie im Satz angegeben. Die gesamte Lösungsmenge erhält man, indem man diese Tupel mit beliebigen reellen Zahlen λ_i multipliziert und aufaddiert. (Wir werden später sagen, dass wir eine Linearkombination der oben konstruierten Spaltenvektoren nehmen.)

Man überzeuge sich (dringend empfohlene Übung!!), dass die Tupel, die in Satz 1.6 gelistet sind, in der Tat Lösungen des linearen Gleichungssystems sind. Zum Beweis, dass dies die gesamte Lösungsmenge ist, kommen wir in Abschnitt 4.4 zurüeck.

Wir können nur Matrizen durch den Gauss-Algorithmus in Zeilenstufenform bringen und von Matrizen in spezieller Zeilenstufenform die Lösungsmenge direkt ablesen. Der Übergang zwischen der (allgemeinen) Zeilenstufenform und der speziellen Zeilenstufenform geschieht durch geeignete Spaltenvertauschungen. **Achtung: Im Gegensatz zu Zeilenumformungen ändern Spaltenvertauschungen die Lösungsmenge eines linearen Gleichungssystems**. Diese Änderung der Reihenfolge der Lösungsvariablen muss also nach Ablesen der Lösungsmenge unbedingt wieder rückgängig gemacht werden. Dies wird in Beispiel 1.9 erläutert.

Den Fall inhomogener Gleichungssystem, d.h. den Fall, dass b nun Null verschieden sein kann, führt man auf den Fall homogener linearer Gleichungssysteme zurück.

Satz 1.7 (Gauß-Algorithmus im inhomogenen Fall) *Jede Matrix* $(A|b)$ *läßt sich nach endlich* *vielen elementaren Zeilenumformungen in die Zeilenstufenform*

$$(\widetilde{A},\widetilde{b}) = \left(\begin{array}{cccccccccccccccccc|c} 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * & 0 & \cdots & 0 & * & \cdots & * & \widetilde{b}_1 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & \cdots & 0 & * & \cdots & * & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & \cdots & 0 & * & \cdots & * & \widetilde{b}_r \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \widetilde{b}_{r+1} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \ddots & 0 & \vdots & & \vdots & 0 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \end{array}\right) \tag{1.9}$$

bringen. Ist $\widetilde{b}_{r+1}$ ungleich Null, so ist $\mathbb{L}_{(A|b)} = \emptyset$, d.h. das lineare Gleichungssystem nicht lösbar. Ist $\widetilde{b}_{r+1} = 0$, so ist

$$\mathbb{L}_{(A|b)} = \left\{ \begin{pmatrix} \widetilde{b}_{1\,n} \\ \widetilde{b}_{2\,n} \\ \vdots \\ \widetilde{b}_{r-1\,n} \\ \widetilde{b}_{r\,n} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right\} + \mathbb{L}_{(A|0)}$$

jede Lösung eine Summe aus der speziellen Lösung und einer beliebigen Lösung des zugehörigen homogengen linearen Gleichungssystems.

Beispiel 1.8 Wir beginnen mit folgendem linearen Gleichungssystem

$$\begin{aligned} 3x_1 + 6x_3 + 3x_5 &= 6 \\ x_1 + x_2 + 2x_3 &= 1 \\ 2x_2 - x_4 + 10x_5 &= 2 \\ x_2 + 3x_3 + 2x_4 + 2x_5 &= 12 \end{aligned}$$

bei dem (zufälligerweise) keine Spaltenvertauschungen notwendig sind. Wir protokollieren die Gaußschritte (von rechts nach links) wie in der Vorlesung:

$$\left(\begin{array}{ccccc|c} 3 & 0 & 6 & 0 & 3 & 6 \\ 1 & 1 & 2 & 0 & 0 & 1 \\ 0 & 2 & 0 & -1 & 10 & 2 \\ 0 & 1 & 3 & 2 & 2 & 12 \end{array}\right) \overset{E_{21}(-1)M_1(\frac{1}{3})}{\mapsto} \left(\begin{array}{ccccc|c} 1 & 0 & 2 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 & -1 & -1 \\ 0 & 2 & 0 & -1 & 10 & 2 \\ 0 & 1 & 3 & 2 & 2 & 12 \end{array}\right) \overset{E_{42}(-1)E_{32}(-2)}{\mapsto}$$

$$\left(\begin{array}{ccccc|c} 1 & 0 & 2 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & -1 & 12 & 4 \\ 0 & 0 & 3 & 2 & 3 & 13 \end{array}\right) \overset{E_{13}(-2)M_3(\frac{1}{3})V_{34}}{\mapsto} \left(\begin{array}{ccccc|c} 1 & 0 & 0 & -\frac{4}{3} & -1 & -\frac{20}{3} \\ 0 & 1 & 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & \frac{2}{3} & 1 & \frac{13}{3} \\ 0 & 0 & 0 & -1 & 12 & 4 \end{array}\right) \overset{E_{14}(\frac{4}{3})E_{34}(-\frac{2}{3})M_4(-1)}{\mapsto}$$

$$\left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & -17 & -12 \\ 0 & 1 & 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & 0 & 9 & 7 \\ 0 & 0 & 0 & 1 & -12 & -4 \end{array}\right)$$

Offenbar ist das LGS lösbar und wir können x_5 als freien Parameter wählen. Der Lösungsraum ist also

$$\mathbb{L} = \left\{ \begin{pmatrix} -12 \\ -1 \\ 7 \\ -4 \\ 0 \end{pmatrix} + t \begin{pmatrix} -17 \\ -1 \\ 9 \\ 12 \\ -1 \end{pmatrix} \in \mathbb{R}^5 \,:\, t \in \mathbb{R} \right\}.$$

Beispiel 1.9 Wir wenden den Gaußalgorithmus an, um das folgende LGS in Zeilenstufenform zu bringen:

$$\left(\begin{array}{cccccccc|c} 0 & 0 & 0 & 0 & 0 & 2 & 3 & 0 & 1 \\ 1 & 1 & 2 & 1 & 0 & 1 & 2 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 3 & 2 & 0 & 5 \\ 2 & 2 & 5 & 1 & 1 & 3 & 0 & 5 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 4 \end{array}\right) \overset{M_5(\frac{1}{2})M_4(\frac{1}{2})E_{21}(-2)V_{12}V_{14}}{\mapsto} \left(\begin{array}{cccccccc|c} 1 & 1 & 2 & 1 & 0 & 1 & 2 & 1 & 3 \\ 0 & 0 & 1 & -1 & 1 & 1 & -4 & 3 & 2 \\ 0 & 0 & 0 & 0 & 1 & 3 & 2 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{3}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \end{array}\right)$$

$$\overset{E_{15}(5)E_{25}(-3)E_{14}(-5)E_{24}(2)E_{34}(-3)E_{13}(2)E_{23}(-1)E_{12}(-2)}{\mapsto} \left(\begin{array}{cccccccc|c} 1 & 1 & 0 & 3 & 0 & 0 & \frac{13}{2} & 0 & \frac{33}{2} \\ 0 & 0 & 1 & -1 & 0 & 0 & -3 & 0 & -8 \\ 0 & 0 & 0 & 0 & 1 & 0 & -\frac{5}{2} & 0 & \frac{7}{2} \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{3}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \end{array}\right)$$

Die spezielle Zeilenstufenform erhalten wir nun, indem wir die Spalten (und damit die Variablen) folgendermaßen vertauschen:

$$\begin{array}{cccc} 1 \mapsto 1 & 2 \mapsto 6 & 3 \mapsto 2 & 4 \mapsto 7 \\ 5 \mapsto 3 & 6 \mapsto 4 & 7 \mapsto 8 & 8 \mapsto 5 \end{array}$$

Somit erhalten wir als spezielle Zeilenstufenform:

$$\left(\begin{array}{cccccccc|c} 1 & 0 & 0 & 0 & 0 & 1 & 3 & \frac{13}{2} & \frac{33}{2} \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 & -3 & -8 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -\frac{5}{2} & \frac{7}{2} \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & \frac{3}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 2 \end{array}\right)$$

Als Lösungsraum der speziellen Zeilenstufenform erhalten wir also

$$\mathbb{L}_{\text{spez}} = \left\{ \begin{pmatrix} \frac{33}{2} \\ -8 \\ \frac{7}{2} \\ \frac{1}{2} \\ 2 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ -1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 3 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 13/2 \\ -3 \\ -5/2 \\ 3/2 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R} \right\}.$$

Indem wir die obige Permutation (rückwärts) auf die Zeilen anwenden, erhalten wir die Lösungsmenge des ursprünglichen Systems:

$$\mathbb{L} = \left\{ \begin{pmatrix} \frac{33}{2} \\ 0 \\ -8 \\ 0 \\ \frac{7}{2} \\ \frac{1}{2} \\ 0 \\ 2 \end{pmatrix} + \lambda_1 \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 3 \\ 0 \\ -1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 13/2 \\ 0 \\ -3 \\ 0 \\ -5/2 \\ 3/2 \\ -1 \\ 0 \end{pmatrix}, \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R} \right\}.$$

Beweis von Satz 1.5 (Gauß-Algorithmus) (***) : Wir beweisen den Satz durch Induktion nach der Anzahl der Spalten. Ist $A = (a_{ij}) \in K^{m\times 1}$ ein Nullspaltenvektor, so ist A in Zeilenstufenform. Andernfalls können wir durch Vertauschen von Zeilen erreichen, dass $a_{11} \neq 0$ ist. Wenden wir nun $M_1(a_{11}^{-1})$ an, so erreichen wir $a_{11} = 1$. Wir wenden nun $E_{j1}(-a_{j1})$ an, d.h. wir addieren $-a_{j1}$ mal die erste Zeile auf die j-te Zeile für $j = 2, \ldots, m$. Damit erreichen wir, dass

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

ist und diese Matrix ist offenbar in Zeilenstufenform.

Wir können nun annehmen, dass wir die Zeilenstufenform für Matrizen mit weniger als n Spalten durch elementare Zeilenumformungen erreichen können und betrachten $A \in$

$K^{m \times n}$. Ist die erste Spalte von A eine Nullspalte, also

$$A = \begin{pmatrix} 0 & \\ \vdots & B \\ 0 & \end{pmatrix},$$

so können wir nach Induktionsvoraussetzung B auf Zeilenstufenform B' bringen. Dann ist auch $A' = \begin{pmatrix} 0 & B' \end{pmatrix}$ in Zeilenstufenform.

Ist die erste Spalte von $A = (a_{ij}) \in K^{m \times n}$ keine Nullspalte, so erreichen wir durch Zeilenvertauschung, dass $a_{11} \neq 0$ ist. Durch anwenden von $M_1(a_{11}^{-1})$ und $E_{1i}(-a_{i1})$ erreichen wir wie im ersten Fall, dass die erste Spalte nur einen von Null verschiedenen Eintrag hat, genauer, dass wir A zu

$$A' = \begin{pmatrix} 1 & B' \\ 0 & \\ \vdots & C' \\ 0 & \end{pmatrix}$$

mit $B' \in K^{1 \times (n-1)}$ und $C' \in K^{(m-1) \times (n-1)}$ umformen.

Wir wenden nun die Induktionshypothese an und erreichen durch elementare Zeilenumformungen, dass C' zu C'' in Zeilenstufenform umgeformt wird. Das ändert nichts an der ersten Zeile, also mit $B'' = B'$ wird A' zu

$$A'' = \begin{pmatrix} 1 & B'' \\ 0 & \\ \vdots & C'' \\ 0 & \end{pmatrix} = (a''_{ij})$$

umgeformt. A'' ist noch nicht ganz in Zeilenstufenform. Wir müssen noch sicherstellen, dass oberhalb der Pivotelemente Nullen stehen. Sei $P_{A''}$ die Menge der Spalten, in denen A'' ein Pivotelement hat und für $j \in P_{A''}$ sei $i = i(j)$ die Zeile von A'', deren Pivotelement in der Spalte j steht. Für alle $j \in P_{A''} \setminus \{1\}$ addieren wir $-a''_{1j}$ Mal die $i(j) - te$ Zeile auf die erste Zeile, um dies zu erreichen, d.h. wir wenden die Zeilenoperation $E_{i(j)1}(-a''_{1j})$ an. Damit haben wir A'' und somit A in Zeilenstufenform umgeformt. $\square$

Der Beweis der ersten Aussage von Satz 1.7 folgt hieraus unmittelbar, indem man nach Anwendung des obigen Algorithmus' auf A noch elementare Zeilenoperationen auf den Zeilen $r+1, \ldots, m$ durchführt und die rechte Seite des LGS, d.h. die $\widetilde{b_i}$ in die gewünschte Gestalt zu bringen. Der Beweis der zweiten Aussage von Satz 1.7 verschieben wir ebenfalls auf Abschnitt 4.4.

2 Vektoren und Vektorräume

2.1 Ein Beispiel

Sei S die Menge der gerichteten Strecken der Ebene $\mathbb{R}^2$ oder des Anschauungsraums $\mathbb{R}^3$. Parallele Strecken gleicher Richtung und Länge (vgl. dazu Abbildung 2.1) sollen identifiziert werden, d.h. wir führen auf S eine Äquivalenzrelation „$\sim$“ ein (= „gehen durch Parallelverschiebung auseinander hervor“) und betrachten nur noch die Äquivalenzklassen.

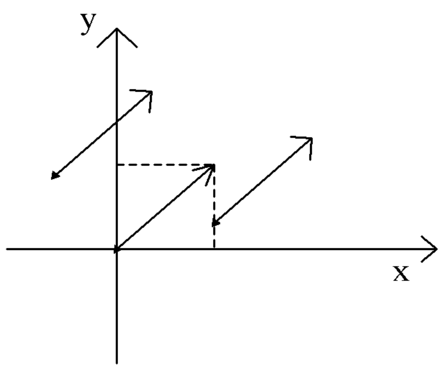

Abbildung 2.1

Andere Sichtweise: Lege den Anfangspunkt der Strecke in den Nullpunkt des Koordinatensystems und beschreibe die Strecke durch das Koordinatenpaar (x, y) des Endpunkts bzw. durch sein Koordinatentripel (x, y, z) im Raum. Wir sprechen dann von **Vektoren**. Vektoren kann man addieren (siehe Abbildung 2.2) und mit (reellen) Zahlen multiplizieren (vgl. Abbildung 2.3).

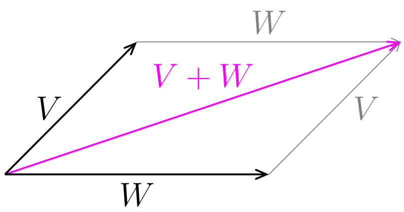

Abbildung 2.2

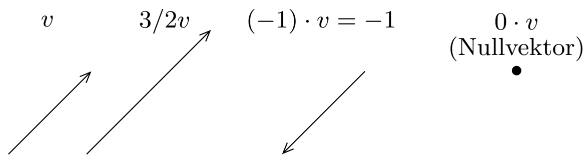

Abbildung 2.3

Mit Zahlenpaaren und -tripeln wird einfach komponenetenweise gerechnet:

$$(x, y) + (u, w) := (x + u, y + w)$$
$$r(x, y) := (rx, ry)$$

Offenbar erfüllen die Vektoren bekannte Rechenregeln, wie zum Beispiel das Assoziativgesetz (für je drei Vektoren u, v und w gilt $v + (u + w) = (v + u) + w$), das Kommutativgesetz (für je zwei Vektoren u und v gilt $v+u = u+v$) und das Distributivgesetz für die Skalarmultiplikation (für eine reelle Zahl λ und Vektoren u und v gilt $\lambda(u+v) = \lambda u+\lambda v$). Des Weiteren gibt es einen ausgezeichneten Nullvektor und jeder Vektor v hat ein additives Inverses $(-v)$.

Dies ist leicht zu verifizieren (zeichnerisch oder rechnerisch) und motiviert den folgenden Abschnitt.

2.2 Die Axiome eines Vektorraums

Im folgenden sei K ein Körper (z.B. $K = \mathbb{Q}$, die rationalen Zahlen, oder $K = \mathbb{R}$, die reellen Zahlen).

Wir definieren zuerst eine Struktur, die die Addition in den ganzen Zahlen, den rationalen Zahlen, den reellen Zahlen oder den Vektoren, aber auch die Multiplikation in $\mathbb{Q} \setminus \{0\}$ oder $\mathbb{R} \setminus \{0\}$ verallgemeinert.

Definition 2.1 *Eine Menge G heißt* abelsche Gruppe, *falls es eine Abbildung*

$$+\colon G \times G \to G$$

gibt, die folgenden Eigenschaften genügt

(GN) es gibt ein Element $0 \in G$, so dass für jedes $x \in G$ gilt: $+(x, 0) = x$ (Neutralelement);

(GI) es gibt zu jedem $x \in G$ ein Element $y \in G$ mit $+(x, y) = 0$; wir schreiben $-x := y$ (Inverses);

(GK) für alle $x, y \in G$ ist $+(x, y) = +(y, x)$ (Kommutativität);

(GA) für alle $x, y, z \in G$ gilt: $+(x, +(y, z)) = +(+(x, y), z)$ (Assoziativität).

In diesem Fall definieren wir $x + y := +(x, y)$ für $x, y \in G$.

Beispiel 2.2 Offenbar bilden die reellen Zahlen $\mathbb{R}$ mit der gewöhnlichen Addition $+$ eine abelsche Gruppe. Auch das kartesische Produkt $\mathbb{R} \times \mathbb{R}$ ist mit komponentenweiser Addition eine abelsche Gruppe.

Im Fall eines Vektorraums bilden die Vektoren eine Gruppe (man kann sie addieren); Vektoren können aber auch mit Skalaren multipliziert werden.

Definition 2.3 *Eine abelsche Gruppe V heißt* K-Vektorraum, *falls es eine Abbildung* („Skalarmultiplikation“)

$$\cdot\colon K \times V \longrightarrow V$$

gibt, die folgenden Eigenschaften genügt:

(N) $1 \cdot a = a$ für alle $a \in V$ („Nichttrivialität der Skalarmultiplikation“)

(A) $(\lambda \cdot \mu) \cdot a = \lambda \cdot (\mu \cdot a)$ für alle $a \in V$ und alle $\lambda, \mu \in K$ („Assoziativität zwischen Skalarmultiplikation und Multiplikation in K“)

(D) $\lambda \cdot (a+b) = \lambda \cdot a + \lambda \cdot b$ *und* $(\lambda + \mu) \cdot a = \lambda a + \mu a$ *für alle* $a, b \in V$ *und alle* $\lambda, \mu \in K$ *(„Distributivgesetze“).*

Beispiel 2.4 Der Vektorraum $\mathbb{R}^n$ für $n \in \mathbb{N}$. Wir haben bereits gesehen, dass das kartesische Produkt $\mathbb{R}^n$ mit komponentenweiser Addition eine abelsche Gruppe ist. Wir definieren für $\lambda \in \mathbb{R}$ und $(x_1, \ldots, x_n) \in \mathbb{R}^n$ die Skalarmultiplikation

$$\lambda \cdot (x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n).$$

Damit ist offenbar (N) erfüllt. Ebenso rechnet man die anderen Vektorraumaxiome leicht nach, z.B.

$$\begin{aligned}(\lambda+\mu)(x_1,\ldots,x_n) &= \big((\lambda+\mu)x_1,\ldots,(\lambda+\mu)x_n\big)\\ &= (\lambda x_1+\mu x_1,\ldots,\lambda x_n+\mu x_n)\\ &= (\lambda x_1,\ldots,\lambda x_n)+(\mu x_1,\ldots,\mu x_n)\\ &= \lambda(x_1,\ldots,x_n)+\mu(x_1,\ldots,x_n)\end{aligned}$$

für alle $\lambda, \mu \in \mathbb{R}$ und alle $(x_1, \ldots, x_n) \in \mathbb{R}^n$.

Beispiel 2.5 Die Abbildung $\mathbb{R} \times \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ gegeben durch

$$\lambda(x_1, x_2, x_3) = (\lambda x_1, \lambda^2 x_2, \lambda^3 x_3)$$

macht aus der abelschen Gruppe $\mathbb{R}^3$ **keinen Vektorraum**, denn für den Vektor $(x_1, x_2, x_3) = (1, 1, 1)$ und die Skalare $\lambda = \mu = 1$ folgt aus $(\lambda + \mu)(1, 1, 1) = \lambda(1, 1, 1) + \mu(1, 1, 1)$ die Gleichheit $(2, 4, 8) = (2, 2, 2)$, ein Widerspruch.

Beispiel 2.6 Der Vektorraum K[X] der Polynome (***) über einem Körper K.
Die Menge der Polynome über einem Körper $K[X]$ bildet einen K-Vektorraum: Wir betrachten zunächst nur die zugrundeliegende kommutative Gruppe $(K[X], +)$ und führen eine Skalarmultiplikation $K \times K[X] \longrightarrow K[X]$ durch

$$\lambda \cdot \sum_{i=1}^{n} b_i X^i = \sum_{i=1}^{n} (\lambda \cdot b_i) X^i$$

ein. Damit wird $K[X]$ zu einem K-Vektorraum, denn die Axiome verifiziert man leicht, z.B.

$$\begin{aligned}(\lambda\cdot\mu)\sum_{i=0}^{n} b_i X^i &= \sum_{i=0}^{n}(\lambda\cdot\mu)\cdot b_i\cdot X^i\\ &= \sum_{i=0}^{n}\lambda(\mu\cdot b_i)X^i\\ &= \lambda\cdot\sum_{i=0}^{n}(\mu\cdot b_i)X^i\\ &= \lambda\cdot(\mu\cdot\sum_{i=0}^{n} b_i X^i)\end{aligned}$$

aufgrund der Assoziativität der Multiplikation in K.

Wir halten noch 2 Konsequenzen der Axiome fest:

Lemma 2.7 *Sei V ein K-Vektorraum. Dann gilt für alle $\lambda \in K$ und $x \in V$:*

i) $\lambda \cdot x = 0$ *genau dann, wenn* $\lambda = 0$ *oder* $x = 0$.

ii) $(-\lambda) \cdot x = -(\lambda \cdot x)$, *insbesondere* $(-1) \cdot x = -x$.

Beweis : Übung. □

Man beachte, dass wir in diesem Lemma mehrfach das scheinbar (?!) überflüssige Vektorraumaxiom (N) verwendet haben.

2.3 Untervektorräume

Wir betrachten noch einmal die Lösungsmenge eines homogenen LGS. Gesucht haben wir Lösungen in dem Vektorraum K^n und die Lösungsmenge selbst ist Teilmenge hiervon und wiederum ein Vektorraum. Daher folgende Begriffsbildung.

Definition 2.8 *Sei V ein K-Vektorraum. Eine Teilmenge $U \subseteq V$ heißt* Untervektorraum, *falls U bezüglich der auf V erklärten Addition und Skalarmultiplikation ein Vektorraum ist.*

Das ist nicht sehr praktisch zu verifizieren, daher beweisen wir sofort folgendes Kriterium.

Proposition 2.9 *Eine Teilmenge $U \subseteq V$ ist genau dann ein Untervektorraum, wenn gilt:*

i) $U \neq \emptyset$

ii) Für alle $x, y \in U$ und alle $\lambda \in K$ gilt $x + y \in U$ und $\lambda \cdot x \in U$.

Man sagt zu ii) auch, dass U bezüglich Addition und Skalarmultiplikation abgeschlossen sein muß.

Beweis : Ist U ein Untervektorraum, also eine abelsche Gruppe, so ist $0 \in U$, also gilt i). Außerdem ist bei einem Vektorraum U Addition eine Abbildung $U \times U \longrightarrow U$ und Skalarmultiplikation eine Abbildung $K \times U \longrightarrow U$, woraus ii) folgt.
Umgekehrt folgen aus den Rechenregeln in V sofort die Axiome (GA) und (GK) einer abelschen Gruppe sowie (N), (A) und (D) der Vektorraumaxiome. Es bleibt die Existenz eines neutralen und inversen Elements bzgl. $+$ zu zeigen. Aus i) folgt, dass es ein $x \in U$ gibt. Nach ii) ist dann auch $0 \cdot x = 0 \in U$. Außerdem ist nach ii) zu jedem $u \in U$ auch $(-1) \cdot u \in U$ und nach obigem Lemma ist $(-1) \cdot u = -u$. Damit haben wir die verbleibenden Axiome einer abelschen Gruppe gezeigt. □

Die Konstruktion von Lösungen eines LGS führt zu einem weiteren Begriff:

Definition 2.10 *Sind $v_1, \dots, v_k \in V$ und $\lambda_1, \dots, \lambda_k \in K$ so nennt man den Vektor*

$$v = \sum_{i=1}^{k} \lambda_i \cdot v_i$$

eine Linearkombination *von $v_1, \dots, v_k$. Man sagt auch, v sei als Linearkombination der v_i* dar-stellbar *oder aus den v_i* linear kombinierbar.

Man beachte, dass in Linearkombination nur Summen von Vektoren mit endlich vielen Summanden auftreten.
Die wichtigste Quelle von Untervektorräumen sind Mengen von Linearkombinationen.

Definition 2.11 *Die* lineare Hülle *oder der* Spann *$[a_1, \dots, a_k]$ der Vektoren $a_1, \dots, a_k \in V$ ist die Menge aller Vektoren von V, die sich als Linearkombinationen von $a_1, \dots, a_k$ schreiben lassen. Allgemeiner, für eine Teilmenge $M \subseteq V$ ist die lineare Hülle $[M]$ die Menge aller Vektoren von V, die sich als Linearkombination von Elementen von M schreiben lassen.*

Wir setzen $[\emptyset] = \{0\}$ und zeigen:

Proposition 2.12 *Für jede Teilmenge $M \subseteq V$ ist $[M]$ ein Untervektorraum von V.*

Beweis : Für $M = \emptyset$ ist dies obenstehende Konvention und für $M \neq \emptyset$ müssen wir noch Eigenschaft ii) des Untervektorraumkriteriums nachprüfen. Sind $x = \sum_{i=1}^{k} \lambda_i a_i, \; a_i \in M$ und $y = \sum_{i=1}^{\ell} \mu_i b_i, \; b_i \in M$. Linearkombination von Elementen aus M, so ist auch

$$x + y = \sum_{i=1}^{k} \lambda_i a_i + \sum_{i=1}^{\ell} \mu_i b_i$$

Linearkombination und für alle $\lambda \in K$ ist

$$\lambda \cdot x = \sum_{i=1}^{k} (\lambda \cdot \lambda_i) a_i$$

wieder eine Linearkombination von Elementen aus M und damit in $[M]$. □

Einige weitere Eigenschaften der linearen Hülle, die direkt aus der Definition folgen, sind in folgendem Lemma festgehalten.

Lemma 2.13 *Für alle Teilmengen M, M_1, M_2 eines Vektorraums gilt:*

i) $M \subseteq [M]$

ii) *Ist $M_1 \subset M_2$, so ist $[M_1] \subseteq [M_2]$.*

iii) *Es gilt $[M] = M$ genau dann, wenn M ein Untervektorraum ist.*

2.4 Anwendungen der Vektorrechnung in der Elementargeometrie:

Satz 2.14 *Die Diagonalen eines Parallelogramms halbieren sich gegenseitig (vgl. Abbildung 2.4).*

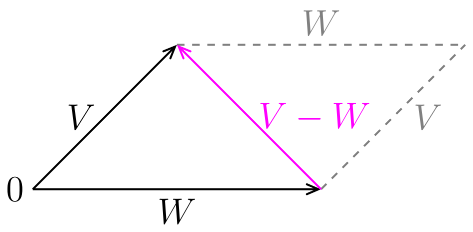

Abbildung 2.4

Beweis : Wir zeigen zunächst, dass die Diagonalen-Vektoren $v + w$ und $v - w$ sind. Der Vektor vom Nullpunkt zum Mittelpunkt der $v+w$-Diagonale ist $\frac{1}{2}(v+w)$, das ist aber gleich $w + \frac{1}{2}(v - w)$, also auch auf halber Strecke der anderen Diagonale. □

Satz 2.15 *Der Schwerpunkt eines Dreiecks ist der Schnittpunkt der drei Seitenhalbierenden (vgl. Abbildung 2.5).*

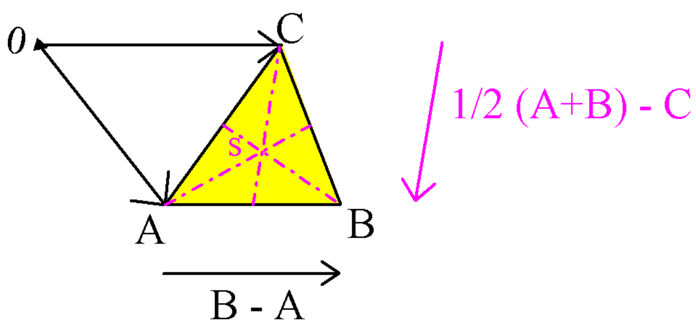

Abbildung 2.5

Beweis : Der Nullpunkt der Ebene sei beliebig gewählt. Wir identifizieren alle Punkte mit den Endpunkten der Vektoren vom Nullpunkt aus. Dann ist

$$\frac{1}{2}(A + B) = A + \frac{1}{2}(B - A)$$

der Seitenmittelpunkt der Strecke $\overline{AB}$. Entsprechend ist $\frac{1}{2}(A+C)$ der Seitenmittelpunkt von $\overline{AC}$ und $\frac{1}{2}(B + C)$ der Seitenmittelpunkt der Strecke $\overline{BC}$. Schließlich ist

$$S := \frac{1}{3}(A + B + C) = C + \frac{2}{3}(\frac{1}{2}(A + B) - C)$$

der Punkt, welcher die Seitenhalbierende der Seite C im Verhältnis $2:1$ teilt. Wegen

$$S = A + \frac{2}{3}(\frac{1}{2}(B+C) - A) = B + \frac{2}{3}(\frac{1}{2}(A+C) - B)$$

liegt er aber auch auf den anderen Seitenhalbierenden und teilt sie im gleichen Verhältnis. □

Genauso zeigt man folgenden Satz.

Satz 2.16 *Gegeben ein Tetraeder im $\mathbb{R}^3$ mit Eckpunkten A, B, C, D (nicht in einer Ebene gelegen). Verbindet man jeden Eckpunkt mit dem Schwerpunkt des gegenüberliegenen Randdreiecks, so schneiden sich diese vier Strecken in einem Punkt $S = \frac{1}{4}(A+B+C+D)$, der alle vier Verbindungsstrecken im Verhältnis $3:1$ teilt.*

3 Erzeugende, Basis, Dimensionen

3.1 Lineare Unabhängigkeit

Wir wollen einen Begriff dafür schaffen, dass eine Menge von Vektoren ihre lineare Hülle ohne Redundanz aufspannt. Dies erkennt man bereits an Linearkombinationen, die den Nullvektor darstellen. Ist z.B. $a = (1,0), b = (0,1)$ und $c = (1,1) \in \mathbb{R}$, so ist

$$0 = 0 \cdot a + 0 \cdot b + 0 \cdot c$$

eine Linearkombination von a, b, c, die den Nullvektor ergibt, genannt die *triviale Linearkombination*. Hier gilt aber auch

$$0 = 1 \cdot a + 1 \cdot b + (-1) \cdot c$$

und in dieser Linearkombination sind nicht alle Koeffizienten gleich Null. Sie wird *nichttrivial* genannt.

Definition 3.1 *Eine Teilmenge M eines K-Vektorraums V heißt* linear unabhängig *(l.u.), falls die Null nur in trivialer Weise eine Linearkombination ist, d.h. falls für alle $k \in \mathbb{N}$, alle $a_i \in M$ und $\lambda_i \in K$ aus $\sum_{i=1}^{k} \lambda_i a_i = 0$ folgt $\lambda_i = 0\ \forall i = 1, \ldots, k$.*

Die Menge M heißt linear abhängig *(l.a.), falls es eine nichttriviale Linearkombination von Vektoren in M gibt, die Null ist, d.h. falls es $a_i \in M$ und $\lambda_i \in K$ gibt, sodass nicht alle $\lambda_i = 0$ sind und*

$$\sum_{i=1}^{k} \lambda_i a_i = 0$$

gilt.

Beispiele 3.2 i) Im einleitenden Beispiel ist $\{a, b, c\}$ l.a., aber die Mengen $\{a, b\}, \{a, c\}, \{a, c\}$ und $\{b, c\}$ sind allesamt l.u.

ii) Die Menge $\{a\}$ ist l.a. genau dann, wenn $a = 0$.

iii) Allgemeiner: Ist $0 \in M$, so ist M l.a., denn $0 = 1 \cdot 0$ ist eine nichttriviale Linearkombination von Elementen in M.

iv) Sind $a, b \in V \setminus \{0\}$ linear abhängig, so gilt $0 = \lambda a + \mu b$ mit $(\lambda, \mu) \neq (0, 0)$ und daher $\lambda \neq 0$ und $\mu \neq 0$. Also ist

$$a = \frac{-\mu}{\lambda} \cdot b \quad \text{und} \quad b = \frac{-\lambda}{\mu} \cdot a.$$

In diesem Fall nennt man die Vektoren auch *proportional*. Sind umgekehrt a und b proportional, d.h. $a = \delta \cdot b$, so ist $0 = 1 \cdot a - \delta \cdot b$ eine nichttriviale Linearkombination, also $\{a, b\}$ l.a.

In der Definition von linearer (Un)abhängigkeit haben wir stets von einer Menge gesprochen, wir haben nie definiert „a_1 ist linear unabhängig von $a_2, \dots, a_p$“. Eine richtige Version solch eines Begriffs liefert die folgende Proposition. Dennoch wird dem Leser nahegelegt, den Nachweis der linearen (Un)abhängigkeit, wenn möglich mit Hilfe der ursprünglichen Definition statt mit der folgenden Proposition zu führen. Dies ist weniger anfällig für elementare Fehler.

Proposition 3.3 *Eine Teilmenge $M \subseteq V$ ist genau dann linear abhängig, wenn es ein Element $a \in M$ gibt, das sich als Linearkombination von $M \setminus \{a\}$ schreiben lässt.*

Beweis : Ist M linear abhängig und $0 = \sum_{i=1}^{k} \lambda_i a_i$, so gibt es einen Index i_0 mit $\lambda_{i_0} \neq 0$. Also ist $a_{i_0} = \sum_{\substack{i=1 \\ i \neq i_0}}^{k} \frac{-\lambda_i}{\lambda_{i_0}} a_i$ die geforderte Linearkombination. Ist umgekehrt $M \ni a = \sum_{i=1}^{k} \lambda_i a_i$, so setzen wir $a_{k+1} = a$ und $\lambda_{k+1} = -1$. Dann ist $0 = \sum_{i=1}^{k+1} \lambda_i a_i$ die gesuchte nichttriviale Darstellung des Nullvektors. □

Wir halten noch folgende direkte Konsequenzen der Definition fest.

Lemma 3.4 *i) Ist $M_1 \subseteq V$ linear abhängig und $M_2 \supseteq M_1$, so ist auch M_2 linear abhängig.*

ii) Ist $M_1 \subseteq V$ linear unabhängig und $M_3 \subseteq M_1$, so ist auch M_3 linear unabhängig.

iii) Ist $n \in \mathbb{N}$ und $v_1, \dots, v_n \in V$ sowie $w_1, \dots, w_{n+1} \in [v_1, \dots, v_n]$, dann ist $\{w_1, \dots, w_{n+1}\}$ linear abhängig.

Beweis : i) und ii) sind klar, iii) beweisen wir durch Induktion. Die Fälle $n = 0$ und $n = 1$ sind (vgl. Beispiel 3.2) klar und wir können annehmen, dass die Aussage für $n-1$ richtig ist. Sei also

$$\begin{aligned} w_1 &= a_{11}v_1 + \ldots + a_{1p}v_p \\ &\vdots \\ w_{p+1} &= a_{p+1,1}v_1 + \ldots + a_{p+1,p}v_p \end{aligned}$$

Sind die $a_{jp} = 0$ für alle j, so sind die w_i alle in $[v_1, \ldots, v_{p-1}]$ und aus der Annahme folgt die Behauptung. Andernfalls können wie nach Vertauschen annehmen, dass $a_{p+1,p} \neq 0$ ist. Dann sind die p Vektoren

$$\widetilde{w}_j = w_j - \frac{a_{jp}}{a_{p+1,p}} \cdot w_{p+1} \quad \text{für} \quad j = 1, \ldots, p$$

allesamt in $[v_1, \ldots, v_{p-1}]$. Nach Induktion gibt es also eine Linearkombination

$$0 = \sum_{j=1}^{p} \lambda_j \widetilde{w}_j = \left(\sum_{j=1}^{p} \lambda_j w_j \right) - \left(\sum_{j=1}^{p} \frac{a_{jp} \cdot \lambda_j}{a_{p+1,p}} \right) \cdot w_{p+1},$$

wobei nicht alle λ_j Null sind. Dies ist die geforderte nichttriviale Linearkombination von $\{w_1, \ldots, w_{p+1}\}$. $\square$

3.2 Basen und der Dimensionsbegriff

In den vorherigen Abschnitten haben wir gesehen, dass es Teilmengen eines Vektorraums gibt, die diesen erzeugen und welche, die linear unabhängig sind. In diesem Abschnitt interessieren wir uns für solche Teilmengen, die beide Bedinungen auf einmal erfüllen.

Beispiel 3.5 Sei $V = \mathbb{R}^2$. Dann ist die Menge $\{(1,0), (0,1), (1,1)\}$ ein Erzeugendensystem, das nicht linear unabhängig ist; die Menge $\{(1,0)\}$ ist zwar linear unabhängig, erzeugt aber nicht ganz V und die Menge $\{(1,0), (0,1)\}$ erzeugt V und ist linear unabhängig. Auch $\{(2,3), (3,2)\}$ ist ein linear unabhängiges Erzeugendensystem von V.

In diesem Abschnitt sei V stets ein K-Vektorraum.

Definition 3.6 *Eine linear unabhängige Teilmenge $B \subseteq V$ mit $[B] = V$ heißt* Basis *von V.*

Dieser Begriff ist zentral bei der Beschreibung von Vektorräumen. Mit seiner Hilfe können wir die „Größe“ von V messen. Wir werden bald sehen, dass es viele Basen von V gibt. Ist $V = K^n$, so ist die folgende besonders „einfach“. Es sei

$$\begin{aligned} e_1 &= (1, 0, \ldots, 0, 0) \\ e_2 &= (0, 1, \ldots, 0, 0) \\ &\vdots \\ e_n &= (0, 0, \ldots, 0, 1) \end{aligned}$$

Dann ist jedes $K^n \ni (a_1, \dots, a_n) = \sum_{i=1}^n a_i \cdot e_i$ im Spann von $\{e_1, \dots, e_n\}$ und aus $\sum_{i=1}^n \lambda_i e_i = 0$ folgt durch Betrachten des j-ten Eintrags $\lambda_j = 0$. Daraus folgt die lineare Unabhängigkeit von $e_1, \dots, e_n$. Diese Menge wird *Standardbasis* von K^n genannt.

Satz 3.7 *Jeder Vektorraum V hat eine Basis.*

Dies ist offenbar ein Spezialfall von $R = \emptyset$ und $E = V$ der folgenden Aussage, die hier ohne Beweis zitiert wird.

Satz 3.8 *Sei $R \subseteq V$ linear unabhängig $R \subseteq E$ und $[E] = V$. Dann gibt es eine Basis B von V mit $R \subseteq B \subseteq E$.*

Wir notieren einige wichtige Konsequenzen.

Korollar 3.9
i) Jede linear unabhängige Teilmenge eines Vektorraums kann man zu einer Basis ergänzen.

ii) Zu jedem Untervektorraum $U_1 \subseteq V$ gibt es ein Komplement, *d.h. einen Untervektorraum U_2 mit der Eigenschaft $[U_1 \cup U_2] = V$ und $U_1 \cap U_2 = \{0\}$.*

iii) Jeder endlich erzeugbare Vektorraum hat eine endliche Basis.

iv) Hat eine Basis von V genau n Elemente, so hat jede Basis von V genau n Elemente.

Aufgrund der letzten Aussage ist folgender Begriff wohldefiniert.

Definition 3.10 *Die* Dimension *eines Vektorraums V ist die Kardinalität einer Basis von V, falls diese endlich ist, und unendlich andernfalls.*

Beweis des Korollars : Die Teile i) und iii) sind unmittelbar klar.

In ii) sei R eine Basis von U_1. Wir wenden den Satz 3.8 auf dieses R und $E = V$ an. Sei also B eine Basis von V mit $R \subseteq B$. Bezeichne $B_2 = B \setminus R$ und $U_2 = [B_2]$. Dann ist sicherlich $B \subseteq U_1 \cup U_2$, also erzeugt die Vereinigung ganz V. Außerdem: Ist $v \in U_1 \cap U_2$, so gibt es, da v sowohl in U_1, als auch in U_2 ist, Elemente $\alpha_i, \beta_i \in K$ $(i \in I)$, so dass

$$\sum_{i=1}^n \alpha_i r_i = v = \sum_{i=1}^m \beta_i b'_i \quad \text{und daher} \quad \sum_{i=1}^n \alpha_i r_i - \sum_{i=1}^m \beta_i b'_i = 0,$$

wobei $r_i \in R$ und $b'_i \in B_2$ sind. Da aber $\{r_1, \dots, r_n, b'_1, \dots, b'_m\}$ als Teilmenge einer Basis linear unabhängig ist, muss die Linearkombination trivial sein, d.h. alle $\alpha_i = \beta_i = 0$ und damit ist $v = 0$.

Zum Beweis von iv) nehmen wir an, B sei eine Basis mit n Elementen. Nach dem Lemma 3.4 sind $n+1$ Linearkombinationen aus diesen Elementen stets linear abhängig. Also hat jede Basis höchstens n Elemente. Gäbe es eine Basis B_2 mit $n_2 < n$ Elementen, so folgt mit der gleichen Überlegung, dass B linear abhängig sein müßte. $\square$

Wir notieren folgenden einfachen Satz, den man sich als „**Eindeutige Darstellbarkeit** eines Vektors in einer (gegebenen) Basis“ merken sollte.

Satz 3.11 *Ist B eine Basis des Vektorraums V, so besitzt jedes $v \in V$ eine eindeutige Darstellung als Linearkombination*

$$a = \sum_{i=1}^{n} \lambda_i b_i$$

mit $\lambda_i \in K$ und $b_i \in B$.

Beweis : Existenz einer solchen Darstellung ist unmittelbare Konsequenz aus B Basis. Zum Beweis der Eindeutigkeit nehmen wir an, dass $a \in V$ zwei Darstellungen

$$a = \sum_{i=1}^{n} \lambda_i b_i = \sum_{i=1}^{n} \mu_i b_i$$

hat. Dabei haben wir die gleichen Basisvektoren für beide Darstellungen verwendet, denn falls ein b_i in der ersten Darstellung auftritt, aber nicht in der zweiten, so können wir $\mu_i = 0$ setzen und zur zweiten Darstellung $0 \cdot b_i$ addieren, und umgekehrt. Daraus folgt nun

$$0 = \sum_{i=1}^{n} (\lambda_i - \mu_i) \cdot b_i$$

und aufgrund der linearen Unabhängigkeit von B folgt $\lambda_i = \mu_i$. Also waren die zwei Darstellungen in Wirklichkeit gleich. □

Beispiele 3.12 Offensichtlich gelten $\dim \mathbb{R}^3 = 3$, $\dim \mathbb{R}^2 = 2$, $\dim \mathbb{R}^1 = 1$, $\dim\{0\} = 0$. Jede Gerade durch den Nullpunkt des $\mathbb{R}^2$ oder $\mathbb{R}^3$ definiert einen eindimensionalen Untervektorraum, Ebenen durch den Nullpunkt im $\mathbb{R}^3$ zweidimensionale Untervektorräume.

4 Matrizen und lineare Abbildungen

Die Koeffizienten von linearen Gleichungssystemen in einem rechteckigen System, einer Matrix, anzuordnen, ist nicht nur eine Frage von Bequemlichkeit und Schreibaufwand. Denn mit solchen Matrizen kann man neben den offensichtlichen Operationen, komponentenweiser Addition und Multiplikation mit einem Skalar, auch eine weniger intuitive Operation durchführen, die Matrizenmultiplikation. Diese Definition der Matrizenmultiplikation ist durch den Bezug zu linearen Abbildungen motiviert.

4.1 Matrizenoperationen

Sei K ein beliebiger Körper (der Leser möge mit Vorstellung $K = \mathbb{R}$ oder $K = \mathbb{Q}$ weiterlesen) und seien $m, n \in \mathbb{N}$.

Definition 4.1 *Eine $m \times n$* Matrix *mit Koeffizienten in K ist ein $m \cdot n$-faches Tupel von Elementen in K, welche wie folgt angeordnet und durchnummeriert sind:*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Die Menge der $m \times n$ Matrizen wird mit $\mathrm{Mat}_{m,n}(K)$ oder mit $K^{m\times n}$ bezeichnet. Ist $m = n$, so heißt die Matrix *quadratisch* und wir bezeichnen die Menge der quadratischen $n \times n$ Matrizen mit $\mathrm{Mat}_n(K)$ oder $K^{n\times n}$.

Wir können eine Matrix in m Elemente von K^n zerlegen. Es sein $(a_{j1} \dots a_{jn})$ für $j = 1, \dots, m$ die m *Zeilen* oder *Zeilenvektoren* der Matrix. Ebenso können wir sie in n Elemente von K^m zerlegen, wir nennen

$$\begin{pmatrix} a_{1\ell} \\ \vdots \\ a_{m\ell} \end{pmatrix}$$

für $\ell = 1, \dots, n$ die n *Spalten* oder *Spaltenvektoren* der Matrix. Wir schreiben auch $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n,}}$ um die Einträge einer Matrix zu spezifizieren.

Die *Addition zweier Matrizen* $A, B \in K^{m\times n}$ definieren wir komponentenweise. Die *Skalarmultiplikation* einer Matrix $A = (a_{ij}) \in K^{m\times n}$ mit einem Element $\lambda \in K$ ist definiert als

$$\lambda \cdot A = (\lambda \cdot a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} = \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix}.$$

Definition 4.2 *Das* Produkt *zweier Matrizen $A = (a_{ij}) \in K^{m\times n}$ und $B = (b_{k\ell}) \in K^{n\times p}$ ist definiert als die Matrix $C = A \cdot B = (c_{i\ell})_{\substack{i=1,\dots,m \\ \ell=1,\dots,p}} \in K^{m\times p}$ mit den Einträgen*

$$c_{i\ell} = \sum_{j=1}^{n} a_{ij} b_{j\ell}$$

d.h. ausgeschrieben

$$c = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + \ldots + a_{1n}b_{n1} & \cdots & a_{11}b_{1\ell} + a_{12}b_{2\ell} + \ldots + a_{1n}b_{np} \\ \vdots & & \vdots \\ a_{m1}b_{11} + a_{m2}b_{21} + \ldots + a_{mn}b_{n1} & \cdots & a_{m1}b_{1\ell} + a_{m2}b_{2\ell} + \ldots + a_{mn}b_{np} \end{pmatrix}$$

Man beachte, dass das Produkt zweier Matrizen nur dann definiert ist, falls die Spaltenzahl der ersten Matrix gleich der Zeilenzahl der zweiten Matrix ist. Ist $B \in K^{n \times p}$ und $A \in K^{m \times n}$ wie oben, so ist $B \cdot A$ nur definiert, falls $p = m$ ist.

Beispiel 4.3 Als Spezialfall der Matrizenmultiplikation betrachten wir zunächst den Fall $n = 1$. In diesem Fall erhalten wir

$$\underbrace{\begin{pmatrix} a_1, \dots, a_n \end{pmatrix}}_{\text{Zeilenvektor}} \cdot \underbrace{\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}}_{\text{Spaltenvektor}} := \underbrace{a_1 b_1 + a_2 b_2 + \ldots + a_n b_n}_{\text{„Multiplikation“}} \in \mathbb{R}$$

Der Fall $m = n = 2$ und $p = 1$ ist ausgeschrieben

$$\underbrace{\begin{pmatrix} a & b \\ c & d \end{pmatrix}}_{2\times 2 \text{ Matrix}} \begin{pmatrix} x \\ y \end{pmatrix} := \underbrace{\begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}}_{\text{Zeilenweise multipliziert}}$$

Ein weiteres Beispiel für ein Matrizenprodukt ist

$$\underbrace{\begin{pmatrix} 3 & 0 & 4 \\ 7 & -1 & -1 \end{pmatrix}}_{3\times 2\text{-Matrix}} \underbrace{\begin{pmatrix} x \\ y \\ z \end{pmatrix}}_{\in \mathbb{R}^3} = \underbrace{\begin{pmatrix} 3x + 0y + 4z \\ 7x - y - z \end{pmatrix}}_{\in \mathbb{R}^2}$$

4.2 Lineare Abbildungen

Wir betrachten Matrixmultiplikation von einem anderen Standpunkt. Sei $A \in K^{m \times n}$ gegeben. Die Abbildung, die einem Spaltenvektor $x \in K^n$ den Spaltenvektor $A \cdot x \in K^m$ zuordnet, ist der Prototyp einer linearen Abbildung, wie in folgender Definition.

Definition 4.4 *Seien V und W K-Vektorräume. Eine Abbildung $f\colon V \longrightarrow W$ heißt* linear, *falls für alle $v_1, v_2 \in V$ und alle $\lambda \in K$ gilt*

$$f(v_1 + v_2) = f(v_1) + f(v_2) \quad \textit{und} \quad f(\lambda v_1) = \lambda f(v_1).$$

Man kann diese zwei Eigenschaften auch zu einer Bedingung

$$f(\lambda v_1 + v_2) = \lambda f(v_1) + f(v_2) \quad \text{für alle} \quad v_1, v_2 \in V, \quad \lambda \in K$$

zusammenfassen.

Beispiele 4.5 i) Die Abbildung $f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^3, (x,y) \longmapsto (x+y, x-y, 2x-y)$ ist linear, denn ist $(x_1,y_1),(x_2,y_2) \in \mathbb{R}^2$ und $\lambda \in K$, so ist

$$\begin{aligned}
&f\big(\lambda(x_1,y_1)+(x_2,y_2)\big) = f(\lambda x_1 + x_2, \lambda y_1 + y_2)\\
&= (\lambda x_1 + x_2 + \lambda y_1 + y_2, \lambda x_1 + x_2 - \lambda y_1 - y_2, 2\cdot\lambda\cdot x_1 + 2\cdot x_2 - \lambda y_1 - y_2)\\
&= \lambda\cdot(x_1+y_1, x_1-y_1, 2x_1-y_1) + (x_2+y_2, x_2-y_2, 2x_2-y_2)\\
&= \lambda f(x_1,y_1) + f(x_2,y_2)
\end{aligned}$$

ii) Die Abbildung $f\colon V \longrightarrow W, x \longmapsto a$ für ein festes $a \in W$ ist linear genau dann, wenn $a = 0$ ist, denn es muss $a = f(0\cdot x) = 0\cdot f(x) = 0\cdot a = 0$ gelten.

iii) Die Abbildung $f\colon \mathbb{R} \longrightarrow \mathbb{R}, x \longmapsto x^2$ ist nicht linear, denn wenn sie linear wäre, müsste gelten

$$4 = f(2) = f(1+1) = f(1) + f(1) = 1^2 + 1^2 = 2,$$

was offensichtlich ein Widerspruch ist.

Als weitere Beispiele für lineare Abbildungen betrachten wir die zwei Abbildungen $f_i : \mathbb{R}^2 \to \mathbb{R}^2$ für $i = 1,2$ gegeben durch $x \mapsto A_i \cdot x$, wobei A_i die Matrizen

$$A_1 = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \quad \text{bzw.} \quad A_2 = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}.$$

sind.

In beiden Fällen wird der Vektor $\binom{1}{0}$ auf den Vektor $\binom{\cos\varphi}{\sin\varphi}$ abgebildet, also um den Winkel φ gedreht. Handelt es sich in beiden Fällen um eine Drehung um den Winkel φ? Das Bild des Vektors $e_2 = \binom{0}{1}$ unter den beiden Abbildungen f_1 bzw. f_2 ist jedoch verschieden, es ist nämlich $f_1(e_2) = \binom{-\sin\varphi}{\cos\varphi}$, und $f_2(e_2) = \binom{\sin\varphi}{-\cos\varphi}$. (Wir halten nebenbei fest, dass wir mit Kenntnis dieser beiden Bilder alle Einträge der Matrix A_1 bzw. A_2 kennen. Das allgemeine Prinzip dahinter fasst der nächste Satz zusammen.) Wir prüfen nach, dass der Vektor $\binom{\cos\frac{\varphi}{2}}{\sin\frac{\varphi}{2}}$ unter der Abbildung f_2 fest bleibt. Es ist nämlich

$$f_2\left(\begin{pmatrix} \cos\frac{\varphi}{2} \\ \sin\frac{\varphi}{2} \end{pmatrix}\right) = \begin{pmatrix} \cos\varphi\cos\frac{\varphi}{2} + \sin\varphi\sin\frac{\varphi}{2} \\ \sin\varphi\cos\frac{\varphi}{2} - \cos\varphi\sin\frac{\varphi}{2} \end{pmatrix} = \begin{pmatrix} \cos(\varphi - \frac{\varphi}{2}) \\ \sin(\varphi - \frac{\varphi}{2}) \end{pmatrix} = \begin{pmatrix} \cos\frac{\varphi}{2} \\ \sin\frac{\varphi}{2} \end{pmatrix}.$$

Die Gerade mit Winkel $\varphi/2$ zur x-Achse ist also eine Fixgerade der Abbildung f_2. Man kann weiter nachprüfen, dass unter f_2 alle Vektoren an dieser Gerade gespiegelt werden. Die Matrix A_2 beschreibt also eine Achsenspiegelung (siehe Abbildung 4.1). Unter der Abbildung f_1 hingegen wird auch f_2 um den Winkel φ gedreht (Nachrechnen!) und damit werden alle Vektoren um diesen Winkel gedreht. Die Matrix f_1 stellt also in der Tat eine Drehung um den Winkel φ dar.

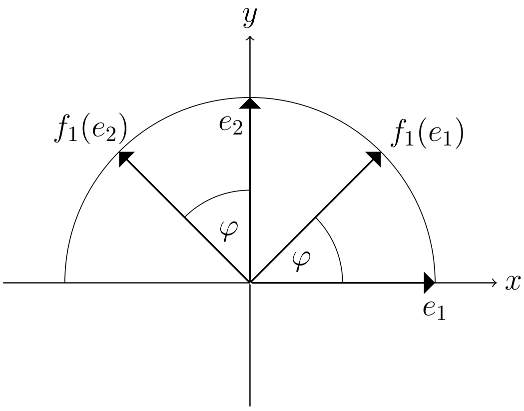

Abbildung 4.1: Drehung in der Ebene

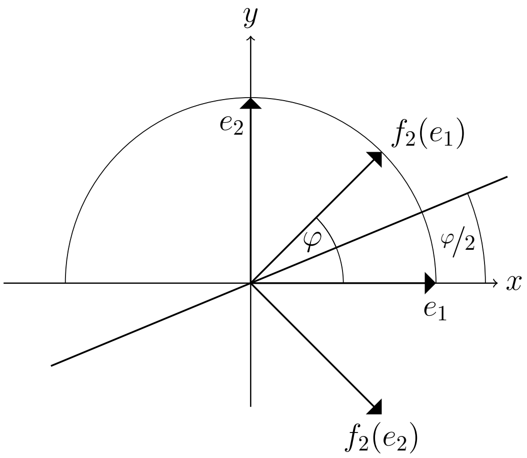

Abbildung 4.2: Achsenspiegelung

Im obigen Beispiel wird implizit behauptet, dass wir die lineare Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$ kennen, in dem wir die Bilder zweier (linear unabhängiger) Elemente kennen. Dies ist ein Spezialfall des folgenden Satzes. Diesen Satz kann man sich als „Lineare Abbildungen sind durch die Bilder einer Basis bestimmt“ merken.

Satz 4.6 *Seien V und W K-Vektorräume und $B = \{b_1, \ldots, b_n\}$ eine Basis von V. Ist $\{c_1, \ldots, c_n\} \subseteq W$ eine beliebige Menge, so gibt es genau eine lineare Abbildung $f\colon V \longrightarrow W$ mit*

$$f(b_i) = c_i \quad \textit{für} \quad i = 1, \ldots, n.$$

Beweis (***) : Ist $x \in V$, so lässt sich x eindeutig als Linearkombination

$$x = \sum_{k=1}^{n} \alpha_k b_k$$

in den Basiselementen schreiben. Ist f gesuchte lineare Abbildung, so muss

$$f(x) = f\left(\sum_{k=1}^{n} \alpha_k b_k\right) = \sum_{k=1}^{n} \alpha_k \; f\left(b_k\right) = \sum_{k=1}^{n} \alpha_k c_k \tag{4.1}$$

gelten. Das heißt, dass $f(x)$ für jedes x aus den Vorgaben auf den Basiselementen und der Linearität eindeutig festgelegt ist. Wir nehmen nun die Gleichung (4.1) als Definition und zeigen, dass die so definierte Abbildung in der Tat linear ist.
Es sei also $y \in V$ die Basisdarstellung

$$y = \sum_{k=1}^{n} \beta_k b_k$$

und es sei $\lambda \in K$. Dann gilt

$$\begin{aligned} f(\lambda x + y) &= f\left(\lambda \sum_{k=1}^{n} \alpha_k b_k + \sum \beta_k b_k\right) \\ &= f\left(\sum_{k=1}^{n} (\lambda\alpha_k + \beta_k) b_k\right) = \sum_{k=1}^{n} (\lambda\alpha_k + \beta_k) c_k \\ &= \lambda \cdot \sum_{k=1}^{n} \alpha_k c_k + \sum_{k=1}^{n} \beta_k c_k = \lambda \cdot f(x) + f(y). \end{aligned}$$

Also ist f linear. $\square$

Matrizenmultiplikation liefert nicht nur ein Beispiel für lineare Abbildungen, sondern jede lineare Abbildung kann durch Matrizen beschrieben werden. Seien nun konkret $V = \mathbb{R}^n$, $W = \mathbb{R}^m$, seien $e_1, \ldots, e_n$ die Standardbasis von V und zur Unterscheidung bezeichnen wir mit Strichen, also $e'_1, \ldots, e'_m$ die Standardbasis von W.

Satz 4.7 *Ist $f : V \to W$ eine lineare Abbildung und die Bilder der Standardbasis von V sind als Linearkombination der Standardbasis von W gegeben, also*

$$\begin{aligned} w_1 &:= f(e_1) = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} = a_{11}e'_1 + a_{21}e'_2 + \ldots + a_{m1}e'_m, \\ &\vdots \\ w_n &:= f(e_n) = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = a_{1n}e'_1 + a_{2n}e'_2 + \ldots + a_{mn}e'_m. \end{aligned} \tag{4.2}$$

Dann wird die Abbildung f beschrieben durch die Matrixmultiplikation $f(x) = A \cdot x$, wobei

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ a_{21} & & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}.$$

Man merkt sich diesen Satz, der den Übergang von einer abstrakten linearen Abbildung zu einer konkret durch Matrizenmultiplikation gegebenen linearen Abbildung, als **„Die Spalten (!!) von A sind die Bilder der Standardbasis des Ausgangsraums V ausgedrückt in der Standardbasis des Bildraums W."**

Beweis : Man rechnet mit der Definition der Matrixmultiplikation nach, dass unter der Abbildung $x \mapsto A \cdot x$ die Standardbasisvektoren e_i in der Tat auf die Vektoren w_i abgebildet werden. Dadurch ist eine lineare Abbildung nach dem vorigen Satz eindeutig bestimmt. Also stimmt die Abbildung $x \mapsto A \cdot x$ mit der Abbildung f überein. $\square$

4.3 Kern, Bild und Rang einer linearen Abbildung

Sei $f \colon V \longrightarrow W$ eine lineare Abbildung zwischen zwei K-Vektorräumen V und W.

Definition 4.8 *Die Menge aller $f(v), v \in V$ wird das* Bild *von f genannt. Die Menge $f^{-1}(0)$ wird* Kern *von f genannt und mit $\mathrm{Ker}(f)$ bezeichnet.*

Proposition 4.9 *Der Kern von f ist ein Untervektorraum von V und das Bild ist ein Untervektorraum von W.*

Beweis : Es ist $0 \in \mathrm{Ker}(f)$ und $0 \in$ Bild(f). Für die Anwendung des Untervektorraumkriteriums seien $x, y \in \mathrm{Ker}(f)$ und $\lambda \in K$ gegeben, d.h. $f(x) = f(y) = 0$. Dann ist

$$f(\lambda x + y) = \lambda f(x) + f(y) = \lambda \cdot 0 + 0 = 0,$$

also ist $\lambda x + y \in \mathrm{Ker}(f)$. Sind andererseits $x, y \in$ Bild(f) gegeben, d.h. gibt es $u, v \in V$ mit $f(u) = x$ und $f(v) = w$, so gilt

$$f(\lambda u + v) = \lambda f(u) + f(v) = \lambda \cdot x + y$$

und demnach ist $\lambda x + y$ ebenfalls im Bild von f. □

Mit dem Dimensionbegriff für Vektorräume und dieser Proposition können wir folgenden Begriff einführen, der ein Maß für die Nichttrivialität, die „Reichhaltigkeit" einer lineare Abbildung ist.

Definition 4.10 *Der* Rang *einer linearen Abbildung f ist die Dimension des Bildes.*

Aus der Definition der Dimension eines Vektorraums und aus dem Satz 4.7 folgt unmittelbar:

Proposition 4.11 *Ist die lineare Abbildung f durch Matrixmultiplikation $f(x) = A \cdot x$ gegeben, so ist der Rang von f gleich der Maximalzahl linear unabhängiger Spalten von A.*

Aufgrund dieser Proposition definiert man auch den *Spaltenrang* (oder kurz *Rang*) einer Matrix als die Maximalzahl linear unabhängiger Spalten von A.

Die verschiedenen Dimensionsbegriffe stehen in der folgenden fundamentalen Beziehung zueinander.

Satz 4.12 (Dimensionssatz für lineare Abbildungen) *Ist $f \colon V \longrightarrow W$ linear und V endlichdimensional, so gilt*

$$\mathit{Rang}(f) + \dim(\mathrm{Ker}(f)) = \dim(V).$$

Wir illustrieren den Satz zunächst am Beispiel $m = n = 2$. Es gibt drei Möglichkeiten für die Gestalt der linearen Abbildung $f(x) = A \cdot x$, nämlich

$$A_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} a & ra \\ c & rc \end{pmatrix}, \quad A_3 = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (ad - bc \neq 0)\,. \tag{4.3}$$

Die Nullmatrix A_1 ist die einzige Matrix, mit der Eigenschaft, dass das Bild der Nullraum, also von Dimension Null ist. In diesem Fall ist der Kern der ganze Ausgangsraum V und wegen $2 + 0 = 2$ ist der Satz in diesem Fall richtig. Ist der Bildraum eindimensional, so ist eine Spalte von A nicht Null, sagen wir die erste. Die zweite Spalte von A ist dann zwingend ein Vielfaches der ersten Spalte, sagen wir das r-fache, wie in A_2 angegeben. In diesem Fall ist

$$\mathrm{Ker}(f) \;=\; \{ \begin{pmatrix} x \\ y \end{pmatrix} \,|\, x = -ry \} \;=\; [\begin{pmatrix} -r \\ 1 \end{pmatrix}]\,,$$

also ein eindimensionaler Vektorraum. Der Dimensionssatz ist wegen $1 + 1 = 2$ auch in diesem Fall richtig.

Es bleibt noch der Fall eines zweidimensionalen Bildraums zu diskutieren. Wir behaupten also, dass für alle Matrizen, die nicht von der Gestalt A_1 oder A_2 (oder die Vertauschung der beiden Spalten hiervon) sind, die Bindung $ad - bc \neq 0$ gilt, dass für Matrizen der Gestalt A_1 und A_2 die Bedingung $ad - bc = 0$ gilt, und dass aus $ad - bc \neq 0$ folgt, dass der Bildraum wirklich zweidimensional ist. Die zweite dieser Aussagen ist ganz offensichtlich. Für die erste dieser Aussagen können wir annehmen, dass die Matrix der linearen Abbildung nicht Null ist (sonst sind wir im Fall A_1). Sei also ein Eintrag der ersten Spalte nicht Null. (Falls es diesen nicht gibt, vertauschen wir die Rollen der ersten und zweiten Spalte im folgenden Argument.) Ist $a \neq 0$, so definieren wir $r = b/a$ und dann ist $d = bc/a = rc$, wie behauptet. Analog, ist $c \neq 0$, so definieren wir $r = d/c$ und dann ist $b = ad/c = ra$, wie behauptet. Für die dritte dieser Aussagen verwenden wir schließlich, dass im Fall $ad - bc \neq 0$ die Gleichung

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \\ s \end{pmatrix}$$

lösbar ist, und zwar durch

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} r \\ s \end{pmatrix} \quad \text{(nachrechnen!)} \tag{4.4}$$

Im Fall $ad - bc = 0$ folgt aus diesen Zeilen auch, dass $\mathrm{Ker}\, f = 0$. Damit ist der Dimensionssatz in diesem Fall die Aussage $0 + 2 = 2$.

Beweis von Satz 4.12 (***) : Für $\dim V = 0$ ist die Behauptung $0 + 0 = 0$, also trivialerweise richtig. Wir wählen eine Basis $\{b_1, \ldots, b_d\}$ des Kerns von f und ergänzen sie mittels

$\{c_{d+1}, \ldots, c_n\}$ zu einer Basis von V. Zu zeigen ist also, dass Rang$(f) = n - d$ ist. Dies wollen wir nachweisen, indem wir zeigen, dass $\{f(c_{d+1}), \ldots, f(c_n)\}$ eine Basis von Bild(f) ist. Sei also $x \in$ Bild(f) und $f(u) = x$. Dann hat u die Basisdarstellung.

$$u = \sum_{i=1}^{d} \alpha_i b_i + \sum_{i=d+1}^{n} \alpha_i c_i.$$

Da $f\Big(\sum \alpha_i b_i\Big) = 0$ ist, gilt $x = f(u) = f\left(\sum_{i=d+1}^n \alpha_i c_i\right)$, also erzeugen $\{f(c_{d+1}), \ldots, f(c_n)\}$ das Bild von f. Diese Menge ist auch linear unabhängig, denn falls es eine Linearkombination $0 = \sum_{i=d+1}^n \lambda_i f(c_i)$ gibt, so ist aufgrund der Linearität $f\left(\sum_{i=d+1}^n \alpha_i c_i\right) = 0$, also $\sum_{i=d+1}^n \alpha_i c_i \in \mathrm{Ker}(f)$. Anders gesagt, es gibt $\alpha_i = \lambda_i \in K$ für $i = 1, \ldots, d$ mit

$$\sum_{i=d+1}^{n} \lambda_i c_i = \sum_{i=1}^{d} \lambda_i b_i.$$

Da aber $\{b_1, \ldots, b_d, c_{d+1}, \ldots, c_n\}$ eine Basis von V war, müssen alle $\lambda_i = 0$ sein. Also war die Linearkombination trivial und dies zeigt die Behauptung. □

Die Eigenschaften des Rang-Begriffs helfen uns im nächsten Abschnitt die Strukturaussagen über die Lösungsmenge eines linearen Gleichungssystems zu beweisen.

Als Hilfsbegriff definiert man auch den *Zeilenrang* einer Matrix als die Maximalzahl linear unabhängiger Zeilen. Bei einer Matrix in spezieller Zeilenstufenform 1.8 ist offensichtlich der Zeilenrang gleich dem Spaltenrang gleich der Anzahl r der Pivotlemente. Offensichtlich ändern Spaltenvertauschungen den Rang einer Matrix nicht. Wir verwenden ohne Beweis, dass die elementaren Zeilenoperationen des Gauss-Algorithmus den Rang einer Matrix ebenfalls nicht ändern. Zusammengesetzt beweisen diese Beobachtungen

Satz 4.13 *Der Zeilenrang einer Matrix A ist gleich dem Spaltenrang von A.*

Aufgrund dieses Satzes sprechen wir ab sofort nur noch vom Rang der Matrix A. Als Beispiel für den Satz betrachten wir die Matrix A_2 aus (4.3), wobei $(a, c) \neq (0, 0)$. Wir hatten dort festgestellt, dass der Spaltenrang gleich 1 ist. Wir prüfen nochmal von Hand nach, dass auch der Zeilenrang gleich 1 ist. Die Voraussetzung über das Nichtverschwinden von (a, c) impliziert, dass der Zeilenrang nicht Null ist. Weiter ist $c(a, ra) - a(c, rc) = 0$ eine nichttriviale Linearkombination der zwei Zeilenvektoren. Also sind diese linear abhängig und der Zeilenrang echt kleiner als zwei.

4.4 Anwendung auf lineare Gleichungssysteme

Zunächst halten wir fest, dass mit Hilfe der Matrixmultiplikation die Schreibweise von linearen Gleichungssystemen aus Abschnitt 1 einen weiteren Sinn bekommt. Die Lösungsmenge $\mathbb{L}_{(A,b)}$ des linearen Gleichungssystems

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

wie in (1.7) besteht genau aus den $x \in K^n$ mit $A \cdot x = b$, wobei wie in Abschnitt

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Im Fall $b = 0$ ist die Lösungsmenge also genau der Kern $\mathrm{Ker}(f)$ der linearen Abbildung $f : K^n \to K^m$ gegeben durch die Matrixmultiplikation $f(x) = A \cdot x$. Nach Proposition 4.9 ist $\mathrm{Ker}(f) = \mathbb{L}_{(A|0)}$ ein Untervektorraum und zwar ein Untervektorraum der Dimension $n - \mathrm{Rang}(A)$ nach dem Dimensionssatz 4.12. Wir haben bereits oben verwendet, dass elementare Zeilenumformungen und Spaltenvertauschungen den Rang von A nicht verändern. Also ist $\mathrm{Rang}(A) = r$ die Anzahl der Pivotelemente und $\dim \mathrm{Ker}(f) = n - r$. Die Lösungsmenge, die in Satz 1.6 angegeben wird, besteht aus allen Linearkombinationen von $n - r$ Elementen, von denen wir bereits geprüft haben, dass sie in $\mathrm{Ker}(f)$ liegen. Da diese $n - r$ Elemente linear unabhängig sind (man beachte die Zeilen mit den -1en) ist die Lösungsmenge aus Satz 1.6 in der Tat ein $n-r$-dimensionaler Vektorraum. Damit ist der Beweis von Satz 1.6 vollständig.

Wir kommen nun nochmal auf inhomogene lineare Gleichungssysteme, d.h. auf den Fall $b \neq 0$ zurück. Sei $f : \mathbb{R}^n \to \mathbb{R}^m$ die lineare Abbildung $x \mapsto A \cdot x$ und x_0 eine („spezielle“) Lösung des linearen Gleichungssystem, d.h. $f(x_0) = Ax_0 = b$. Ist x_1 eine weitere Lösung des linearen Gleichungssystem, so gilt $f(x_1) = b$ und damit nach Definition einer linearen Abbildung

$$f(x_1 - x_0) = b - b = 0$$

mit anderen Worten $x_1 - x_0$ ist eine Lösung des zugehörigen homogenen Gleichungssystems. Ist umgekehrt x_h eine Lösung des zugehörigen homogenen Gleichungssystems, so hat $x_0 + x_h$ die Eigenschaft, dass

$$f(x_0 + x_h) = f(x_0) + f(x_h) = 0 + b = b$$

also eine Lösung des inhomogenen Gleichungssystems ist. Diese beiden Beobachtungen zusammen, rechtfertigen die bisher noch nicht bewiesene zweite Behauptung in Satz 1.7.

Mit Hilfe der Vokabel „Rang“ können wir auch die erste Behauptung dieses Satzes 1.7 nochmal kompakt wie folgt formulieren.

Satz 4.14 *Ein lineares Gleichungssystem $Ax = b$ besitzt genau dann eine Lösung $x \in \mathbb{R}^n$, wenn $\operatorname{Rang} A = \operatorname{Rang}(A|b)$, d.h. wenn der Rang der Matrix A gleich dem Rang der erweiterten Matrix $(A|b)$ ist.*

5 Längen und Winkel

Die Längen und Winkel im $\mathbb{R}^2$ oder $\mathbb{R}^3$ kann man aus den Koordinaten eines Punktes unter Verwendung des Satzes von Pythagoras bzw. aus der Definition der trigonometrischen Funktionen bestimmen. Wir fassen diese Rechnungen zusammen und benutzen die Motivation, auf einem beliebigen Vektorraum den Begriff eines Skalarprodukts zu definieren und damit Längen und Winkel zu bestimmen.

Der $\mathbb{R}^2$ oder $\mathbb{R}^3$ sei mit einem kartesischen Koordinatensystem versehen, d.h. die Achsen stehen senkrecht aufeinander, und die Vektoren der Standardbasis mögen die Länge 1 haben. Nach dem Satz von Pythagoras hat die Stecke $\overline{OP}$ in Abbildung 5.1) die Länge $\sqrt{x_1^2 + x_2^2}$. Wir fassen P als Vektor in $\mathbb{R}$ auf und definieren dessen euklidische Norm als

$$||P|| = ||(x_1, x_2)|| := |\overline{OP}| = \sqrt{x_1^2 + x_2^2}\,.$$

Im $\mathbb{R}^3$ definiert man entsprechend die euklidische Norm von $P = (x_1, x_2, x_3)$ als $||P|| :=$

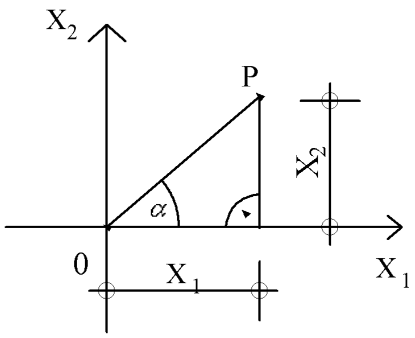

Abbildung 5.1: Länge der Strecke $\overline{OP}$

$|\overline{OP}| = \sqrt{x_1^2 + x_2^2 + x_3^2}$. Damit ist auch der Abstand zweier Punkte im $\mathbb{R}^3$ definiert als

$$|\overline{PQ}| := ||P - Q|| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}\,.$$

Im $\mathbb{R}^2$ ist der Winkel α zwischen $P = (x_1, x_2)$ als Vektor und der ersten Koordinatenachse bestimmt durch $||P|| \cdot \cos(\alpha) = x_1$. Der Winkel zwischen den Vektoren $P, Q \in \mathbb{R}^2$ im Spezialfall, dass Q auf der ersten Koordinatenachse mit Koordinaten $(y_1, 0)$ liegt, ist also

$$\|P\| \cdot \|Q\| \cos\alpha = x_1 y_1.$$

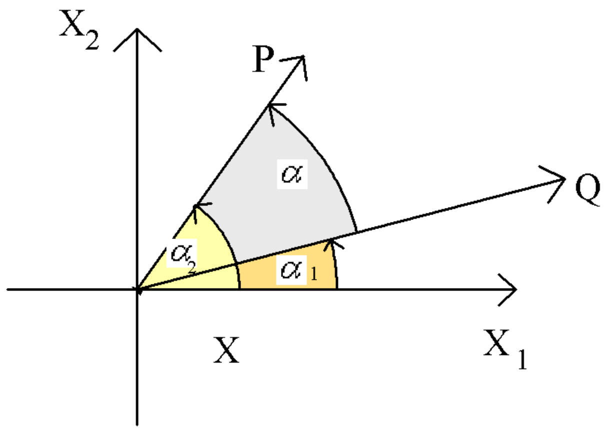

Abbildung 5.2: Bestimmung des Winkels POQ

Für zwei Punkte $P = (x_1, x_2), Q = (y_1, y_2) \in \mathbb{R}^2$ in beliebiger Lage ist α bestimmt als $\alpha_2 - \alpha_1$, wobei (vgl. Abbildung 5.2)

$$\begin{aligned} \cos(\alpha) = \cos(\alpha_2 - \alpha_1) &= \cos\alpha_1 \cos\alpha_2 + \sin\alpha_1 \sin\alpha_2 \\ &= \frac{x_1 y_1 + x_2 y_2}{\|P\| \cdot \|Q\|} \end{aligned}$$

nach dem Additionstheorem für $\cos$ und der Definition der trigonometischen Funktionen. Im $\mathbb{R}^3$ rechnet man nach, indem man eine Ebene durch 0, P und Q legt, dass

$$\cos \measuredangle(P, Q) = \frac{x_1 y_1 + x_2 y_2 + x_3 y_3}{\|P\| \cdot \|Q\|} \,.$$

Wir definieren nun für alle Vektoren $x = (x_1, \ldots, x_n),\ \ y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ das *Standard-Skalarprodukt*:

$$\langle x, y \rangle \ = \ x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

und die zugehörige *euklidischen Norm*

$$||x||^2 = \langle x, x \rangle, \text{ also } ||x|| = \sqrt{\langle x, x \rangle},$$

wobei die Quadratwurzel stets nicht-negativ gewählt wird. Die Längen bestimmen die Winkel eindeutig (vgl. Abbildung 5.3) vermöge des Kosinussatzes

$$||x - y||^2 = ||x||^2 + ||y||^2 - 2||x|| \cdot ||y|| \cos\alpha \,.$$

Verwendet man, dass das Standard-Skalarprodukt bilinear und symmetrisch (siehe unten) ist, so folgt aus

$$||x - y||^2 = \langle x - y, x - y \rangle = \langle x, x \rangle - 2\langle x, y \rangle + \langle y, y \rangle = ||x||^2 + ||y||^2 - 2\langle x, y \rangle \,,$$

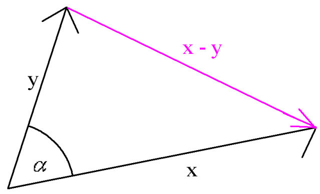

Abbildung 5.3: Vektor zur Verbindungsstrecke

also

$$\cos \measuredangle(x, y) = \frac{< x, y >}{||x|| \cdot ||y||}\,.$$

Das Standard-Skalarprodukt ist Beispiel einer positiv definiten symmetrische Bilinearform im Sinne der folgenden Definition. Sei V ein Vektorraum.

Definition 5.1 *Eine* Bilinearform *auf V ist eine Abbildung*

$$F\colon V \times V \longrightarrow \mathbb{R}, \quad (v, w) \mapsto F(v, w),$$

die in beiden Argumenten linear ist, d.h. es gilt für alle $v_1, v_2, w_1, w_2 \in V,\ \alpha, \beta \in \mathbb{R}$:

$$\begin{aligned} F(\alpha v_1 + \beta v_2, w_1) &= \alpha\, F(v_1, w_1) + \beta\, F(v_2, w_1) \\ F(v_1, \alpha w_1 + \beta w_2) &= \alpha\, F(v_1, w_1) + \beta\, F(v_1, w_2). \end{aligned}$$

Die Bilinearform heißt symmetrisch*, falls für alle $v, w \in V$ gilt: $F(v, w) = F(w, v)$. Wir nennen eine Bilinearform F* positiv definit*, wenn $F(a, a) > 0$ für alle $v \in V \setminus \{0\}$ ist. Eine positive definite, symmetrische Bilinearform wird* Skalarprodukt *genannt.*

Dann definiert man die *Länge* oder *Norm* $||x||$ eines Vektors $x \in V$ via $||x||^2 = \langle x, x\rangle$ und den *Winkel* zwischen x und y für $x, y \in V \setminus \{0\}$ durch

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

Hierfür ist zu zeigen $|\langle x, y\rangle| \leq ||x|| \cdot ||y||$, sonst wäre der Winkel nicht definiert, da $\cos$ nur Werte in $[-1, 1]$ annimmt! Dafür sorgt folgender Satz.

Satz 5.2 *(Cauchy-Schwarz-Ungleichung). In einem Vektorraum V mit Skalarprodukt gilt für alle $v, w \in V$ die Abschätzung*

$$|\langle v, w\rangle| \leq \|v\| \cdot \|w\|.$$

Es gilt Gleichheit genau dann, wenn v und w linear abhängig sind.

Beweis : Ist $v = 0$, so gilt offenbar Gleichheit. Sind v, w linear abhängig und $v \neq 0$, so gibt es $\lambda \in \mathbb{R}$ mit $w = \lambda \cdot v$, also ist $\langle w, w\rangle = \lambda^2 \langle v, v\rangle$ und $\langle v, w\rangle = \lambda\langle v, v\rangle$. Zusammen multipliziert erhält man also

$$|\langle v, w\rangle|^2 = \lambda^2 \|v\|^2 \cdot \langle v, v\rangle = \|v\|^2 \|w\|^2$$

und durch Quadratwurzelziehen die Gleichheit in der Cauchy-Schwarz-Ungleichung.
Sind v, w linear unabhängig, so ist $v - \lambda w \neq 0$ für alle $\lambda \in \mathbb{R}$ und $\langle w, w\rangle > 0$, da $w \neq 0$. Also

$$0 < \langle v - \lambda w, v - \lambda w\rangle = \langle v, v\rangle - 2\lambda\langle w, v\rangle + \lambda^2\langle w, w\rangle.$$

Nimmt man speziell $\lambda = \frac{\langle v,w\rangle}{\langle w,w\rangle}$, so erhält man

$$0 < \langle v, v\rangle - \frac{\langle v, w\rangle^2}{\langle w, w\rangle}.$$

Durchmultiplizieren mit $\langle w, w\rangle$ ergibt

$$0 < \langle v, v\rangle \cdot \langle w, w\rangle - \langle v, w\rangle^2 = \|v\|^2 \cdot \|w\|^2 - \langle v, w\rangle^2$$

und damit die Behauptung. □

Aus der Definition des Winkels folgt insbesondere, dass x und y genau dann aufeinander senkrecht stehen, wenn $\langle x, y\rangle = 0$ ist.

Satz 5.3 *(**Höhenschnittpunkt** im Dreieck) Die drei Höhenlinien im Dreieck schneiden sich in einem Punkt.*

Beweis : Wir bezeichnen im Dreieck ABC die Fußpunkte der Höhen mit A_0, B_0 und C_0. Des Weiteren wählen wir das Koordinatensystem so, dass der Schnittpunkt der Höhe h_a (d.h. die Verbindungsstrecke AA_0) und der Höhe h_b (d.h. die Verbindungsstrecke BB_0) im Ursprung liegen (vgl. Abbildung 5.4). Dann ist aufgrund der rechten Winkel

$$\langle A - A_0, B - C\rangle = \langle A, B - C\rangle = 0 = \langle B - B_0, A - C\rangle = \langle B, A - C\rangle$$

(da die Gerade durch A und A_0, bzw. B und B_0 gleich der Geraden durch A bzw. B und den Ursprung sind), also folgt aus der Bilinearität des Skalarproduktes

$$\langle A, B - C\rangle = \langle A, B\rangle - \langle A, C\rangle = \langle B, A - C\rangle = \langle B, A\rangle - \langle B, C\rangle$$

und wegen Symmetrie somit $\langle A, C\rangle = \langle B, C\rangle$, bzw. $\langle A - B, C\rangle = 0$. Damit steht aber die Gerade durch C und 0 senkrecht auf der Seite AB und muss daher die Höhe h_c enthalten. Insbesondere liegt also 0 auf h_c und damit schneiden sich alle drei Höhen in diesem Punkt. □

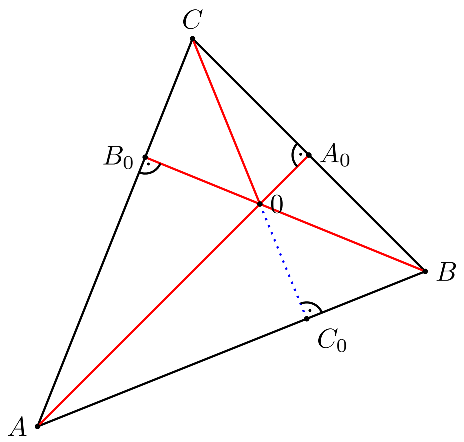

Abbildung 5.4: Höhenschnittpunkt im Dreieck

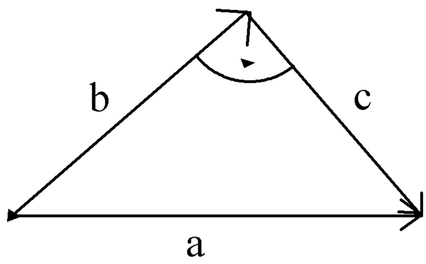

Abbildung 5.5: Rechwinkliges Dreieck zum Satz von Pythagoras

Satz 5.4 *(Pythagoras) In einem Vektorraum mit Skalarprodukt* $(V, \langle\cdot,\cdot\rangle)$ *gilt für Vektoren* a, b, c *in einem Dreieck (d.h. mit* $a = b + c$*), bei dem der* a *gegenüberliegende Winkel ein rechter Winkel ist, die Beziehung*

$$||a||^2 \; = \; ||b||^2 + ||c||^2 \, .$$

Beweis : Wegen $a = b + c$ als Vektoren und $b \perp c$ ist

$$||a||^2 = \langle b + c, b + c\rangle = ||b||^2 + ||c||^2 + \underbrace{2\langle b, c\rangle}_{=0}$$

und damit folgt die Behauptung. □

Diesen Beweis sollte man nicht für den gewöhnlichen Satz des Pythagoras für Punkte im $\mathbb{R}^2$ zitieren, da dies ein Ringschluss wäre: Wir haben oben in der Definition des Winkelbegriffs mit Hilfe des Standardskalarprodukts den Kosinussatz und damit den Satz von Pythagoras bereits verwendet!

Skalarprodukte erlauben auch eine Darstellung von Geraden im $\mathbb{R}^2$ und Ebenen im $\mathbb{R}^3$ zu geben, in der man den Abstand von Punkten zu dieser Geraden schnell berechnen kann. Ab sofort seien $\mathbb{R}^2$ und $\mathbb{R}^3$ mit den Standardskalarprodukten versehen.

Definition 5.5 *Eine Gerade $g = g(w, r) \subset \mathbb{R}^2$ ist in* Hesse'scher Normalform, *falls sie als*

$$g = \{v \in \mathbb{R}^2 : \langle w, v\rangle = r\}$$

mit $w \in \mathbb{R}^2$, $r \in \mathbb{R}$ und $||w|| = 1$ geschrieben ist. Eine Ebene $E = E(w, r) \subset \mathbb{R}^3$ ist in Hesse'scher Normalform, *falls sie als*

$$E = \{v \in \mathbb{R}^3 : \langle w, v\rangle = r\}$$

mit $w \in \mathbb{R}^2$, $r \in \mathbb{R}$ und $||w|| = 1$ geschrieben ist.

Zunächst halten wir fest, dass jede Gerade $g \subset \mathbb{R}^2$ und jede Ebene $E \subset \mathbb{R}^3$ in Hesse'scher Normalform" geschrieben werden kann. Wir verwenden dabei, dass jede Gerade durch eine Geradengleichung

$$g = \left\{v = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : ax + by = d\right\} \tag{5.1}$$

mit $(a, b) \neq (0, 0)$ geschrieben werden kann. Unter Verwendung von $\widetilde{w} = \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$ kann man also

$$g = \left\{v = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : \langle \widetilde{w}, v\rangle = d\right\}$$

schreiben. Der Vektor $\widetilde{w}$ hat noch nicht die geforderte Norm eins. Wir setzen also $w = \widetilde{w}/||\widetilde{w}|| = \widetilde{w}/\sqrt{a^2 + b^2}$ und $r = d/\sqrt{a^2 + b^2}$. Dann ist

$$g = \left\{v = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : \langle w, v\rangle = r\right\},$$

wie verlangt. Ganz analog schreibt man eine Ebene in $\mathbb{R}^3$ als Hyperebenengleichung

$$E = \left\{v = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : ax + by + cz = d\right\} \tag{5.2}$$

mit Hilfe von $\widetilde{w} = \left(\begin{smallmatrix} a \\ b \\ c \end{smallmatrix}\right)$ als

$$E = \left\{v = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \langle \widetilde{w}, v\rangle = d\right\}$$

und schliesslich mit Hilfe von $w = \widetilde{w}/||\widetilde{w}|| = \widetilde{w}/\sqrt{a^2 + b^2 + c^2}$ und $r = d/\sqrt{a^2 + b^2 + c^2}$ als

$$E = \left\{v = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \langle w, v\rangle = r\right\}.$$

Satz 5.6 *Ist die Gerade $g = g(w, r)$ in Hessescher Normalform gegeben und $P \in \mathbb{R}^2$, so ist der Abstand von P zu g gleich $|\langle w, P\rangle - r|$.*
Ist die Ebene $E = E(w, r)$ in Hessescher Normalform gegeben und $P \in \mathbb{R}^3$, so ist der Abstand von P zu E gleich $|\langle w, P\rangle - r|$.

Beweis : Der Abstand zwischen einem Punkt P und einer Geraden g ist $d(P, g) = ||P - Q||$, wobei Q derjenige Punkt auf g ist, sodass $P - Q$ senkrecht auf g steht. Es ist $\langle w, Q\rangle = r$, da $Q \in g$ und außerdem ist $P - Q$ parallel zu w, der zugehörige Vektor also ein Vielfaches von w. Da zudem $||w|| = \langle w, w\rangle = 1$ laut Definition der Hesseschen Normalform ist, folgt mit der Hilfe der Gleichheitsaussage in der Cauchy-Schwarzschen Ungleichung

$$||P - Q|| = ||P - Q|| \cdot \langle w, w\rangle = |\langle w, P - Q\rangle| = |\langle w, P\rangle - \langle w, Q\rangle| = |\langle w, P\rangle - r| \,.$$

Das Argument im Fall der Ebene im $\mathbb{R}^3$ ist identisch. □

Wir schließen diesen Abschnitt mit dem Nachweis, dass die oben verwendete Darstellung von Geraden und Ebenen sich in die (vielleicht) gebräuchlichere Darstellung von Geraden und Ebenen mittels *Aufpunkt* und *Richtungsvektor(en)* (vgl. Abbildung 5.6) umrechnen lässt.

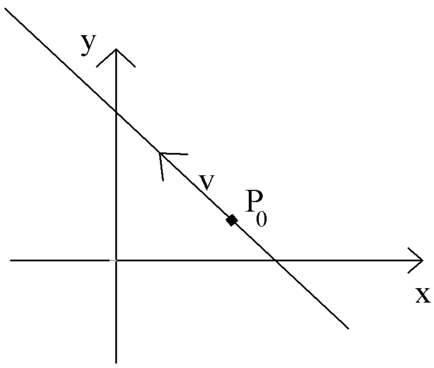

Abbildung 5.6: Eine Gerade gegeben durch Aufpunkt P_0 und Richtungsvektor v

Die Behauptung (oder Definition, je nach Standpunkt) ist also, dass eine Gerade g im $\mathbb{R}^2$ sich als Punktmenge

$$g \;=\; P_0 + [v] \;:=\; \{P_0 + \lambda v \,|\, \lambda \in \mathbb{R}\}, \tag{5.3}$$

schreiben lässt, wobei P_0 ein Punkt („Aufpunkt“) auf der Geraden und $v \neq 0$ ein Differenzvektor zweier verschiedener Punkte auf g ist. Analog ist eine Ebene E im $\mathbb{R}^3$ als Punktmenge

$$E \;=\; P_0 + [v_1, v_2] \;:=\; \{P_0 + \lambda_1 v_1 + \lambda_2 v_2 \,|\, \lambda_1, \lambda_2 \in \mathbb{R}\}, \tag{5.4}$$

mit $\{v_1, v_2\}$ linear unabhängig gegeben.

Beweis : Sei eine Gerade g wie in (5.1) gegeben. Dann folgt aus der Struktur der Lösungsmenge linearer Gleichungssysteme wegen $(a, b) \neq 0$, dass g die Gestalt (5.3) hat. (Konkreter, ist z.B. $a \neq 0$, so leistet $P_0 = (r/a, 0)$ und $v = (-b, a)$ das Verlangte.)

Ist eine Ebene wie in (5.2) gegeben, so folgt ebenso aus der Struktur der Lösungsmenge linearer Gleichungssysteme wegen $(a, b, c) \neq 0$, dass g die Gestalt (5.3) hat. (Konkreter, ist z.B. $a \neq 0$, so leistet $P_0 = (r/a, 0, 0)$ und $v_1 = (-b, a, 0)$, $v_2 = (-c, 0, a)$ das Verlangte.)

Ist umgekehrt die Gerade g wie in (5.3) gegeben, sagen wir mit $v = \left(\begin{smallmatrix} v_x \\ v_y \end{smallmatrix}\right)$ und $P_0 = \left(\begin{smallmatrix} p_x \\ p_y \end{smallmatrix}\right)$, so hat g die Darstellung (5.1) mit $a = -v_y$, $b = v_x$ und $r = ap_y + bp_x$, wie man leicht nachrechnet.

Ist schließlich die Ebene E wie in (5.4) gegeben, so bestimmt man durch Lösen eines linearen Gleichungssystems einen Vektor $w = \left(\begin{smallmatrix} v_x \\ v_y \\ v_z \end{smallmatrix}\right)$, der senkrecht auf $[v_1, v_2]$ steht und wir normieren w auf $||w|| = 1$. (Einen solchen Vektor gibt es, da „senkrecht stehen“ nur zwei Bedingungen an die drei Komponenten stellt, und w ist bis auf das Vorzeichen eindeutig, da $\{v_1, v_2\}$ linear unabhängig ist.) Man setzt nun $r = \langle w, P_0 \rangle$ und rechnet wiederum nach, dass die Ebene nun auch durch (5.1) gegeben ist. $\square$

6 Bewegungsgruppen und Symmetrie

In Abschnitt 2 haben wir den Begriff abelsche Gruppe als Teil des Vektorraumbegriffs eingeführt. Wir isolieren hieraus nochmals die Axiome einer Gruppe, da wir zwei Beispiele von Gruppen detaillierter untersuchen wollen, Permutationsgruppen und Bewegungsgruppen.

Definition 6.1 *Eine Menge G heißt* Gruppe, *falls es eine Abbildung*

$$\circ \colon G \times G \to G$$

gibt, die folgenden Eigenschaften genügt

(GN) es gibt ein Element $e \in G$, so dass für jedes $x \in G$ gilt: $\circ(x, 0) = x$ (Neutralelement);

(GI) es gibt zu jedem $x \in G$ ein Element $y \in G$ mit $\circ(x, y) = e$; wir schreiben $x^{-1} := y$ (Inverses);

(GA) für alle $x, y, z \in G$ gilt: $\circ(x, \circ(y, z)) = \circ(\circ(x, y), z)$ (Assoziativität).

In diesem Fall definieren wir $x \circ y := \circ(x, y)$ für $x, y \in G$.

Häufig wird „$\circ$“ als „$\cdot$“ geschrieben oder überhaupt weggelassen (z.B. bei Gruppen von Zahlen mit der Multiplikation als Verknüpfung). Dann schreibt man „1“ anstelle „e“. Bei abelschen (kommutativen) Gruppen schreibt man als Verknüpfung häufig „$+$“ anstelle von „$\cdot$“, (siehe Abschnitt 2) und „0“ statt „e“, sowie „$-x$“ statt „x^{-1}“.

Man kann zeigen, dass x^{-1} auch immer $x^{-1} \circ x = e$ erfüllt, man merke sich „Ein Rechtsinverses Element ist stets auch ein Linksinverses“. Ebenso kann man zeigen, dass $e \circ x = x$ für alle $x \in G$ („Eine Rechtseins ist auch immer eine Linkseins“), sowie dass $(xy)^{-1} = y^{-1}x^{-1}$ und $(x^{-1})^{-1} = x$ ist für alle $x, y \in G$. Gruppen sind z.B.

$$(\mathbb{Z}, +),\ (\mathbb{Q}, +),\ (\mathbb{R}, +),\ (\mathbb{Q}^*, \cdot),\ (\mathbb{R}^*, \cdot),$$

aber auch

$$(\{0\}, +),\ (\{1\}, \cdot),\ (\mathbb{Z}/m\mathbb{Z}, +),\ ((\mathbb{Z}/m\mathbb{Z})^*, \cdot).$$

Hierbei bezeichnet der Stern die Menge alle von Null verschiedenen Elemente in $\mathbb{Q}$ bzw. $\mathbb{Z}/m\mathbb{Z}$. Beispiele von Verknüpfungen, die keine Gruppe definieren sind

$$(\mathbb{N}, +),\ (\mathbb{Z}, \cdot),\ (\mathbb{R}, \cdot).$$

Bisher waren das alles abelsche Gruppen. Nicht-abelsche Gruppen erhält man am leichtesten auf dem Weg über Gruppen von Bijektionen $f : M \longrightarrow M$ mit Hintereinanderausführung als Verknüpfung.

Definition 6.2 *Die* Permutationsgruppe *oder* symmetrische Gruppe S_n *besteht aus den Bijektionen* $\{1, 2, \ldots, n\} \longrightarrow \{1, 2, \ldots, n\}$ *von* n *Objekten mit der Hintereinanderausführung als Verknüpfung.*

Die S_n ist in der Tat eine Gruppe mit der Identität als neutralem Element und der inversen Bijektion als inversem Element. Die S_n ist aber keine abelsche Gruppe. Beispielsweise ist

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

Die Klammerschreibweise symbolisiert die Permutationen

$$\begin{array}{ccc} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 2 & 1 & 3 \end{array} \circ \begin{array}{ccc} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 3 & 2 & 1 \end{array} = \begin{array}{ccc} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 3 & 1 & 2 \end{array}$$

Anderseits ist

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

und damit haben wir gezeigt, dass S_3 keine abelsche Gruppe ist.

Die Elemente links schreibt man kürzer in „Zykelschreibweise“ als $(12)(3)$ bzw. $(13)(2)$. Bekanntlich ist $|S_n| = n! =: \operatorname{ord} S_n$ die Mächtigkeit (man sagt auch *Gruppenordnung*) dieser endlichen Gruppe.

Ganz analog kann man $\mathrm{Bij}(M)$, die Menge alle Bijektionen $M \longrightarrow M$, zu einer Gruppe machen. Ist M groß (z.B. $M = \mathbb{R}^n$), so ist man meistens an Bijektionen mit speziellen Zusatzeigenschaften, d.h. an Untergruppen von $\mathrm{Bij}(M)$ interessiert.

Definition 6.3 *Sei $(G, \circ)$ eine Gruppe. Eine Teilmenge $H \subseteq G$ von G heißt* Untergruppe, *falls gilt*

i) Das neutrale Element $e \in H$.

ii) Für alle $g, h \in H$ liegt $g \circ h \in H$.
(„H ist bzgl. Verknüpfung abgeschlossen").

iii) Für alle $g \in H$ ist das Inverse $g^{-1} \in H$.

Zwei Beispiele sind für die lineare Algebra besonders wichtig: Die *allgemeine lineare Gruppe* $\mathrm{GL}(V)$ eines Vektorraums V, im Fall $V = \mathbb{R}^n$ auch $\mathrm{GL}_n(\mathbb{R})$ genannt, besteht aus allen bijektiven linearen Abbildungen $f : V \longrightarrow V$. Damit diese Menge zu einer Untergruppe von $\mathrm{Bij}(V)$ wird, ist zu zeigen, dass mit $f, g \in \mathrm{GL}_n(V)$ auch $f \circ g \in GL_n V$ und f^{-1} ist. Bijektivität der Verkettung zweier bijektiver Abbildungen und der Umkehrabbildung ist offensichtlich, es bleibt die Linearität zu zeigen. Wir rechnen nach, dass

$$(f \circ g)(v + w) = f(g(v + w)) = f(g(v) + g(w)) = f(g(v)) + f(g(w)) = (f \circ g)(v) + f \circ g(w)$$

und analog $(f \circ g)(rv) = r(f \circ g)(v)$ für alle $v, w \in V$ und für alle $r \in \mathbb{R}$. Außerdem ist zu prüfen, dass mit f auch f^{-1} linear ist. Seien dazu $x, y \in V$ beliebig. Dann gibt es, da f bijektiv ist, also insbesondere surjektiv ist, zwei Elemente $v, w \in V$ mit $f(v) = x$ und $f(w) = y$. Dann ist

$$\begin{aligned} f^{-1}(x + y) &= f^{-1}(f(v) + f(w)) = f^{-1}(f(v + w)) = v + w \\ &= f^{-1}(f(v)) \; + \; f^{-1}(f(w)) = f^{-1}(x) + f^{-1}(y)\,, \end{aligned}$$

und entsprechend rechnet man die Verträglichkeit von f^{-1} mit der Skalarmultiplikation nach.

Wir erinnern daran, dass wir in Abschnitt 4.2 jede lineare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$ (oder allgemeiner zwischen zwei beliebigen Vektorräumen V und W, in denen wir eine Basis gewählt haben) als Multiplikation mit einer Matrix schreiben können. Die genaue Formulierung davon ist Satz 4.7. Aus diesem Satz folgert man auch, dass die Verkettung von linearen Abbildungen durch das Produkt der zugehörigen Matrizen gegeben ist.

Die Matrix zur Identitätsabbildung ist die Einheitsmatrix E_n. Die Matrix zur Abbildung f^{-1} bezeichnet man mit A^{-1}. Um sie zu bestimmen, kann man, zu gegebenem A die Koeffizienten von A^{-1} als Unbekannte ansetzen. Dann fasst man $A \cdot A^{-1} = E_n$ als n lineare Gleichungssysteme (mit n Gleichungen und n Unbestimmten) mit der gleichen linken Seite A und n verschiedenen rechten Seiten (den Einheitsspaltenvektoren) auf und löst diese mit dem Gauss-Algorithmus. Für für $n = 2$ gibt es eine geschlossene Formel, die wir im Prinzip in Gleichung 4.4 schon gesehen haben, nämlich

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

Anschauliche Beispiele von Gruppen, sind *Symmetriegruppen* z.B. von Figuren in der Ebene oder im Raum. Eine *Symmetrie* bedeutet die Existenz einer nicht-trivialen (endlichen oder unendlichen Gruppe) G von euklidische Bewegungen (längentreuen Selbstabbildungen der Ebene oder des Raumes), so dass alle $f \in G$ die Figur wieder in sich selbst überführen, also dass $f(F) = F$ gilt. Ist beispielsweise F das regelmäßige Fünfeck (vgl. Abbildung 6.1), dann besteht G aus fünf Drehungen und fünf Spiegelungen. Nummeriert man die Ecken, lassen sich die $f \in G$ auch durch Permutationen aus S_5 beschreiben:

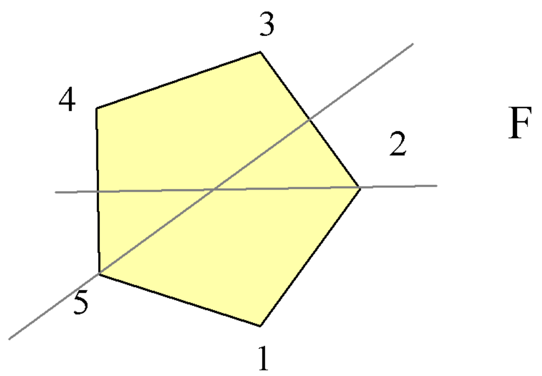

Abbildung 6.1: Symmetrien des regelmässigen Fünfecks

- Drehungen: $(1), (12345), (13524), (14253), (15432)$
- Spiegelungen: $(25)(34), (13)(45), (15)(24), (12)(35), (14)(23)$

Unendliche Symmetriegruppen kommen beim Kreis ($S = O_2$) (vgl. Abbildung 6.2) oder

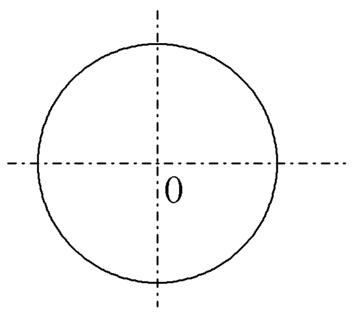

Abbildung 6.2: Kreis

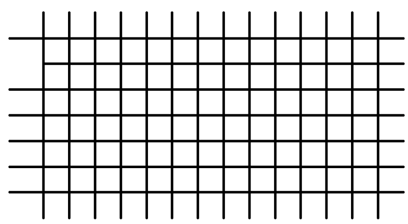

Abbildung 6.3: Parkettierung

bei Bänderfriesen und Parkettierungen vor. Bei der Parkettierung des $\mathbb{R}^2$ mit Quadraten (vgl. Abbildung 6.3) der Kantenlänge Eins wird die Symmetriegruppe S erzeugt von $\mathbb{Z}^2 = T$ (der Translationsuntergruppe) und der Symmetriegruppe $D_4 \subset O_2$ des Quadrats. Die Diedergruppe D_4 besteht aus 4 Drehungen und 4 Spiegelungen.

Wir untersuchen die Gruppe der euklidischen Bewegungen des $\mathbb{R}^n$ genauer. Diese enthält zum einen die Translationen. Durch Verkettung mit einer Translation können wir jede euklidischen Bewegungen zu einer Bewegung modifizieren, die den Ursprung festlässt. Die Untergruppe der euklidischen Bewegungen, die den Ursprung festlassen, kann man auch als orthogonale Gruppe wie folgt beschreiben.

Die *orthogonale Gruppe* $O_n \subset \mathrm{GL}_n(\mathbb{R})$ besteht aus jenen bijektiven linearen Abbildungen: $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, welche das Standardskalarprodukt invariant lassen, d.h. für welche

$$\langle f(v), f(w) \rangle = \langle v, w \rangle, \quad \text{für alle} \quad v, w \in V$$

gilt. Elemente der orthogonalen Gruppe nennt man auch *Isometrien*. Daraus folgt insbesondere, dass unter Isometrien Längen und Winkel erhalten bleiben.

Die Standardbasis des $\mathbb{R}^n$ hat bzgl. des Standardskalarprodukts die rechnerisch nützliche Eigenschaft $\langle e_i, e_j \rangle = \delta_{ij}$ (Kronecker-Delta). Allgemein nennt man eine Basis $B = \{b_1, \ldots, b_n\}$ eines Vektorraums $(V, \langle \cdot, \cdot \rangle$ mit Skalarprodukt, welche die Eigenschaft $\langle e_i, e_j \rangle = \delta_{ij}$ besitzt, eine *Orthonormalbasis*.

Aus der Definition einer Isometrie folgt sofort, dass die Bilder einer Orthonormalbasis, also z.B. $f(e_1), \ldots, f(e_n)$ wieder eine Orthonormalbasis bilden. Auch die Umkehrung gilt, und damit kann man die orthogonale Gruppe charakterisieren.

Satz 6.4 *Wenn $v_1, \ldots, v_n$ irgendeine Orthonormalbasis des $\mathbb{R}^n$ ist, gibt es genau eine lineare Abbildung $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ mit $e_i \longmapsto v_i$ für alle $i = 1, \ldots, n$. Diese ist zudem eine orthogonale Abbildung, d.h. $f \in O_n$.*

Beweis : Linearität und die Vorgabe $v_i = f(e_i)$ erzwingt, dass die Abbildung eindeutig ist, nämlich gegeben durch

$$f(x_1 e_1 + \ldots + x_n e_n) = x_1 v_1 + \ldots + x_n v_n \,.$$

Außerdem gilt

$$\begin{aligned} \langle f(x), f(y) \rangle &= \langle x_1 v_1 + \ldots + x_n v_n, y_1 v_1 + \ldots + y_n v_n \rangle \\ &= \sum_{i,j=1}^{n} x_i y_j \langle v_i, v_j \rangle \;=\; \sum_{i,j=1}^{n} x_i y_j \delta_{i,j} = \sum_{i=1}^{n} x_i y_i = \langle x, y \rangle \end{aligned}$$

aufgrund der Bilinearität des Skalarprodukts unter Eigenschaft der Orthonormalbasis. Daraus folgt die zweite Behauptung. □

Man erkennt also eine *orthogonale Matrix*, d.h. die Matrix einer orthogonalen Transformation $f \in O_n$ daran, dass ihre Spalten eine Orthogonalbasis bilden. Diese Orthogonalbasiseigenschaft drückt sich in Matrizenschreibweise aus, dass eine $n \times n$-Matrix B genau dann in

O_n ist, wenn $B^T \cdot B = E$ gilt, wobei B^T die zu B transponierte Matrix (gespiegelt an der Diagonalen von links oben nach rechts unten) beschreibt. Es gilt also

$$B \in O_n \iff B^T = B^{-1}.$$

Es bleibt noch zu zeigen, dass O_n wirklich eine Untergruppen von $\mathrm{GL}_n \mathbb{R}$ ist, also für alle $f, g \in O_n$ auch $f \circ g \in O_n$ und $f^{-1} \in O_n$. Für alle $v, w \in \mathbb{R}^n$ gilt

$$\langle f \circ g(v), f \circ g(w) \rangle = \langle f(g(v)), f(g(w)) \rangle = \langle g(v), g(w) \rangle = \langle v, w \rangle$$

und, indem wir $v = f(x)$ und $w = f(y)$ schreiben, erhalten wir

$$\langle f^{-1}(x), f^{-1}(y) \rangle = \langle f^{-1} f(v), f^{-1} f(w) \rangle = \langle v, w \rangle = \langle f(v), f(w) \rangle = \langle x, y \rangle.$$

Da offensichtlich $E_n \in O_n$ ist, sind alle Eigenschaften einer Untergruppe erfüllt.

Wir untersuchen den Fall der ebenen Isometrien $O_2 \subset \mathrm{GL}_2(\mathbb{R})$ im Detail. Die Eigenschaft

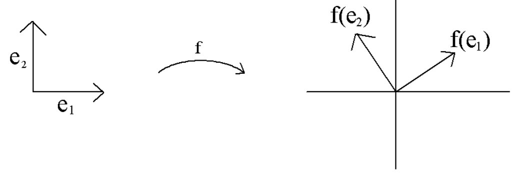

Abbildung 6.4: Drehung

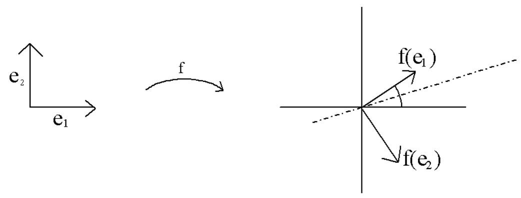

Abbildung 6.5: Spiegelung

für eine Matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ eine Isometrie, also ein Element von O_2 zu sein, ist äquivalent zu

$$a^2 + c^2 = 1, \quad b^2 + d^2 = 1, \quad \text{und} \quad ab + cd = 0.$$

Daraus folgt, dass es ein $\varphi \in [0, 2\pi]$ gibt, mit $a = \cos(\varphi)$ und $c = \sin(\varphi)$. Aus der dritten Gleichung folgt, dass (b, d) zu $(\sin(\varphi), -\cos(\varphi))$ proportional sein muss. Aus der zweiten Gleichung folgt, dass es nur zwei Lösungen gibt, nämlich der Fall I), dass $(b, d) = (-\sin(\varphi), \cos(\varphi))$ und der Fall II), dass $(b, d) = (-\sin(\varphi), \cos(\varphi))$, vgl. die Abbildungen 6.4 und 6.5.

Man beobachtet, dass im Fall I $ad - bc = 1$ gilt. Dies sind die sogenannten „eigentlichen“ oder „orientierungserhaltenden“ orthogonalen Transformationen (Drehungen). Im Fall II gilt $ad - bc = -1$. Dies sind die „uneigentlichen“ oder „orientierungsumkehrenden“ orthogonalen Transformationen (Spiegelungen).

7 Determinanten und Eigenwerte

Eine von vielen Möglichkeiten den Begriff des Eigenwerts und des Eigenvektors zu motivieren, besteht in der Klassifikation von euklidischer Bewegungen im $\mathbb{R}^3$. Wir haben im letzten Abschnitt die Untergruppen $T = (\mathbb{R}^3, +)$ der Translationen und die Untergruppe der orthogonalen Bewegungen O_3 kennengelernt. Die Translationen wollen wir nicht feiner klassifizieren, aber unter den orthogonalen Bewegungen gibt es eine natürliche Feinklassifikation. Diese besteht, neben der Identität, aus folgenden Klassen:

i) Spiegelungen an einem 2-dimensionalen Unterverktorraum.

ii) Drehungen um eine Achse (vgl. Abbildung 7.1, links);

iii) Kombination von i) und ii), d.h. die Spiegelung an einer Ebene und die Drehung um eine Achse, die senkrecht dazu steht (vgl. Abbildung 7.1, rechts).

iv) Eine Punktspiegelung am Ursprung.

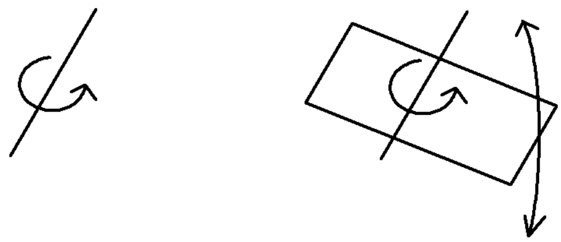

Abbildung 7.1: Drehung um eine Achse, mit und ohne Ebenenspiegelung

Nach der Besprechung von Eigenwerten können wir zeigen, dass diese Liste vollständig ist.

Definition 7.1 *Sei $f : V \longrightarrow V$ eine lineare Abbildung. Ein Element $0 \neq v \in V$ heißt* Eigenvektor *von f zum* Eigenwert $\lambda \in \mathbb{R}$, *wenn $f(v) = \lambda v$.*

Wir diskutieren diesen Begriff an den oben eingeführten drei Klassen orthogonaler Transformationen. Sei f eine solche Transformation. Ist $v \neq 0$ ein Element der Fixachse im Fall ii), so gilt (per Definition von 'Fix') $f(v) = v$. Also ist v ein Eigenvektor zum Eigenwert 1. Sind v, w zwei linear unabhängige Vektoren, die eine Fixebene im Fall i) aufspannen, so sind alle Elemente in $[v, w] \setminus \{0\}$ Eigenvektoren zum Eigenwert 1. Sei $u \neq 0$ ein Vektor, der dazu senkrecht steht, also gespiegelt wird. D.h. nach Definition von Spiegelung, dass $f(u) = -u$. Also ist u ein Eigenvektor zum Eigenwert -1

Bemerkung 7.2 Wenn $f \in O_n$, kommt nur $\lambda = \pm 1$ in Frage wegen

$$||f(v)|| = |\lambda| \cdot ||v|| = ||v||.$$

Mit dieser Bemerkung kann man die Vollständigkeit obiger Liste zu großen Teilen beweisen, siehe Satz 7.7 unten.

Wie findet man Eigenwerte und Eigenvektoren? Da bei den orthogonalen Matrizen nur $\lambda = \pm 1$ in Frage kommen, kann man sich hier auf das Lösen der zwei linearen Gleichungssysteme

$$Ax = x \quad \text{oder äquivalent} \quad (A - E)x = 0$$

und

$$Ax = -x \quad \text{oder äquivalent} \quad (A + E)x = 0$$

beschränken. Hat das erste lineare Gleichungssystem eine nicht-triviale Lösung (d.h. eine Lösung ungleich Null), so ist 1 ein Eigenwert. Analog, hat das zweite Gleichungssystem eine nicht-triviale Lösung so ist -1 ein Eigenwert. Ist f eine beliebige lineare Abbildung, und will man feststellen, ob λ ein Eigenwert ist, so muss man das lineare Gleichungssysteme $(A - \lambda E)x = 0$ lösen und testen, ob es eine nicht-triviale Lösung gibt. Die Menge aller Lösungen von $(A - \lambda E)x = 0$ wird auch der *Eigenraum* zu λ genannt. Offenbar ist der Eigenraum ein Untervektorraum von $\mathbb{R}^n$.

Aber wie findet man die Eigenwerte einer beliebigen linearen Abbildung? Für welche λ soll man versuchen, diese Gleichung zu lösen. Die Anwort lautet, dass man es für alle λ simultan tun sollte, mit Hilfe von Determinanten. Die folgende Definition ist eigentlich ein Satz mit Definition, d.h. wir behaupten auch (ohne Beweis) die Existenz und Eindeutigkeit einer solchen Abbildung.

Definition 7.3 *Eine* Determinante *ist eine Abbildung* $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$ *mit folgenden Eigenschaften.*

i) $\det E = 1$

ii) Die Determinante ist linear in den Zeilen, d.h. sind a_i die Zeilen einer Matrix, so ist für alle $i = 1, \dots, n$ und für alle $r \in \mathbb{R}$

$$\det \begin{pmatrix} a_1 \\ \cdots \\ ra_i + a_i' \\ \cdots \\ a_n \end{pmatrix} = \det \begin{pmatrix} a_1 \\ \cdots \\ a_i \\ \cdots \\ a_3 \end{pmatrix} + r \det \begin{pmatrix} a_1 \\ \cdots \\ a_i' \\ \cdots \\ a_3 \end{pmatrix}.$$

iii) Bei Vertauschung zweier Zeilen ändert sich das Vorzeichen.

Für kleine n kann man Determinanten leicht direkt angeben, und zwar ist für $n = 1$ einfach $\det(a) := a$. Für $n = 2$ ist

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc\,.$$

Schließlich ist für $n = 3$

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12},$$

was man sich merken kann, indem man die Diagonalen (mit Vorzeichen!) betrachtet, auf denen die Elemente stehen, die aufmultipliziert werden ('Sarrus-Regel').

Aus den Axiomen leitet man folgende Eigenschaften her.

Proposition 7.4 *i) Die Determinante einer Matrix ist gleich der Determinanten der transponierten Matrix, d.h.* $\det A = \det A^T$.

ii) Die Eigenschaften i)) bis iii)) gelten ebenso für Spalten wie für Zeilen.

iii) Die Determinante ändert sich nicht, wenn man das r-fache einer Zeile zu einer anderen addiert

iv) Wenn zwei Zeilen gleich sind, dann ist die Determinante gleich Null.

v) Die Determinante ist gleich Null, genau dann wenn der Rang der Matrix strikt kleiner als n ist.

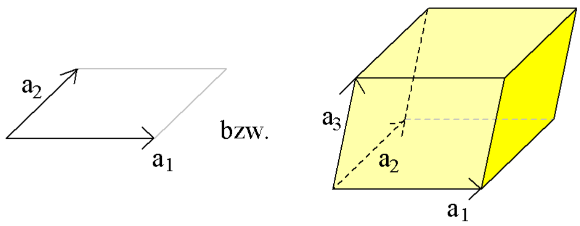

Abbildung 7.2: Determinante als Volumen eines Parallelogramms

Satz 7.5 *Der Betrag der Determinante ist das Volumen des Parallelotops, das von den Zeilen aufgespannt wird. D.h. für $n = 2$ und $n = 3$ seien a_i die Zeilen der Matrix A und*

$$P = \{xa_1 + ya_2 | x, y \in [0, 1]\} \quad bzw. \quad P = \{xa_1 + ya_2 + za_3 | x, y, z \in [0, 1]\}\,.$$

Dann gilt $|\det(A)| = vol(P)$.

Beweis : Die Behauptung stimmt offenbar für die Standardbasis und bleibt unter Addition eines Vielfachen einer Zeile auf eine andere Zeile, unter Multiplikation einer Zeile mit einem Faktor $r \in \mathbb{R}$ sowie unter Zeilenvertauschungen wahr. Das gleiche gilt für Transponieren und die entsprechenden Spaltenoperationen. Mit Hilfe solcher Operationen kann man nach den Überlegungen im Abschnitt zur Rangbestimmung jede Matrix von Rang n erreichen. Für alle anderen Matrizen sind beide Seiten in der Behauptung des Satzes gleich Null. □

Satz 7.6 *Eine reelle Zahl λ ist genau dann ein Eigenwert der $n \times n$-Matrix A, wenn λ Nullstelle des* charakteristischen Polynoms $\chi_A(x) = \det(A - xE)$ *ist.*

Speziell für $n = 2$ und

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{ist} \quad \chi_A(x) = \det \begin{pmatrix} a-x & b \\ c & d-x \end{pmatrix} \quad .$$

Beweis : Ist λ ein Eigenwert mit zugehörigem Eigenvektor $v \neq 0$, so hat das Gleichungssystem $Av = \lambda v$ eine nicht-triviale Lösung. Also ist $(A - \lambda E)v = 0$ ein homogenes lineares Gleichungssystem vom Rang kleiner n. Die bedeutet nach obiger Proposition, dass $\det(A - \lambda E) = 0$ ist, also λ eine Nullstelle von χ_A ist.

Diese Implikationen sind in Wirklichkeit alles Äquivalenzen. Ist also umgekehrt λ eine Nullstelle von χ_A, so ist $\det(A - \lambda E) = 0$ und damit der Rang von $A - \lambda E$ kleiner als n. Damit hat $(A-\lambda E)v = 0$ eine nicht-triviale Lösung $v \neq 0$. Dieser Vektor v ist ein Eigenvektor zum Eigenwert λ. □

Das charakteristische Polynom χ_A einer $n \times n$-Matrix ist ein Polynom vom Grad n. Für $n = 3$ hat es also eine reelle Nullstelle. Eine 3×3-Matrix hat also stets einen Eigenwert.

Satz 7.7 *Eine orthogonale Transformation $A \in O_3$ hat eine spezielle Gerade $g = [v]$ durch 0 und eine Ebene $E \perp g$, und ist entweder*

0) die Identität (mit einem dreidimensionalen Eigenraum zum Eigenwert 1).

1) Spiegelungen an einem 2-dimensionalen Unterverktorraum E. (zwei Eigenwerte 1 und ein Eigenwert -1)

2) Drehungen um eine Achse g (es gibt nur einen Eigenwert $+1$ mit einem eindimensionalen Eigenraum $[v]$).

3) Drehungen um eine Achse g und gleichzeitiger Spiegelung an E (es gibt nur einen Eigenwert -1 mit einem eindimensionalen Eigenraum $[v]$).

4) Punktspiegelung am Ursprung (mit einem dreidimensionalen Eigenraum zum Eigenwert -1).

Beweis : Die Anzahl der Eigenwerte ist stets eins oder drei, da reelles Polynom dritten Grades stets ein oder drei reelle Nullstellen besitzt. Damit sieht man, dass in obiger Liste alle möglichen Kombinationen von Eigenwerten auftreten. Um einen vollständigen Beweis zu geben, müsste man noch zeigen, dass zweidimensionale Eigenräume zum Eigenwert ± 1 unmöglich sind und dass im Falle von eindimensionalen Eigenräumen die Matrix in der Tat eine Drehung ist. $\square$

Stichwortverzeichnis