

# Lineare Algebra Wintersemester 2018/19

Prof. Dr. Annette Huber-Klawitter

Fassung vom 28. März 2019

**Dies ist ein Vorlesungsskript und kein Lehrbuch.  
Mit Fehlern muss gerechnet werden!**

Math. Institut  
Ernst-Zermelo-Str. 1  
79104 Freiburg

0761-203-5560  
annette.huber@math.uni-freiburg.de



# Inhaltsverzeichnis

<b>0</b>	<b>Einleitung</b>	<b>1</b>
<b>1</b>	<b>Lineare Gleichungssysteme und Matrizen</b>	<b>5</b>
<b>2</b>	<b>Körper und Vektorräume</b>	<b>17</b>
<b>3</b>	<b>Basen und Dimension</b>	<b>31</b>
<b>4</b>	<b>Lineare Abbildungen und Dimensionsformel</b>	<b>41</b>
<b>5</b>	<b>Darstellende Matrizen</b>	<b>49</b>
<b>6</b>	<b>Determinanten</b>	<b>57</b>
<b>7</b>	<b>Eigenwerte und Eigenvektoren</b>	<b>67</b>
<b>8</b>	<b>Etwas Gruppentheorie</b>	<b>77</b>
<b>9</b>	<b>Etwas affine und projektive Geometrie</b>	<b>85</b>



# Kapitel 0

## Einleitung

### Inhalt

In der linearen Algebra geht es um das Lösen von linearen Gleichungssystemen:

$$\begin{aligned}2x + 3y &= 4 \\ 3x + y &= 2\end{aligned}$$

Das System heißt linear, da kein  $x^2$  oder gar  $\sin(x)$  vorkommt. Wir lösen die zweite Gleichung nach  $y$  auf und setzen in die erste ein:

$$\begin{aligned}y &= 2 - 3x \\ 4 &= 2x + 3(2 - 3x) = -7x + 6 \Rightarrow x = \frac{2}{7}, y = \frac{8}{7}\end{aligned}$$

Eigentlich ist damit das Wesentliche gesagt: mit diesem Verfahren kann man jedes System von linearen Gleichungen in 2, 3 oder auch 29 Variablen lösen - oder feststellen, dass sie unlösbar sind.

Lineare Gleichungssysteme sind so wichtig, weil sie einerseits überall auftauchen und andererseits auch beherrschbar sind. In der allgemeinen Theorie bemühen wir uns oft, die Frage auf lineare Gleichungssysteme zurückzuführen. Das ganze Konzept der Ableitung beruht auf dieser Idee. Spätestens in der Analysis II wird unsere Theorie benutzt.

Für den Mathematiker stellen sich zwei Arten von Fragen:

- Können wir die Eigenschaften der Gleichungssysteme und ihrer Lösungsräume beschreiben?
- Gibt es effiziente und schnelle Verfahren, die auch real mit 100 oder 1000 Variablen funktionieren? Wie gut sind sie?

Die erste Frage werden wir in der linearen Algebra bearbeiten. Die zweite gehört in das Gebiet der Numerik, die wir nur am Rande betrachten werden.

Eine andere Quelle für die Fragen der linearen Algebra ist die analytische Geometrie: Ebene oder Raum werden mit Koordinaten versehen. Geometrische Objekte wie Geraden werden durch lineare Gleichungen gegeben. Geometrische Fragen (Was ist der Schnitt der folgenden beiden Geraden?) übersetzen sich wieder in lineare Gleichungssysteme. Dies genügt noch nicht, um unsere physikalische Welt zu beschreiben. Wir benötigen zusätzlich einen Abstands- und Winkelbegriff. Die zugehörige Mathematik (Skalarprodukte, bilineare Algebra) werden wir im zweiten Semester studieren.

## Was Sie nebenbei lernen werden

Sie lernen in der lineare Algebra die strenge axiomatische Argumentationsweise kennen. Wir definieren Begriffe und formulieren Sätze über die definierten Objekte. Jeder dieser Sätze muss bewiesen werden. Dabei dürfen nur die Definition und bereits bewiesene Sätze verwendet werden.

Sie sollen in dieser Vorlesung nichts glauben, weil es Ihnen gesagt wird - es muss Ihnen logisch einwandfrei bewiesen werden. Wenn Ihnen ein Argument unvollständig oder unklar erscheint, so ist es ihr Recht und ihre Pflicht nachzufragen. Beim Lösen der Übungsaufgaben üben Sie das Formulieren von Beweisen auch ein. Außerdem ist es die Funktion der Übungsaufgaben, Sie bei der Beschäftigung mit dem Stoff anzuleiten. In den Übungsgruppen lernen Sie das Sprechen über Mathematik.

## Warum muss ich die Vorlesung hören?

*B. Sc. Physic:* Weil Ihr Fach sehr viel und sehr schwierige Mathematik benötigt. Sie müssen daher nicht nur Mathematik lernen, sondern auch die Kommunikation mit Mathematikern. Hier legen wir die Grundlagen.

*Lehramt:* Weil in Deutschland ein Gymnasiallehrer einen akademischen Abschluss in seinem Fach haben soll. Dafür fallen mir mehrere Gründe ein: Damit er/sie kompetent ist, wenn Schüler/innen schwierige Fragen haben. Damit er/sie beraten kann, wenn es um die Wahl eines mathematiklastigen Studienfachs geht. Damit er/sie eine Chance hat in Auseinandersetzungen mit Eltern, die ein solches Fach studiert haben. Wenn Sie im Laufe des Semesters merken, dass es Ihnen keinen Spaß macht, gehen Sie in die Studienberatung, um herauszubekommen, woran es liegt. Wenn es nur Anfangsschwierigkeiten sind: durchhalten. Wenn Sie im falschen Fach sind: schnell wechseln.

*B. Sc. Mathematik:* Sie haben es so gewollt.

## Literatur

Die Verwendung von weiterer Literatur neben der Vorlesungsmitschrift empfiehlt sich sehr! Dieses Skript ist *kein Ersatz* - mit kleinen und größeren Fehlern muss jederzeit gerechnet werden. Es gibt eine Fülle von Lehrbüchern. Die deutschen Texte sind meist sehr ähnlich im Inhalt. Suchen Sie nach einem, dessen Stil Ihnen zusagt. In der Lehrbuchsammlung können Sie sich verschiedene Bücher

ausleihen und testen. Aufpassen sollten Sie nur, dass es wirklich ein Text für Mathematiker und nicht für Ingenieure ist. Kontrolle: Es sollten andere Körper als  $\mathbb{Q}$  und  $\mathbb{R}$  vorkommen, der Dualraum sollte definiert werden. Bücher, auf denen 100 Seiten lang Matrizenoperationen geübt werden, sind nicht geeignet. Zum Material der Vorlesung gibt es auch unzählige Youtube-Videos. Je nach Qualität können diese sehr erhellend sein. Die Vorlesung selbst ist auch ein Taktgeber und stellt klar, was wichtig ist.

- (i) G. Fischer: Lineare Algebra, Vieweg Verlag 1989. (Eine Standardquelle in Deutschland)
- (ii) F. Lorenz: Lineare Algebra I, BI-Wissenschafts-Verlag 1988 (Aufbau entspricht dem Vorgehen der Vorlesung)
- (iii) K. Jänich: Lineare Algebra, 10. Auflage, Springer Verlag 2004 (sehr flüssig geschrieben, Teile des Textes gezielt für Physiker)
- (iv) S. Lang: Algebra, Addison-Wesley 1993. (alternativer Zugang, Stoff geht weit über die lineare Algebra hinaus - ein treuer Begleiter für ein ganzes Mathematikerleben)





# Kapitel 1

## Lineare Gleichungssysteme und Matrizen

Wir nehmen zunächst einen etwas naiven Standpunkt ein. Wir rechnen in rationalen oder reellen Zahlen. Wir beginnen der Einfachheit halber mit dem Fall von zwei Gleichungen in zwei Variablen.

**Definition 1.1.** *Seien  $a, b, c, d, e, f$  Zahlen. Dann heißt*

$$\begin{aligned}ax + by &= e \\ cx + dy &= f\end{aligned}$$

*lineares Gleichungssystem für die Variablen  $x$  und  $y$ .*

Wir machen uns an das Lösen:

- (i) Der Fall  $a = b = c = d = 0$ . Das Gleichungssystem lautet

$$0 = e, 0 = f$$

Ist dies erfüllt, so sind alle Paare  $(x, y)$  Lösungen. Andernfalls ist das System unlösbar.

- (ii) Der Fall: Einer der Koeffizienten  $a, b, c, d$  ist ungleich 0. Durch Vertauschen können wir annehmen, dass  $a \neq 0$ . Dann lösen wir die erste Gleichung nach  $x$  auf:

$$x = \frac{e}{a} - \frac{b}{a}y$$

und setzen in die zweite Gleichung ein:

$$c \left( \frac{e}{a} - \frac{b}{a}y \right) + dy = \frac{ce}{a} + \frac{-cb + ad}{a}y = f$$

bzw.

$$\frac{ad - bc}{a}y = f - \frac{cd}{a}.$$

Wieder sind zwei Fälle zu unterscheiden:

(a)  $ad - bc = 0$ . Die Gleichung lautet also

$$\frac{ce}{a} = f \Leftrightarrow ce = af$$

Ist diese Gleichheit nicht erfüllt, so ist das System unlösbar. Ist sie erfüllt, so ist  $y$  beliebig und  $x$  berechnet sich aus  $y$ . Wir sagen: Wir haben einen "Freiheitsgrad". Oder: Der Lösungsraum ist "eindimensional".

(b)  $ad - bc \neq 0$ . Dann lösen wir weiter:

$$y = \frac{af}{ad - bc} - \frac{ce}{ad - bc}, x = \frac{de}{ad - bc} + \frac{-bf}{ad - bc},$$

Insbesondere gibt es eine Lösung und diese ist eindeutig.

Wir fassen zusammen:

**Satz 1.2.** *Seien  $a, b, c, d, e, f$  Zahlen. Das Gleichungssystem*

$$\begin{aligned} ax + by &= e \\ cx + dy &= f \end{aligned}$$

*ist genau dann eindeutig lösbar, falls  $\delta = ad - bc \neq 0$ . In diesem Fall gilt*

$$x = \frac{d}{\delta}e + \frac{-b}{\delta}f, y = \frac{-c}{\delta}e + \frac{a}{\delta}f$$

*Beweis:* Den Beweis haben wir oben geführt. Zu bemerken ist zusätzlich, dass die Aussagen mit dem Vertauschen von  $a, b, c, d$  verträglich sind. D.h. der Schritt "ohne Einschränkung ist  $a \neq 0$ " ist richtig. Auch im trivialen ersten Fall gilt die Aussage.  $\square$

Nun wenden wir uns dem allgemeinen Fall zu.

**Definition 1.3.** *Seien  $n, m$  natürliche Zahlen. Seien  $a_{ij}$  und  $b_i$  für  $1 \leq i \leq m$  und  $1 \leq j \leq n$  Zahlen. Dann heißt*

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

lineares Gleichungssystem *in den Unbekannten  $x_1, \dots, x_n$ .*

Wir sprechen von  $m$  Gleichungen in  $n$  Unbekannten.

**Definition 1.4.** *Zwei lineare Gleichungssysteme mit  $n$  Unbekannten heißen äquivalent, wenn ihre Lösungsmengen übereinstimmen.*

Lösen eines Gleichungssystems bedeutet die Suche nach besonders einfachen Systemen, die zum gegebenen System äquivalent sind.

**Beispiel.** Multipliziert man eine Gleichung mit einer Zahl  $\alpha \neq 0$ , so ändert dies nichts an der Lösungsmenge. Die Gleichung  $0 = 0$  kann weggelassen werden.

**Definition 1.5.** *Wir betrachten ein lineares Gleichungssystem. Eine elementare Zeilenumformung ist der Übergang zu einem neuen System durch eine der Operationen:*

- (i) Vertauschen zweier Gleichungen.
- (ii) Multiplikation einer Gleichung mit einer Zahl ungleich 0.
- (iii) Addition des Vielfachen einer Gleichung zu einer anderen Gleichung.

**Lemma 1.6.** *Geht ein Gleichungssystem aus einem anderen durch eine elementare Zeilenumformung hervor, so sind sie äquivalent.*

*Beweis:* Eigentlich klar. Zur Übung führen wir die Details für (iii) aus. Wir betrachten ein Gleichungssystem von  $m$  Gleichungen in  $n$  Unbekannten. Die Operation benutzt nur zwei Gleichungen und lässt alle anderen unverändert. Es genügt daher, zur Vereinfachung der Notation, den Fall  $m = 2$  zu betrachten. Wir betrachten also ein Gleichungssystem

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \end{aligned}$$

Wir addieren das  $\alpha$ -fache der ersten Zeile zur zweiten. Das neue Gleichungssystem lautet

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ (a_{21} + \alpha a_{11})x_1 + (a_{22} + \alpha a_{12})x_2 + \cdots + (a_{2n} + \alpha a_{1n})x_n &= b_2 + \alpha b_1 \end{aligned}$$

Ist  $(x_1, \dots, x_n)$  eine Lösung des ersten System, dann erfüllt es automatisch die unveränderte erste Gleichung des neuen Systems. Die Lösung erfüllt auch das System

$$\begin{aligned} \alpha a_{11}x_1 + \alpha a_{12}x_2 + \cdots + \alpha a_{1n}x_n &= \alpha b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \end{aligned}$$

und dann auch die Summe der beiden Gleichungen. Dies ist genau die zweite Gleichung des transformierten Systems.

Ist  $(x_1, \dots, x_n)$  ein Lösungsvektor des zweiten Systems, so ist umgekehrt zu zeigen, dass alle Gleichungen des ersten Systems erfüllt sind. Das erste System geht aus dem zweiten ebenfalls durch eine elementare Zeilentransformation hervor, nämlich die Addition des  $-\alpha$ -fachen der ersten Zeile zur zweiten Zeile. Diese Aussage gilt also nach dem bereits betrachteten Fall.  $\square$

**Bemerkung.** Mit etwas Tricksen führt man (i) und (ii) auf (iii) zurück. Das wird z.B. in der Informatik beim Vertauschen von Speicherinhalten benutzt.

**Theorem 1.7** (Gauß-Algorithmus). *Gegeben sei ein Gleichungssystem von  $m$  Gleichungen in  $n$  Unbekannten.*

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

Dann gibt es eine Zahl  $r \leq n, m$ , so dass das System bis auf Vertauschen der Unbekannten äquivalent ist zu einem System der Form

$$\begin{aligned} x_1 + a'_{12}x_2 + a'_{13}x_3 + \cdots + a'_{1n}x_n &= b'_1 \\ x_2 + a'_{23}x_3 + \cdots + a'_{2n}x_n &= b'_2 \\ &\dots \\ x_r + \cdots + a'_{rn}x_n &= b'_r \\ 0 &= b'_{r+1} \\ &\dots \\ 0 &= b'_m \end{aligned}$$

Das neue System entsteht aus dem alten durch elementare Zeilentransformationen. Das Gleichungssystem ist genau dann lösbar, wenn  $b'_{r+1} = \cdots = b'_m = 0$ . Dann gibt es für jede Wahl von  $x_{r+1}, \dots, x_n$  einen eindeutigen Lösungsvektor  $(x_1, \dots, x_n)$ .

**Bemerkung.** Genauso wichtig wie die Aussage ist das im Beweis verwendete Lösungsverfahren. Es ist von großer praktischer Bedeutung.

*Beweis:* Wir bringen der Reihe nach die erste, dann die zweite, dann die dritte Zeile und Spalte usw. in die gewünschte Form. Wir benutzen die Notation  $a_{ij}$ ,  $b_i$  jeweils auch für die neuen Gleichungssysteme. Dies sollte nicht zu Verwirrung führen.

**1. Schritt:** Durch Vertauschen von Gleichungen erreichen wir, dass die erste Zeile einen Koeffizienten ungleich 0 enthält. (Dies ist unmöglich, wenn alle Koeffizienten  $a_{ij} = 0$  sind. Dann gilt der Satz bereits.) Durch Vertauschen von Variablen erhalten wir ein Gleichungssystem mit  $a_{11} \neq 0$ . Wir dividieren die erste Zeile durch  $a_{11}$ . Damit hat die erste Zeile die gewünschte Form. Wir subtrahieren nun für  $2 \leq i \leq m$  das  $a_{i1}$ -fache der (neuen) ersten Zeile von der  $i$ -ten Zeile. Im neuen Gleichungssystem hat die erste Zeile und die erste Spalte die gewünschte Form.

**2. Schritt:** Durch Vertauschen der Gleichungen 2 bis  $m$  erreichen wir, dass die zweite Zeile einen Eintrag ungleich 0 enthält. (Dies ist unmöglich, wenn alle Koeffizienten  $a_{ij} = 0$  sind für  $i \geq 2$ . Dann gilt der Satz bereits.) Durch Vertauschen der Unbekannten 2 bis  $n$  erhalten wir ein Gleichungssystem mit

$a_{21} = 0$ ,  $a_{22} \neq 0$ . Wir dividieren die zweite Gleichung durch  $a_{22}$ . Damit hat die zweite Zeile die gewünschte Form. Wir subtrahieren nun für  $3 \leq i \leq m$  das  $a_{i2}$ -fache der neuen zweiten Zeile von der  $i$ -ten Zeile. Bei der neuen Matrix haben die ersten beiden Zeilen und Spalten die gewünschte Form

**3. bis  $m$ -ter Schritt:** Genauso

Die Aussage über die Lösbarkeit ist klar. Die übrigen  $r$  Gleichungen lassen sich jeweils auflösen nach der ersten vorkommenden Variable  $x_i$ . Durch Einsetzen von unten nach oben erhält man für jedes  $x_{r+1}, \dots, x_n$  eine eindeutige Lösung  $(x_1, \dots, x_n)$ .  $\square$

**Bemerkung.** Ist  $r = m$ , so ist das System automatisch lösbar. Ist das System eindeutig lösbar, so muss  $r = n$  sein, insbesondere also auch  $m \geq n$ .

**Frage.** Ist die Zahl  $r$  aus dem Theorem eindeutig?

Diese Frage wird uns nun für einige Zeit beschäftigen. Zu ihrer Beantwortung werden wir uns auf die *Struktur* der Lösungsmenge konzentrieren und einiges an abstrakter Begrifflichkeit einführen.

## Matrizen

Wir führen gleich eine abkürzende Schreibweise ein, die uns in der gesamten linearen Algebra begleiten wird.

**Definition 1.8.** Seien  $n, m$  natürliche Zahlen. Seien  $a_{ij}$  für  $1 \leq i \leq m$  und  $1 \leq j \leq n$  Zahlen. Dann heißt

$$A = (a_{ij})_{i=1, j=1}^{m, n} = (a_{ij})_{m, n} = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

$m \times n$ -Matrix mit den Einträgen  $a_{ij}$ .

$n \times n$ -Matrizen heißen quadratisch. Eine  $m \times 1$ -Matrix

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_m \end{pmatrix}$$

heißt auch Spaltenvektor der Länge  $m$ . Eine  $1 \times n$ -Matrix

$$w = (w_1 w_2 \dots w_n) = (w_1, w_2, \dots, w_n)$$

heißt auch Zeilenvektor der Länge  $n$ .

Matrizen kann man leicht addieren und mit Zahlen multiplizieren.

**Definition 1.9.** Für  $m, n \in \mathbb{N}$  seien  $A = (a_{ij}), B = (b_{ij})$   $m \times n$ -Matrizen, sowie  $\alpha$  eine Zahl. Dann heißt

$$A + B = (c_{ij}) \text{ mit } c_{ij} = a_{ij} + b_{ij} \text{ für alle } i, j$$

die Summe der Matrizen. Weiter heißt

$$\alpha A = (\alpha a_{ij})_{i=1, j=1}^{m, n}$$

skalares Vielfaches von  $A$ . Die Nullmatrix ist die  $n \times m$ -Matrix mit allen Einträgen gleich 0. Die Einheitsmatrix ist die  $n \times n$ -Matrix

$$E_n = (\delta_{ij})_{n, n} \text{ wobei } \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

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

**Bemerkung.** Eine Variable  $x, a, \dots$  kann also sowohl für eine Zahl, einen Zeilen- oder Spaltenvektor oder sogar eine Matrix stehen. Dies muss jeweils im Kontext klar sein bzw. definiert werden. Aus Gründen der Schreibökonomie

schreiben wir auch oft  $(b_1, \dots, b_n)$  für den Spaltenvektor  $\begin{pmatrix} b_1 \\ \dots \\ b_n \end{pmatrix}$ . Das Symbol 0

kann eine Zahl, ein Nullvektor oder eine Nullmatrix sein - auch hier unterscheiden wir nicht in der Notation.

Matrizen lassen sich multiplizieren:

**Definition 1.10.** Seien  $m, n, k \in \mathbb{N}$ ,  $A = (a_{ij})_{m, p}$  eine  $m \times p$ -Matrix und  $B = (b_{jk})_{p, n}$  eine  $p \times n$ -Matrix. Dann ist  $AB = (c_{ik})_{m, n}$  die  $m \times n$ -Matrix mit den Einträgen

$$c_{ik} = \sum_{j=1}^p a_{ij} b_{jk} = a_{i1} b_{1k} + a_{i2} b_{2k} + a_{i3} b_{3k} + \dots + a_{ip} b_{pk}$$

Merkregel:  $i$ -te Zeile mal  $k$ -te Spalte für den Eintrag  $ik$ .

**Beispiel.** (i)

$$\begin{pmatrix} 1 & 2 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} 5 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 5+2 & 1+4 \\ 4 & 8 \end{pmatrix} = \begin{pmatrix} 7 & 5 \\ 4 & 8 \end{pmatrix}$$

(ii)

$$(1, 0, 1, 9) \begin{pmatrix} 2 \\ \frac{1}{2} \\ 1 \\ 4 \end{pmatrix} = (2 + 0 + 1 + 36) = (39)$$

(iii)

$$\begin{pmatrix} 2 \\ \frac{1}{2} \\ 1 \\ 4 \end{pmatrix} (1, 0, 1, 9) = \begin{pmatrix} 2 & 0 & 2 & 18 \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{9}{2} \\ 1 & 0 & 1 & 9 \\ 4 & 0 & 4 & 36 \end{pmatrix}$$

(iv)

$$\begin{pmatrix} 1 & 0 & 2 \\ 7 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \text{ ist nicht definiert}$$

Zunächst zurück zu unserem Thema.

Das lineare Gleichungssystem aus Definition 1.8 schreiben wir jetzt:

$$Ax = b$$

mit  $A = (a_{ij})_{m,n}$ ,  $b = \begin{pmatrix} b_1 \\ \dots \\ b_n \end{pmatrix}$  und dem Lösungsvektor  $x = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}$ . Die Matrix

$$(A|b) = \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ a_{21} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix}$$

heißt *erweiterte Koeffizientenmatrix* des Gleichungssystems.

**Bemerkung.** In Satz 1.2 ging es um die Matrix  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  und die Gleichung

$$A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} e \\ f \end{pmatrix}$$

Die erweiterte Koeffizientenmatrix ist

$$\begin{pmatrix} a & b & e \\ c & d & f \end{pmatrix}$$

Die Zahl  $\delta = ad - bc$  heißt *Determinante*  $\det(A)$  von  $A$ . Auch die Lösungsformel schreibt sich am einfachsten mit Matrizen. Ist  $\det(A) \neq 0$ , so gilt

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{d}{\det(A)} & \frac{-b}{\det(A)} \\ \frac{-c}{\det(A)} & \frac{a}{\det(A)} \end{pmatrix} \begin{pmatrix} e \\ f \end{pmatrix}$$

Wie für Zahlen vereinbaren wir Punkt vor Strichrechnung.

**Satz 1.11** (Rechenregeln für Matrizen). *Seien  $A, B, C, D, F, G$  Matrizen, so dass jeweils die Rechenoperationen definiert sind. Seien  $\alpha, \beta$  Zahlen.*

(i) (*Assoziativität der Addition*)

$$A + (B + C) = (A + B) + C$$

(ii) (Kommutativität der Addition)

$$A + B = B + A$$

(iii) (Assoziativität der Multiplikation)

$$A(BC) = (AB)C$$

(iv) (neutrales Element der Multiplikation)

$$E_m A = A, A E_n = A$$

(v) (Distributivitätsgesetze der Matrixmultiplikation)

$$\begin{aligned} A(B + C) &= AB + AC \\ (D + F)G &= DG + FG \end{aligned}$$

(vi) (Assoziativgesetz der skalaren Multiplikation)

$$\begin{aligned} (\alpha\beta)A &= \alpha(\beta A) \\ \alpha(AB) &= (\alpha A)B = A(\alpha B) \end{aligned}$$

(vii) (Distributivgesetz der skalaren Multiplikation)

$$\begin{aligned} (\alpha + \beta)A &= \alpha A + \beta A \\ \alpha(A + B) &= \alpha A + \alpha B \end{aligned}$$

**Bemerkung.** Die Multiplikation ist nicht kommutativ! Meist ist nur eines der beiden Produkte überhaupt definiert.

*Beweis:* Hinschreiben, nachrechnen mit den Rechenregeln für Zahlen. Wir führen den Beweis für die Regel (iii) aus. Seien

$$A = (a_{ij})_{m,n}, B = (b_{jk})_{n,p}, C = (c_{kl})_{p,q}$$

Nach Definition gilt

$$AB = \left( \sum_j a_{ij} b_{jk} \right)_{i=1, k=1}^{m,p}, \quad (AB)C = \left( \sum_k \left( \sum_j a_{ij} b_{jk} \right) c_{kl} \right)_{i=1, l=1}^{m,q}$$

Und andererseits gilt

$$BC = \left( \sum_k b_{jk} c_{kl} \right)_{j=1, l=1}^{n,q}, \quad A(BC) = \left( \sum_j a_{ij} \left( \sum_k b_{jk} c_{kl} \right) \right)_{i=1, l=1}^{m,q}$$

Die Behauptung gilt, da für jedes  $i, l$  gilt

$$\sum_k \left( \sum_j a_{ij} b_{jk} \right) c_{kl} = \sum_{k,j} a_{ij} b_{jk} c_{kl} = \sum_j a_{ij} \left( \sum_k b_{jk} c_{kl} \right)$$

□



**Bemerkung.** In der Physik kommen viele solche Rechnungen vor. Dort gibt es eine abkürzende Notation: In einem Ausdruck wie  $a_{ij}b_{jk}$  oder  $a_{ij}b_{jk}c_{kl}$  wird über alle Indizes summiert, die doppelt vorkommen. Damit vereinfacht sich unsere Rechnung zu

$$a_{ij}(b_{jk}c_{kl}) = (a_{ij}b_{jk})c_{kl}$$

$a_{ii}$  steht für  $\text{Spur}(a_{ij}) = \sum_i a_{ii}$ , die Spur der Matrix. In einer Verfeinerung unterscheidet man noch zwischen Zeilen und Spaltenindizes.  $a_i^j$  ist die Matrix  $(a_i^j)_{i=1, j=1}^{m, n}$ . Damit ist klar, dass  $b_i$  ein Spaltenvektor ist,  $b^i$  jedoch ein Zeilenvektor. Matrixmultiplikation ist  $a_i^j b_j^k$ , d.h. es wird summiert über doppelte Indizes, aber das ist nur erlaubt, falls ein Index oben, der andere unten steht ("Kontraktion"). Diese Notation gibt es auch mit mehr als zwei Indizes, dann spricht man von Tensoren. Mathematisch steht dahinter das Tensorprodukt, das wir in der LA I wohl nicht mehr behandeln werden.

Nun wenden wir dies auf unsere Gleichungssysteme an.

**Beispiel.** Sei  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ,  $A' = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ . Dann gilt

$$A'A = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} da - bc & bd - bd \\ -ca + ac & -cb + da \end{pmatrix} = \det(A)E_2$$

Für  $B = \begin{pmatrix} e \\ f \end{pmatrix}$  rechnen wir also

$$Ax = B \Rightarrow \det(A)x = \det(A)E_2x = A'Ax = A'B \Rightarrow x = \det(A)^{-1}A'B$$

Genau das hatten wir vorher per Hand hergeleitet.

**Definition 1.12.** Sei  $Ax = b$  ein lineares Gleichungssystem. Es heißt homogen, falls  $b = 0$ , andernfalls inhomogen.  $Ax = 0$  heißt zu  $Ax = b$  korrespondierendes homogenes Gleichungssystem.

**Lemma 1.13.** Sei  $Ax = 0$  ein homogenes Gleichungssystem.

(i) Sind  $x, x'$  Lösungsvektoren des Systems, so auch  $x + x'$ .

(ii) Ist  $\alpha$  eine Zahl,  $x$  ein Lösungsvektor, so auch  $\alpha x$ .

*Beweis:* Dieses Lemma ist so wichtig, dass wir den Beweis genau durchgehen. Nach Voraussetzung gilt  $Ax = Ax' = 0$ . Dann folgt

$$0 = Ax + Ax' \stackrel{(v)}{=} A(x + x') \quad 0 = \alpha 0 = \alpha(Ax) \stackrel{(vi)}{=} A(\alpha x)$$

□

Wir sagen: Die Lösungen eines homogenen Systems bilden einen Vektorraum (nächstes Kapitel).

**Lemma 1.14.** *Sei  $Ax = b$  ein lineares Gleichungssystem. Hat es eine Lösung  $x_0$ , so hat jede Lösung eindeutig die Form  $x + x_0$ , wobei  $x$  eine Lösung des korrespondierenden homogenen Systems ist.*

*Beweis:* Wir setzen  $Ax_0 = b$  voraus. Sei  $x$  eine Lösung des homogenen Systems, also  $Ax = 0$ . Dann folgt

$$A(x + x_0) = Ax + Ax_0 = 0 + b = b$$

Sei umgekehrt  $y$  eine andere Lösung des inhomogenen Systems. Wir schreiben  $y = x + x_0$ . Dann folgt

$$Ax = A(y - x_0) = Ay - Ax_0 = b - b = 0$$

□

Wir sagen: Die Lösungen des inhomogenen Systems bilden einen affinen Raum. Alle Operationen mit Gleichungssystemen sind natürlich auch Operationen mit Matrizen.

**Definition 1.15.** *Sei  $A$  eine  $m \times n$ -Matrix. Eine elementare Zeilenumformung ist der Übergang zu einer  $m \times n$ -Matrix durch eine der Operationen:*

- (i) Vertauschen zweier Zeilen.
- (ii) Multiplikation einer Zeile mit einer Zahl ungleich 0.
- (iii) Addition des Vielfachen einer Zeile zu einer anderen Zeile.

Analog definieren wir elementare Spaltenumformungen.

**Satz 1.16** (Matrizennormalform). *Sei  $A$  eine  $m \times n$ -Matrix. Dann kann  $A$  durch eine Kette von elementaren Zeilenumformungen und Vertauschen von Spalten überführt werden in eine Matrix  $A'$  der Form*

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

wobei  $B$  eine  $r \times r$ -Matrix ist der Gestalt

$$B = \begin{pmatrix} 1 & b_{12} & \dots & b_{1r} \\ 0 & 1 & \dots & b_{2r} \\ & & \dots & \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

und  $C$  eine beliebige  $r \times (n - r)$ -Matrix.

Durch eine Kette von elementaren Zeilen- und Spaltenumformungen kann  $A$  überführt werden in eine Matrix  $A''$  der Form

$$A'' = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

*Beweis:* Die erste Aussage erhalten wir durch die Anwendung des Gauß-Algorithmus auf  $Ax = 0$ .

Danach subtrahieren wir von der ersten Zeile das  $b_{21}$ -fache der zweiten Zeile, dann das  $b_{31}$ -fache der dritten Zeile usw. Wir subtrahieren das  $b_{32}$ -fache der dritten Zeile von der zweiten, dann das  $b_{42}$ -fache der vierten Zeile von der zweiten usw. Mit diesem Verfahren erhalten wir eine Matrix der Form

$$\begin{pmatrix} E_r & C' \\ 0 & 0 \end{pmatrix}$$

Nun wenden wir elementare Spaltentransformationen an. Durch Subtraktion von Vielfachen der ersten  $r$  Spalten werden alle Einträge von  $C'$  gelöscht.  $\square$

**Frage.** Wie oben: Ist  $r$  eindeutig?



## Kapitel 2

# Körper und Vektorräume

### Grundbegriffe

Die gesamte moderne Mathematik wird (meist) auf dem Begriff der Menge aufgebaut. *Mengenlehre* ist ein eigenes Teilgebiet, in dem es um die genaue Formulierung der Axiome geht. Hier ist nicht der richtige Ort, um die Probleme und Lösungen zu diskutieren. Wir führen nur die allernötigsten Dinge ein. Mengen haben Elemente. Wir schreiben  $x \in M$ , wenn  $x$  ein Element der Menge  $M$  ist. Zwei Menge sind gleich, wenn sie diesselben Elemente haben:

$$M = N \Leftrightarrow (x \in M \Leftrightarrow x \in N).$$

Daher können Mengen durch die Angabe ihrer Elemente beschrieben werden.

#### Beispiel.

$$M = \{1, 2, 7, 100\}, N = \{0, -1, 1, 2\}, X = \{0, -1, 0, 1, 2\}$$

Es ist  $M \neq N$ , denn  $0 \in N$ , aber  $0 \notin M$ . Es ist  $N = X$ , denn jedes Element von  $N$  liegt in  $X$  und umgekehrt.

Wir sagen  $N \subset M$  ( $N$  ist *Teilmenge* von  $M$ ), wenn jedes Element von  $N$  auch ein Element von  $M$  ist, also:

$$x \in N \Rightarrow x \in M.$$

In dieser Definition ist Gleichheit erlaubt. (Achtung, manche Autoren schreiben statt dessen  $N \subseteq M$  und meinen mit  $N \subset M$  nur echte Inklusion, bei denen es also ein  $x \in M$  gibt, das nicht in  $N$  liegt.)

Aus Mengen können neue Mengen gebildet werden. Seien  $M, N$  Mengen. Die *Vereinigungsmenge*  $M \cup N$  besteht aus allen Elementen, die in  $M$  oder in  $N$  liegen:

$$x \in M \cup N \Leftrightarrow (x \in M \text{ oder } x \in N).$$

Die *Schnittmenge*  $M \cap N$  besteht aus allen Elementen, die in  $M$  und in  $N$  liegen:

$$x \in M \cap N \Leftrightarrow (x \in M \text{ und } x \in N).$$

Auch unendliche Vereinigungen und Schnitte können gebildet werden. Sei  $I$  eine Menge (*Indexmenge*). Für jedes  $i \in I$  sei  $M_i$  eine Menge. Dann gibt es die Vereinigungsmenge  $M = \bigcup_{i \in I} M_i$

$$x \in \bigcup_{i \in I} M_i \Leftrightarrow (\text{Es gibt ein } i \in I \text{ so dass } x \in M_i)$$

und die Schnittmenge  $N = \bigcap_{i \in I} M_i$

$$x \in \bigcap_{i \in I} M_i \Leftrightarrow (x \in M_i \text{ für alle } i \in I).$$

**Beispiel.** Seien  $M, N$  wie oben. Dann ist

$$M \cup N = \{0, -1, 1, 2, 7, 100, 0, 2\}, \quad M \cap N = \{1, 2\}.$$

Weiter gibt es die *Differenzmenge*  $M \setminus N$  (lies:  $M$  ohne  $N$ ), die alle Elemente von  $M$  enthält, die nicht in  $N$  liegen:

$$x \in M \setminus N \Leftrightarrow (x \in M \text{ und } x \notin N).$$

Besonders wichtig ist auch das *kartesische Produkt*  $M \times N$ . Seine Elemente sind Paare von Elementen  $(x, y)$  mit  $x \in M, y \in N$ . Es gilt  $(x, y) = (x', y')$  genau dann, wenn  $x = x'$  und  $y = y'$ . Sind allgemeiner  $M_1, M_2, \dots, M_n$  Mengen, dann ist  $M_1 \times \dots \times M_n$  die Menge der  $n$ -Tupel  $(x_1, \dots, x_n)$  mit  $x_i \in M_i$  für  $i = 1, \dots, n$ . Speziell für  $M_1 = M_2 = \dots = M_n$  schreiben wir

$$M \times \dots \times M = M^n.$$

Ist  $M$  eine Menge und  $P$  eine Eigenschaft, so gibt es die Menge

$$M(P) = \{x \in M \mid x \text{ erfüllt } P\}.$$

**Beispiel.** (i) Sei  $\mathbb{N}$  die Menge der natürlichen Zahlen. Dann gibt es die Menge

$$N = \{x \in \mathbb{N} \mid 2 \text{ teilt } x\} = \{2, 4, 6, 8, \dots\}$$

der geraden natürlichen Zahlen.

(ii) Sei  $M$  eine Menge. Wir betrachten die Diagonale

$$\Delta = \{(x, y) \in M^2 \mid x = y\}$$

Seien  $M, N$  Mengen. Eine *Abbildung*  $f : M \rightarrow N$  ist eine Vorschrift, die jedem  $x \in M$  ein  $f(x) \in N$  zuordnet. Wir schreiben auch  $x \mapsto f(x)$ . Zu einer Abbildung gehört ihr *Graph*

$$\Gamma_f = \{(x, y) \in M \times N \mid y = f(x)\}.$$

**Beispiel.** Die Diagonale  $\Delta$  ist der Graph der *Identität*  $x \mapsto x$ .

Alternativer (besserer) Standpunkt: Eine *Relation* ist eine Teilmenge  $R$  von  $M \times N$ . Eine Relation ist eine *Abbildung*, wenn es für jedes  $x \in M$  genau ein  $y \in N$  gibt mit  $(x, y) \in R$ . Wir sagen,  $f(x)$  ist das Bild von  $x$  und  $x$  ist ein Urbild von  $f(x)$ .

Ist  $f : M \rightarrow N$  eine Abbildung, so heißt  $M$  Definitionsbereich und  $N$  Wertebereich von  $f$ . Das *Bild*  $f(M)$  ist

$$f(M) = \{y \in N \mid \text{es gibt } x \in M \text{ mit } y = f(x)\} = \{f(x) \mid x \in M\}.$$

Für  $T \subset N$  heißt

$$f^{-1}(T) = \{x \in M \mid f(x) \in T\}$$

*Urbildmenge* von  $T$ .

- Eine Abbildung heißt *surjektiv*, wenn  $f(M) = N$ , d.h. für jedes  $y \in N$  gibt es  $x \in M$  mit  $y = f(x)$ .
- Eine Abbildung heißt *injektiv*, wenn jedes Element von  $N$  höchstens ein Urbild hat, also:

$$f(x) = f(x') \Rightarrow x = x'.$$

- Eine Abbildung heißt *bijektiv*, wenn sie injektiv und surjektiv ist. In diesem Fall definieren wir die *Umkehrabbildung*  $f^{-1} : N \rightarrow M$ , indem wir  $y \in N$  das eindeutige  $x \in M$  zuordnen mit  $f(x) = y$ . Es existiert wegen der Surjektivität und ist eindeutig wegen der Injektivität. In Relationensprache:

$$\Gamma_{f^{-1}} = \{(y, x) \in N \times M \mid (x, y) \in \Gamma_f\}.$$

Sind  $f : M \rightarrow N$  und  $g : N \rightarrow K$  Abbildungen, so definieren wir die *Verkettung*  $g \circ f : M \rightarrow K$  als

$$g \circ f(x) = g(f(x)).$$

**Beispiel.** Nach Definition ist  $f^{-1} \circ f = \text{id}_M$ . Man überprüft leicht  $f \circ f^{-1} = \text{id}_N$ .

Das Verketteten von Abbildungen ist *assoziativ*: für  $f : M \rightarrow N$ ,  $g : N \rightarrow K$ ,  $h : K \rightarrow L$  gilt

$$f \circ (g \circ h) = (f \circ g) \circ h$$

denn für jedes  $x \in M$  gilt

$$(f \circ (g \circ h))(x) = f((g \circ h)(x)) = f(g(h(x)))$$

und

$$((f \circ g) \circ h)(x) = (f \circ g)(h(x)) = f(g(h(x))).$$

Zum Abschluss noch etwas allgemeine mathematische Notation:

$A \Rightarrow B$  Lies: aus  $A$  folgt  $B$ . D.h.: Wenn die Aussage  $A$  gilt, dann auch die Aussage  $B$ . Wenn  $A$  nicht gilt, kann  $B$  gelten oder nicht. Wenn  $B$  nicht gilt, dann gilt auch  $A$  nicht.

In älterer Sprache:  $A$  ist *hinreichende Bedingung* für  $B$ .  $B$  ist *notwendige Bedingung* für  $A$ . Die meisten Denkfehler beruhen auf der Verwechslung von notwendig und hinreichend.

$A \Leftrightarrow B$  Lies:  $A$  ist äquivalent zu  $B$ . D.h.: Die Aussage  $A$  gilt genau dann, wenn  $B$  gilt. Ist eines wahr, dann das andere. Ist eines falsch, dann das andere.  $A$  ist notwendige und hinreichende Bedingung für  $B$ .

$A \wedge B$  Lies:  $A$  und  $B$ . D.h.: Die Aussage ist wahr, wenn sowohl  $A$  als auch  $B$  gelten.

$A \vee B$  Lies:  $A$  oder  $B$ . D.h.: Die Aussage gilt, wenn  $A$  gilt, wenn  $B$  gilt oder auch beide.

$\neg A$  Lies: nicht  $A$  D.h.  $\neg A$  ist genau dann wahr, wenn  $A$  falsch ist.

$\forall, \exists$  Lies: für alle; es existiert. Diese Symbole nennt man *Quantoren*.

*Schlecht*:  $Ax = b$  für  $x$ .

*Richtig*: Eines von:

- Es gibt  $x \in M$  mit  $Ax = b$ .
- Für alle  $x \in M$  gilt  $Ax = b$ . Oder: Es gilt  $Ax = b$  für alle  $x \in M$ .
- Sei  $x \in M$  mit  $Ax = b$ .

Es handelt sich um verschiedene Aussagen! In der schlechten Fassung ist nicht klar, welche gemeint ist. Manchmal wird es aus dem Kontext klar, besser ist präzise Sprache.

## Erste algebraische Strukturen

Beim Lösen unserer linearen Gleichungssysteme kam es nicht auf den konkreten Zahlbereich an, sondern nur auf die erlaubten Rechenarten und die zugehörigen Rechenregeln, nämlich  $+$ ,  $-$ ,  $\cdot$  und auch  $\div$ . Dies formalisieren wir. Wir wollen dabei gleich ein wenig ausholen.

**Definition 2.1.** Eine Gruppe ist ein Paar  $(G, \cdot)$  bestehend aus einer Menge  $G$  und einer Abbildung

$$\cdot : G \times G \rightarrow G$$

so dass die folgenden Axiome erfüllt sind:

(i) (Assoziativität) Für alle  $a, b, c \in G$  gilt

$$a(bc) = (ab)c$$



(ii) (neutrales Element) Es gibt ein Element  $e \in G$ , so dass für alle  $a \in G$  gilt

$$ea = ae = a$$

(iii) (inverses Element) Für jedes  $a \in G$  gibt es ein Element  $a' \in G$  mit

$$aa' = a'a = e$$

$e$  heißt neutrales Element,  $a'$  heißt inverses Element zu  $a$ . Wir schreiben auch  $a^{-1}$ .

Eine Gruppe heißt kommutativ oder abelsch, wenn zusätzlich für alle  $a, b \in G$  gilt

(iv)

$$ab = ba$$

Meist schreiben wir kurz  $G$  statt  $(G, \cdot)$ .

Im Fall von abelschen Gruppen schreiben wir oft auch  $+$  statt  $\cdot$ , das neutrale Element heißt dann  $0$ , das Inverse von  $a$  heißt  $-a$ .

**Beispiel.** (i)  $(\mathbb{Z}, +)$  ist eine abelsche Gruppe, ebenso  $(\mathbb{Q}, +)$ ,  $(\mathbb{R}, +)$ .

(ii)  $(\mathbb{N}, +)$  ist keine Gruppe (kein neutrales Element),  $(\mathbb{N}_0, +)$  auch nicht (neutrales Element  $0$ , aber keine Inversen)

(iii)  $(\mathbb{Z}, \cdot)$  ist keine Gruppe (neutrales Element  $1$ , aber keine Inversen).

(iv)  $(\mathbb{Q}, \cdot)$  ist keine Gruppe (neutrales Element  $1$ , aber  $0$  hat kein Inverses.  $\mathbb{Q}^* = (\mathbb{Q} \setminus \{0\}, \cdot)$  ist eine abelsche Gruppe.

(v) Sei  $M$  eine Menge,  $S(M) = \{f : M \rightarrow M \mid f \text{ bijektiv}\}$  ist eine Gruppe, nicht kommutativ, falls  $M$  mehr als zwei Elemente hat. Die Identität  $x \mapsto x$  ist das neutrale Element, die Umkehrabbildung  $f^{-1}$  von  $f$  das Inverse von  $f$ .

(vi)  $G =$  Menge der Kongruenzabbildungen der Ebene ist eine Gruppe.

Die wirklich interessanten Gruppen sind Symmetriegruppen von irgendwelchen Objekten, so wie die letzten beiden Beispiele. Je mehr Symmetrien, desto größer die Gruppe. Dies ist ein mächtiges Hilfsmittel bei Studium physikalischer Systeme. So steckt hinter der Energieerhaltung eine Translationsinvarianz eines mechanischen Systems bezüglich der Zeit. Wir werden später darauf zurückkommen.

**Lemma 2.2.** Sei  $G$  eine Gruppe. Dann ist das neutrale Element  $e$  eindeutig. Ist  $a \in G$ , so ist das inverse Element  $a^{-1}$  eindeutig.

*Beweis:* Seien  $e, e'$  neutrale Elemente. Dann gilt

$$e = ee' = e'$$

nach Eigenschaft (ii) für  $e'$  und für  $e$ . Seien  $a', a''$  inverse Elemente von  $a$ . Dann gilt

$$a'' = ea'' = (a'a)a'' = a'(aa'') = a'e = a'$$

□

**Definition 2.3.** Ein Ring ist ein Tripel  $(R, +, \cdot)$  bestehend aus einer Menge  $R$  und Abbildungen

$$+ : R \times R \rightarrow R, \quad \cdot : R \times R \rightarrow R$$

so dass die folgenden Axiome erfüllt sind:

(i)  $(R, +)$  ist eine abelsche Gruppe. Das neutrale Element nennen wir  $0$ .

(ii) (Assoziativität der Multiplikation) Für alle  $a, b, c \in R$  gilt

$$a(bc) = (ab)c$$

(iii) (Distributivgesetze) Für alle  $a, b, c, d, e, f \in R$  gilt

$$a(b + c) = ab + ac, (d + e)f = df + ef$$

Ein Ring  $R$  heißt kommutativ, falls zusätzlich gilt

(iv) (Kommutativität der Multiplikation) Für alle  $a, b \in R$  gilt

$$ab = ba$$

Ein Ring heißt unitär, wenn es ein  $1 \in R$  gibt, so dass für alle  $a \in R$

$$1a = a1 = a$$

Ein Körper ist ein unitärer kommutativer Ring, so dass  $(R \setminus \{0\}, \cdot)$  eine abelsche Gruppe ist.

Ein nicht notwendig kommutativer Ring, so dass  $(R \setminus \{0\}, \cdot)$  eine Gruppe ist, heißt *Schiefkörper* oder *Divisionsalgebra*. Wir werden uns in dieser Vorlesung auf Körper konzentrieren.

**Beispiel.**

$\mathbb{Z}$  ist ein kommutativer, unitärer Ring.

$\mathbb{Q}, \mathbb{R}$  sind sogar Körper.

**Bemerkung.** Alle Sätze aus Kapitel 1 gelten für Gleichungssysteme und Matrizen über Körpern. Die Rechenregeln für Matrizen gelten (mindestens) für Matrizen mit Einträgen in einem kommutativen, unitären Ring. Der Gauß-Algorithmus benutzt jedoch das Invertieren von Elementen. Er gilt nicht für Ringe (z.B. nicht über  $\mathbb{Z}$ ).

**Definition 2.4.** Sei  $R$  ein kommutativer unitärer Ring,  $n, m \in \mathbb{N}$ . Dann schreiben wir

$M_{m \times n}(R) =$  Menge der  $m \times n$ -Matrizen mit Einträgen in  $R$

$M_n(R) =$  Menge der quadratischen  $n \times n$ -Matrizen mit Einträgen in  $R$

$R^n =$  Menge der Spaltenvektoren mit Einträgen in  $R$

**Beispiel.** Sei  $R$  ein kommutativer unitärer Ring,  $n \in \mathbb{N}$ . Dann ist  $M_n(R)$  nach Satz 1.11 ein unitärer Ring mit Einselement  $E_n$ . Für  $n \geq 2$  ist es weder ein kommutativer Ring noch eine Divisionsalgebra.

Aus den Axiomen folgen weitere Rechenregeln.

**Lemma 2.5.** Sei  $R$  ein Ring,  $a, b \in R$ . Dann gilt

$$0 \cdot a = a \cdot 0 = 0, (-b)a = b(-a) = -(ba), (-1)a = -a$$

*Beweis:* Wir rechnen geschickt:

$$0 \cdot a + 0 \cdot a = (0 + 0) \cdot a = 0 \cdot a$$

Addition von  $-(0 \cdot a)$  ergibt

$$0 \cdot a = 0$$

Ebenso für  $a \cdot 0$ . Die zweite Formel folgt ähnlich:

$$(-b)a + ba = (-b + b)a = 0a = 0 \Rightarrow (-b)a = -(ba)$$

Ebenso für  $b(-a)$ . Schließlich:

$$0 = 0a = (1 - 1)a = 1a + (-1)a = a + (-1)a \Rightarrow -a = (-1)a.$$

□

Wir wollen über beliebigen Körpern rechnen, daher sollten Sie andere Beispiele als  $\mathbb{Q}$  und  $\mathbb{R}$  kennen.

**Definition 2.6.** Die komplexen Zahlen  $\mathbb{C}$  bestehen aus der Menge

$$\mathbb{C} = \mathbb{R}^2 = \{(x, y) | x, y \in \mathbb{R}\}$$

mit der komponentenweisen Addition

$$(x, y) + (x', y') = (x + x', y + y') \text{ für alle } x, x', y, y' \in \mathbb{R}$$

und der Multiplikation

$$(x, y)(x', y') = (xx' - yy', xy' + x'y) \text{ für alle } x, x', y, y' \in \mathbb{R}$$

Die Abbildung  $x \mapsto (x, 0)$  ist injektiv und verträglich mit Addition und Multiplikation. Wir identifizieren  $\mathbb{R}$  mit dem Bild unter dieser Abbildung. Das Element  $i = (0, 1)$  heißt *imaginäre Einheit*. Die komplexe Zahl  $(x, y)$  lässt sich nun eindeutig schreiben als

$$(x, y) = x + iy$$

Als einzige Rechenregel muss man sich merken

$$i^2 = (0, 1)(0, 1) = (0 - 1, 0 + 0) = -1$$

**Lemma 2.7.**  $\mathbb{C}$  ist ein Körper.

*Beweis:*  $(\mathbb{C}, +)$  ist die abelsche Gruppe  $\mathbb{R}^2$ . Zu verifizieren sind die Rechenregeln der Multiplikation. Diese kann man einfach nachrechnen. (Eigentlich haben wir Assoziativgesetz und Distributivgesetze zur Definition der Multiplikation schon benutzt.) Nur die Existenz von multiplikativen Inversen ist nicht sofort klar.

Sei  $\alpha = x + iy$  eine komplexe Zahl. Dann setzen wir  $\bar{\alpha} = x - iy$ . Es folgt

$$|\alpha|^2 := \alpha\bar{\alpha} = (x + iy)(x - iy) = x^2 - i^2y^2 = x^2 + y^2 \in \mathbb{R}$$

Es folgt also

$$\alpha = 0 \Leftrightarrow \alpha\bar{\alpha} = 0 \Leftrightarrow |\alpha| = 0$$

Für  $\alpha \neq 0$  gilt dann

$$\alpha \frac{\bar{\alpha}}{|\alpha|^2} = \alpha \frac{\bar{\alpha}}{\alpha\bar{\alpha}} = 1$$

□

Es gibt auch Körper mit nur endlichen vielen Elementen.

**Definition 2.8.** Sei  $\mathbb{F}_2 = \{0, 1\}$  mit Addition und Multiplikation

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

Sei  $\mathbb{F}_3 = \{0, 1, 2\}$  mit der Addition und Multiplikation

$$\begin{array}{c|ccc} + & 0 & 1 & 2 \\ \hline 0 & 0 & 1 & 2 \\ 1 & 1 & 2 & 0 \\ 2 & 2 & 0 & 1 \end{array} \qquad \begin{array}{c|ccc} \cdot & 0 & 1 & 2 \\ \hline 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 2 \\ 2 & 0 & 2 & 1 \end{array}$$

Hinter diesen Operationen steckt die Addition und Multiplikation von *Restklassen*. Wir identifizieren ganze Zahlen, die denselben Rest bei der Division durch 2 bzw. 3 haben. Also in  $\mathbb{F}_3$ :

$$2 + 2 = 4 = 1$$

denn 4 hat bei Division durch 3 den Rest 1.

**Lemma 2.9.**  $\mathbb{F}_2$  und  $\mathbb{F}_3$  sind Körper.

*Beweis:* Aus den Tabellen liest man ab, dass 0 jeweils ein neutrales Element der Addition ist, 1 ein neutrales Element der Multiplikation. Die Operationen sind kommutativ, da die Tabellen symmetrisch bezüglich der Diagonale sind. Die Existenz von Inversen liest man ab, da bezüglich der Addition in jeder Zeile jede Zahl genau einmal vorkommt, also auch 0. Bezüglich der Multiplikation stimmt dies für die Restklassen ungleich 0: In jeder Zeile kommt 1 vor.

Zu überprüfen sind die Assoziativ- und Distributivgesetze. Dies geht mit etwas Aufwand in der Tabelle. z.B. in  $\mathbb{F}_2$ :

$$0 + (0 + 1) = 0 + 1 = (0 + 0) + 1$$

Klarer ist es, wenn man die Restklasseninterpretation nutzt. Dann folgen die Rechenregeln aus den Regeln für  $\mathbb{Z}$ . Wir führen dies hier nicht aus.  $\square$

**Bemerkung.** Für jedes  $n \in \mathbb{N}$  bilden die Restklassen modulo  $n$  einen Ring. Er ist genau dann ein Körper, wenn  $n$  eine Primzahl ist. Allgemein gibt es für jede Primzahlpotenz  $p^k$  einen Körper mit  $p^k$  Elementen. In der Informatik spielen  $\mathbb{F}_2$  und  $\mathbb{F}_{256}$  eine große Rolle.

Nun ist es endlich so weit.

**Definition 2.10.** Sei  $k$  ein Körper. Ein  $k$ -Vektorraum ist eine abelsche Gruppe  $(V, +)$  zusammen mit einer skalaren Multiplikation

$$\cdot : k \times V \rightarrow V$$

so dass gilt:

(i) (Assoziativgesetz) Für alle  $v \in V$ ,  $a, b \in k$  gilt

$$a(bv) = (ab)v$$

(ii) (Distributivgesetze) Für alle  $a, b \in k$ ,  $x, y \in V$  gilt

$$(a + b)x = ax + bx, a(x + y) = ax + ay$$

(iii) Für alle  $x \in V$  gilt

$$1x = x$$

Ist  $k$  nur ein Ring, so heißt  $V$  Modul.

**Beispiel.** (i)  $k^n$  mit der Addition und skalaren Multiplikation von Matrizen ist ein Vektorraum.

(ii) Sei  $Ax = 0$  ein homogenes lineares  $m \times n$ -Gleichungssystem. Dann ist die Lösungsmenge

$$V = \{x \in k^n \mid Ax = 0\}$$

ein Vektorraum. (Lemma 1.13).

(iii)  $\mathbb{C}$  ist ein  $\mathbb{R}$ -Vektorraum,  $\mathbb{R}$  und  $\mathbb{C}$  sind  $\mathbb{Q}$ -Vektorräume.

(iv) Sei  $M$  eine Menge. Dann ist

$$V = \text{Abb}(M, k)$$

mit der Addition und skalaren Multiplikation von Funktionen

$$(f + g)(m) = f(m) + g(m), (af)(m) = af(m)$$

ein  $k$ -Vektorraum.

(v) Sei  $I \subset \mathbb{R}$  ein Intervall. Dann ist

$$C(I, \mathbb{R}) = \{f : I \rightarrow \mathbb{R} \mid f \text{ stetig} \}$$

ein  $\mathbb{R}$ -Vektorraum mit der Addition und skalaren Multiplikation von Funktionen (Analysis). Ebenso für differenzierbare Funktionen, stetig differenzierbare Funktionen, . . .

(vi) Die Lösungen einer linearen, homogenen Differentialgleichung bilden einen  $\mathbb{R}$ -Vektorraum (Analysis 2?)

Wieder ein paar elementare Regeln.

**Lemma 2.11.** *Sei  $k$  ein Körper,  $V$  ein  $k$ -Vektorraum. Für alle  $v \in V, a \in k$  gilt*

$$0v = 0, (-1)v = -v, a0 = 0$$

Ist  $a \in k, v \in V$  mit  $av = 0$ , so folgt  $a = 0$  oder  $v = 0$ .

*Beweis:* Wie im Fall von Ringen:

$$0v + 0v = (0 + 0)v = 0v \Rightarrow 0v = 0$$

$$(-1)v + v = (-1)v + 1v = (-1 + 1)v = 0v = 0 \Rightarrow (-1)v = -v$$

$$a0 + a0 = a(0 + 0) = a0 \Rightarrow a0 = 0$$

Ist  $av = 0, a \neq 0$ , so existiert  $a^{-1}$ . Es folgt

$$0 = a^{-1}0 = a^{-1}(av) = (a^{-1}a)v = 1v = v$$

□

**Definition 2.12.** *Sei  $k$  ein Körper,  $V$  ein  $k$ -Vektorraum. Eine Teilmenge  $U \subset V$  heißt Untervektorraum, wenn  $U$  durch Einschränkung der Addition und skalaren Multiplikation auf  $V$  zu einem Vektorraum wird.*

**Lemma 2.13.** *Sei  $k$  ein Körper,  $V$  ein  $k$ -Vektorraum. Eine nichtleere Teilmenge  $U \subset V$  ist genau dann ein Untervektorraum, wenn gilt: Für alle  $a, b \in k, x, y \in U$  liegt*

$$ax + by \in U$$

*Beweis:* Für  $a, b = 1$  sehen wir, dass die Addition auf  $V$  eine Addition auf  $U$  definiert. Für  $b = 0$  sehen wir, dass die skalare Multiplikation auf  $V$  eine Multiplikation auf  $U$  definiert. Die Assoziativ- und Distributivgesetze gelten, da sie auf  $V$  gelten. Schließlich ist  $U$  eine Gruppe, da  $0v = 0 \in U$  und  $-v = -1(v) \in U$ .  $\square$

**Beispiel.** In unserer Liste ist das Beispiel (ii) ein Untervektorraum von (i), (v) ist ein Untervektorraum von (iv) (mit  $k = \mathbb{R}$ ,  $M = I$ ).

Aus Vektorräumen können wir neue Vektorräume definieren.

**Definition 2.14.** Sei  $k$  ein Körper und  $V, W$  seien  $k$ -Vektorräume. Dann heißt der Vektorraum

$$V \times W = V \oplus W = \{(v, w) | v \in V, w \in W\}$$

mit der Addition

$$(v, w) + (x, y) = (v + x, w + y) \text{ für alle } v, x \in V, w, y \in W$$

und der skalaren Multiplikation

$$\lambda(v, w) = (\lambda v, \lambda w) \text{ für alle } \lambda \in k, v \in V, w \in W$$

heißt direkte Summe oder direktes Produkt von  $V$  und  $W$ .

**Beispiel.**  $k^2 = k \times k$ .

Allgemeiner:

**Definition 2.15.** Sei  $I$  eine Menge, für jedes  $i \in I$  sei  $V_i$  ein  $k$ -Vektorraum. Dann heißt

$$\prod_{i \in I} V_i = \{\phi : I \rightarrow \bigcup_{i \in I} V_i | \phi(i) \in V_i \text{ für alle } i \in I\} = \{(v_i)_{i \in I} | v_i \in V_i \text{ für alle } i \in I\}$$

mit der komponentenweisen Addition und der diagonalen Multiplikation mit Skalaren das direkte Produkt der  $V_i$ . Der Untervektorraum

$$\bigoplus_{i \in I} V_i = \{(v_i)_{i \in I} | v_i = 0 \text{ für fast alle } i \in I\}$$

heißt direkte Summe der  $V_i$ .

”Fast alle” bedeutet ”alle bis auf endlich viele”. Endliche direkte Summen und Produkte stimmen überein.

**Beispiel.**

$$k^n = \prod_{i=1}^n k = \bigoplus_{i=1}^n k$$

$$\text{Abb}(M, k) = \prod_{i \in M} k = k^M$$

Das folgende Beispiel bereitet erfahrungsgemäß Probleme. Die Konstruktion ist aber in vielen Beweisen nützlich.

**Definition 2.16.** Sei  $k$  ein Körper,  $V$  ein  $k$ -Vektorraum,  $U \subset V$  ein Untervektorraum. Für jedes  $v \in V$  heißt

$$\bar{v} = v + U = \{v + u \mid u \in U\} \subset V$$

Nebenklasse von  $v$  bezüglich  $U$ . Die Menge aller Nebenklassen  $V/U$  heißt Quotientenvektorraum.

**Beispiel.** Sei  $k = \mathbb{F}_2 = \{0, 1\}$ ,  $V = \mathbb{F}_2^2 = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$  und  $U = \{(a, a) \in V \mid a \in \mathbb{F}_2\} = \{(0, 0), (1, 1)\}$ . Wir betrachten die Nebenklassen:

$$\overline{(0, 0)} = (0, 0) + U = \{(0, 0) + (0, 0), (0, 0) + (1, 1)\} = \{(0, 0), (1, 1)\}$$

$$\overline{(1, 0)} = (1, 0) + U = \{(1, 0) + (0, 0), (1, 0) + (1, 1)\} = \{(1, 0), (0, 1)\}$$

$$\overline{(0, 1)} = (0, 1) + U = \{(0, 1) + (0, 0), (0, 1) + (1, 1)\} = \{(0, 1), (1, 0)\}$$

$$\overline{(1, 1)} = (1, 1) + U = \{(1, 1) + (0, 0), (1, 1) + (1, 1)\} = \{(1, 1), (0, 0)\}$$

Es gilt also  $\overline{(0, 0)} = \overline{(1, 1)}$ ,  $\overline{(1, 0)} = \overline{(0, 1)}$ . Daher folgt

$$V/U = \{\overline{(0, 0)}, \overline{(1, 0)}, \overline{(0, 1)}, \overline{(1, 1)}\} = \{\overline{(0, 0)}, \overline{(1, 0)}\}$$

Der Quotientenvektorraum hat also zwei Elemente.

**Lemma 2.17.** In der Situation der Definition ist  $V/U$  mit der Addition

$$\overline{x} + \overline{y} = \overline{x + y} \text{ für alle } x, y \in V$$

und der Multiplikation

$$a\overline{x} = \overline{ax} \text{ für alle } a \in k, x \in V$$

zu einem  $k$ -Vektorraum.

*Beweis:* Alle Axiome folgen aus den Axiomen für  $V$ . Z.B.

$$a(\overline{x} + \overline{y}) = \overline{a(x + y)} = \overline{ax + ay} = \overline{ax} + \overline{ay} = a\overline{x} + a\overline{y} = a\overline{x + y}$$

Das Problem ist ein anderes: Definiert unsere Vorschrift überhaupt eine Abbildung? Wir sagen: ist  $+$  wohldefiniert? Seien  $\overline{x} = \overline{x'}$  und  $\overline{y} = \overline{y'}$ . Zu zeigen ist

$$\overline{x + y} = \overline{x' + y'}$$

Dafür untersuchen wir zunächst die Frage, wann zwei Nebenklassen gleich sind.

**Behauptung.** Es gilt  $\overline{x} = \overline{x'}$  genau dann, wenn  $x - x' \in U$ .



Sei zunächst  $\bar{x} = \overline{x'}$ . Dann ist

$$x = x + 0 \in x + U = x' + U$$

d.h. es gibt  $u \in U$  mit  $x = x' + u$ . Hieraus folgt  $x - x' = u \in U$ . Sei nun umgekehrt  $x - x' \in U$ . Wir überprüfen  $x + U \subset x' + U$ . Sei also  $x + v \in x + U$  beliebig, d.h.  $v \in U$ . Dann gilt

$$x + v = x' + (x - x') + v \in x' + U$$

denn  $(x - x') + v \in U$  (Untervektorraumeigenschaft). Die umgekehrte Inklusion sieht man genauso.

Nun können wir uns wieder der Wohldefiniertheit zuwenden. Sei  $\bar{x} = \overline{x'}$ ,  $\bar{y} = \overline{y'}$ . Es gilt

$$(x + y) - (x' + y') = (x - x') + (y - y') \in U$$

da  $x - x' \in U$ ,  $y - y' \in U$  und  $U$  ein Untervektorraum. Damit haben wir  $\overline{x + y} = \overline{x' + y'}$  gezeigt.

Sei nun  $a \in k$ ,  $\bar{x} = \overline{x'}$ . Dann folgt

$$(ax) - (ax') = a(x - x') \in U$$

da  $x - x' \in U$  und  $U$  Untervektorraum. Damit haben wir  $a\bar{x} = \overline{ax'}$  gezeigt.  $\square$

Wir notieren uns für späteren Referenz:

**Lemma 2.18.** *Sei  $k$  ein Körper,  $V$  ein  $k$ -Vektorraum und  $U \subset V$  ein Untervektorraum. Für alle  $x, x' \in V$  gilt*

$$\bar{x} = \overline{x'} \Leftrightarrow x - x' \in U$$

*Beweis:* Siehe oben.  $\square$

**Bemerkung.** Analog definiert man

$$\mathbb{Z}/n\mathbb{Z} = \{x + n\mathbb{Z} \mid x \in \mathbb{Z}\}$$

wobei

$$x + n\mathbb{Z} = \{x + na \mid n \in \mathbb{Z}\}$$

Es handelt sich um eine Ring, für Primzahlen  $n$  sogar um Körper.

Diesselbe Konstruktion betrachten wir auch von einem anderen Standpunkt.

## Relationen

Sei  $X$  eine Menge,  $R \subset X \times X$  eine Relation (also einfach eine Teilmenge). Wir schreiben auch

$$xRy \Leftrightarrow (x, y) \in R.$$

**Definition 2.19.** *Die Relation  $R$  auf einer Menge  $X$  heißt:*

- (i) reflexiv, wenn  $xRx$  für alle  $x \in X$ ;
- (ii) symmetrisch, wenn  $xRy \Rightarrow yRx$  für alle  $x, y \in X$ ;
- (iii) transitiv, wenn  $xRy, yRz \Rightarrow xRz$  für alle  $x, y, z \in X$ ;
- (iv) Äquivalenzrelation, wenn sie reflexiv, symmetrisch und transitiv ist.

Wir schreiben dann oft  $x \sim_R y$  und sagen:  $x$  ist äquivalent zu  $y$  (bezüglich der Relation  $R$ ).

Sei  $R$  eine Äquivalenzrelation auf  $X$ ,  $x \in X$ . Dann heißt

$$\bar{x} = \{y \in X \mid y \sim_R x\}$$

Äquivalenzklasse von  $x$ .

**Lemma 2.20.** *Zwei Äquivalenzklassen sind entweder gleich oder disjunkt. Für alle  $x, y \in X$  gilt  $\bar{x} = \bar{y} \Leftrightarrow x \in \bar{y}$ .*

*Beweis:* Sei  $x, y, z \in X$  mit  $z \in \bar{x} \cap \bar{y}$ . Nach Definition gilt also

$$z \sim_R x, z \sim_R y.$$

Wegen Symmetrie und Transitivität folgt

$$x \sim_R z \sim_R y \Rightarrow x \sim_R z, z \sim_R x$$

Sei nun  $v \in \bar{x}$ ,  $w \in \bar{y}$ , also  $v \sim_R x$  und  $w \sim_R z$ . Dann folgt

$$\begin{aligned} v \sim_R x \sim_R z &\Rightarrow v \sim_R z \Rightarrow v \in \bar{z}, \\ w \sim_R z \sim_R x &\Rightarrow w \sim_R x \Rightarrow w \in \bar{x}. \end{aligned}$$

Daher gilt  $\bar{x} \subset \bar{z}$  und  $\bar{z} \subset \bar{x}$ , also Gleichheit.

Die zweite Aussage ist ein Spezialfall, da  $x \in \bar{x}$  (Reflexivität). □

**Definition 2.21.** *Sei  $V$  ein Vektorraum,  $U \subset V$  ein Untervektorraum. Für  $x, y \in V$  sagen wir, dass*

$$x \equiv y \pmod{U}$$

*(lies:  $x$  kongruent zu  $y$  modulo  $U$ ), wenn  $x - y \in U$ .*

**Lemma 2.22.** *Dies ist eine Äquivalenzrelation. Die Äquivalenzklassen sind die Nebenklassen.*

*Beweis:* Nachrechnen und Lemma 2.18. □

## Kapitel 3

# Basen und Dimension

Wir haben gesehen, das es in der Lösungsmenge von linearen Gleichungssystemen "Freiheitsgrade" gibt. Das ist die Dimension des Vektorraums, die wir nun definieren wollen.

Zunächst präzisieren wir eine Sprechweise.

**Definition 3.1.** Sei  $I$  eine Menge (Indexmenge),  $X$  eine andere Menge. Eine Familie von Elementen aus  $X$  indiziert durch  $I$  ist eine Abbildung

$$\phi : I \rightarrow X$$

Wir sagen: Sei  $x_i \in X$  für  $i \in I$  eine Familie von Elementen aus  $X$ . Dabei ist  $\phi(i) = x_i$ .

Ist  $J \subset I$  eine Teilmenge, so ist  $\phi|_J$  (Einschränkung von  $\phi$  auf  $J$ ) eine Teilfamilie von  $\phi$ . Wir sagen  $x_i$  für  $i \in J$  ist eine Teilfamilie von  $x_i$  für  $i \in I$ .

Eine Familie heißt endlich, falls  $I$  endlich viele Elemente hat.

Ist spezielle  $I = \mathbb{N}$ , so lautet die Familie

$$x_1, x_2, x_3, \dots$$

Eine solche Familie heißt *Folge*. Ist speziell  $I = \{1, \dots, n\}$ , so lautet die Familie

$$(x_1, x_2, \dots, x_n)$$

Es handelt sich also um eine endliche Folge bzw. ein  $n$ -Tupel.

**Definition 3.2.** Sei  $k$  ein Körper,  $V$  ein  $k$ -Vektorraum,  $(v_1, \dots, v_n)$  eine Familie von Elementen aus  $V$ . Ein Element  $v \in V$  heißt Linearkombination von  $(v_1, \dots, v_n)$ , falls es  $\lambda_1, \dots, \lambda_n \in k$  gibt mit

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n$$

Sei

$$\langle v_1, \dots, v_n \rangle_k = \langle v_1, \dots, v_n \rangle = \langle v_i | i = 1, \dots, n \rangle$$

die Menge aller Linearkombinationen von  $v_1, \dots, v_n$ .

Sei  $I$  eine beliebige Indexmenge,  $v_i \in V$  für  $i \in I$  eine Familie von Vektoren.

Dann heißt

$$\langle v_i | i \in I \rangle_k = \langle v_i | i \in I \rangle = \bigcup_{J \subset I \text{ endlich}} \langle v_i | i \in J \rangle$$

Menge der Linearkombinationen der  $i \in I$ . Wir sagen auch:  $\langle v_i | i \in I \rangle$  wird von  $v_i$  für  $i \in I$  aufgespannt.

Ist  $v_i \in V$  für  $i \in I$  eine unendliche Familie, so ist  $v$  genau dann eine Linearkombination der  $v_i$ , wenn es für alle  $i \in I$  Elemente  $\lambda_i \in k$  gibt, die fast alle Null sind (d.h. alle bis auf endlich viele sind 0), so dass

$$v = \sum_{i \in I} \lambda_i v_i$$

Da nur endlich viele Summanden ungleich 0 sind, handelt es sich um eine endliche Summe, die in jedem Vektorraum gebildet wird. Mit Konvergenz von Reihen hat dies nichts zu tun.

**Lemma 3.3.** Sei  $k$  ein Körper,  $V$  ein Vektorraum und  $v_i \in V$  für  $i \in I$  eine Familie von Vektoren. Dann ist  $\langle v_i | i \in I \rangle \subset V$  ein Untervektorraum. Ist  $U \subset V$  ein Untervektorraum mit  $v_i \in U$  für alle  $i \in I$ , so folgt  $\langle v_i | i \in I \rangle \subset U$ .

Mit anderen Worten:  $\langle v_i | i \in I \rangle$  ist der kleinste Untervektorraum von  $V$ , der alle  $v_i$  für  $i \in I$  enthält.

*Beweis:* Seien  $v, v' \in \langle v_i | i \in I \rangle$ . D.h. es gibt  $\lambda_i, \lambda'_i \in k$ , fast alle 0, so dass gilt

$$v = \sum_{i \in I} \lambda_i v_i \quad v' = \sum_{i \in I} \lambda'_i v_i$$

Dann sind auch fast alle  $\lambda_i + \lambda'_i$  gleich 0, und es folgt

$$v + v' = \sum_{i \in I} \lambda_i v_i + \sum_{i \in I} \lambda'_i v_i = \sum_{i \in I} (\lambda_i + \lambda'_i) v_i$$

Damit ist  $v + v'$  ebenfalls eine Linearkombination der  $v_i$  für  $i \in I$ .

Sei  $U$  Untervektorraum,  $v_i \in U$  für alle  $i \in I$ . Sei  $v = \sum_{i \in I} \lambda_i v_i$  eine Linearkombination in  $V$ . Da  $U$  ein Untervektorraum ist, liegen alle  $\lambda_i v_i$  in  $U$ . Fast alle verschwinden. Mit jedem Summanden liegt dann auch die Summe im Unterraum  $U$ .  $\square$

**Definition 3.4.** Sei  $k$  ein Körper,  $V$  ein  $k$ -Vektorraum. Eine Familie  $v_i \in V$  für  $i \in I$  heißt Erzeugendensystem von  $V$ , wenn

$$V = \langle v_i | i \in I \rangle$$

Wir sagen,  $v_i$  für  $i \in I$  erzeugen  $V$ .  $V$  heißt endlich erzeugt, falls es ein endliches Erzeugendensystem gibt.

**Beispiel.** (i)  $k^n$  wird erzeugt von den Vektoren  $e_i = (0, \dots, 0, 1, 0, \dots, 0)$  mit der 1 an der  $i$ -ten Stelle. Es gilt nämlich

$$(a_1, \dots, a_n) = a_1(1, 0, \dots, 0) + a_2(0, 1, 0, \dots, 0) + \dots + a_n(0, \dots, 0, 1)$$

Der Vektorraum ist endlich erzeugt.

(ii)  $\mathbb{C}$  wird aufgespannt von 1 und  $i$ .

(iii) Jeder Vektorraum hat ein Erzeugendensystem, nämlich ganz  $V$ . Formal:  $I = V$ ,  $v_v = v$  für alle  $v \in V$ . Es ist

$$v = 1v$$

Nun kommt die entscheidende Definition:

**Definition 3.5.** Sei  $k$  ein Körper,  $V$  ein Vektorraum. Eine endliche Familie  $v_1, \dots, v_n$  in  $V$  heißt linear unabhängig, wenn für alle  $(\lambda_1, \dots, \lambda_n) \in k^n$  gilt:

$$0 = \lambda_1 v_1 + \dots + \lambda_n v_n \Rightarrow \lambda_i = 0 \text{ für } i = 1, \dots, n$$

Mit anderen Worten: Es gibt keine nicht-triviale Darstellung von 0 als Linearkombination. Ein beliebige Familie  $v_i \in V$  für  $i \in I$  heißt linear unabhängig, wenn jede endliche Teilfamilie  $v_i$  für  $i \in J \subset I$  linear unabhängig ist. Eine Familie  $v_i \in V$   $i \in I$  heißt linear abhängig, wenn sich nicht linear unabhängig ist.

**Beispiel.** (i) In  $k^n$  sind die  $e_i$  des letzten Beispiels linear unabhängig. Ist nämlich

$$0 = \sum a_i e_i = (a_1, \dots, a_n)$$

so sind alle  $a_i = 0$ .

(ii) Gibt es  $i, j \in I$  mit  $i \neq j$ , aber  $v_i = v_j$ , so ist die Familie linear abhängig, da  $0 = 1v_i - 1v_j$ .

(iii) Eine einelementige Familie  $v_1$  ist linear unabhängig genau dann, wenn  $v_1 \neq 0$ : ist  $v_1 = 0$ , so ist  $0 = 1v_1$  eine nicht-triviale Linearkombination. Ist  $0 = \lambda_1 v_1$  und  $v_1 \neq 0$ , so folgt  $\lambda_1 = 0$ .

(iv) Ist  $I = \emptyset$ , so ist die Familie linear unabhängig, denn es gibt keine Linearkombinationen von Elementen, d.h. die Bedingung ist leer. Dies ist natürlich ein pathologischer Fall, der aber auch vorkommt!

**Bemerkung.** Ist  $v_i \in V$  für  $i \in I$  linear abhängig, so gibt es einen Index  $i$ , so dass sich  $v_i$  als Linearkombination der  $v_j$  für  $j \in J = I \setminus \{i\}$  schreiben lässt. Es gibt nämlich  $\lambda_j$ , nicht alle 0, so dass

$$0 = \sum_{j \in I} \lambda_j v_j$$

Sei  $i$  ein Index, für den  $\lambda_i \neq 0$ . Dann folgt

$$-\lambda_i v_i = \sum_{j \in J} \lambda_j v_j \Rightarrow v_i = \sum_{j \in J} -\lambda_i^{-1} \lambda_j v_j$$

**Lemma 3.6.** *Sei  $V$  ein  $k$ -Vektorraum. Sei  $v_1, \dots, v_n$  eine linear unabhängige Familie in  $V$ . Ist  $v \in \langle v_1, \dots, v_n \rangle$ , d.h. gibt es  $\lambda_1, \dots, \lambda_n \in k$  mit*

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n$$

so sind die  $\lambda_i$  für  $i = 1, \dots, n$  eindeutig bestimmt.

*Beweis:* Sei

$$v = \sum_{i=1}^n \lambda_i v_i = \sum_{i=1}^n \lambda'_i v_i \Rightarrow 0 = \sum_{i=1}^n (\lambda_i - \lambda'_i) v_i$$

Die Voraussetzung der linearen Unabhängigkeit impliziert, dass  $\lambda_i - \lambda'_i = 0$  für alle  $i$ , also  $\lambda_i = \lambda'_i$  für alle  $i$ .  $\square$

**Definition 3.7.** *Sei  $V$  ein  $k$ -Vektorraum. Eine Basis ist ein linear unabhängiges Erzeugendensystem.*

**Beispiel.** (i)  $k^n$  hat die Basis  $e_1, \dots, e_n$ .

(ii) Der Vektorraum  $0 = \{0\}$  hat die Basis  $\emptyset$ .

(iii) Sei  $Ax = 0$  ein homogenes lineares  $m \times n$ -Gleichungssystem. Wir haben gesehen (Theorem 1.7), dass das Gleichungssystem bis auf Umm nummerieren der Unbekannten äquivalent ist zu einem Gleichungssystem der Form

$$A'x = 0 \quad A' = \begin{pmatrix} 1 & a'_{12} & \dots & & a'_{1n} \\ 0 & 1 & a'_{23} & \dots & a'_{2n} \\ & & & \dots & \\ 0 & \dots & 1 & a'_{rr+1} & \dots & a'_{rn} \\ 0 & \dots & & \dots & 0 & \\ & & & \dots & & \\ 0 & \dots & & \dots & 0 & \end{pmatrix}$$

Zu jeder Wahl von  $x_{r+1}, \dots, x_n$  gibt es eine eindeutig bestimmte Lösung  $(x_1, \dots, x_n)$ . Dies bedeutet, dass wir eine Basis aus  $n-r$  Vektoren angeben können, nämlich für  $i = r+1, \dots, n$  die Lösungen  $e_i$  mit  $x_i = 1$ ,  $x_j = 0$  für  $r+1, \dots, n$ ,  $j \neq i$ .

(iv)  $k^{\mathbb{N}}$  ist nach Definition der Vektorraum der Folgen

$$(x_1, x_2, \dots)$$

In ihm gibt es die Folgen  $\delta^i$  mit  $\delta^i_i = 1$ ,  $\delta^i_j = 0$  für  $i \neq j$ . Diese sind linear unabhängig, aber keine Basis, denn die Folge  $(1, 1, 1, \dots)$  ist keine endliche Linearkombination der  $e_i$ .

**Lemma 3.8.** Sei  $k$  ein Körper,  $V$  ein  $k$ -Vektorraum,  $v_1, \dots, v_n$ . Ist  $v_1, \dots, v_n$  linear abhängig, so gibt es ein  $1 \leq i \leq n$  mit

$$\langle v_j | j = 1, \dots, n \rangle = \langle v_j | j = 1, \dots, n, j \neq i \rangle$$

*Beweis:* Nach Voraussetzung gibt es eine nicht-triviale Relation

$$\lambda_1 v_1 + \dots + \lambda_n v_n = 0$$

in der  $\lambda_i \neq 0$  für ein  $1 \leq i \leq n$ . Ohne Einschränkung der Allgemeinheit (Vertauschen der  $v_i$ ) ist dies  $\lambda_n$ . Dann können wir die Relation nach  $v_n$  auflösen:

$$v_n = \sum_{i=1}^{n-1} -\lambda_n^{-1} \lambda_i v_i$$

**Behauptung.**  $V = \langle v_1, \dots, v_{n-1} \rangle$ .

Sei  $v \in V$ . Dann gibt es nach Voraussetzung  $a_i \in k$  für  $i = 1, \dots, a_n$  mit

$$v = \sum_{i=1}^n a_i v_n = \sum_{i=1}^{n-1} a_i v_i + a_n \left( \sum_{i=1}^{n-1} -\lambda_n^{-1} \lambda_i v_i \right) = \sum_{i=1}^{n-1} (a_i - a_n \lambda_n^{-1} \lambda_i) v_i$$

□

**Bemerkung.** Ab jetzt ist es wichtig, dass  $k$  ein Körper ist! Wir haben mit dem Schluss  $\lambda_n \neq 0 \Rightarrow \lambda_n^{-1} \in k$  gearbeitet.

**Satz 3.9.** Sei  $V$  ein Vektorraum,  $v_i \in V$  für  $i \in I$  eine Familie. Äquivalent sind:

- (i) Die Familie ist eine Basis.
- (ii) Für jedes  $v \in V$  gibt es eine eindeutige Familie  $\lambda_i \in k$  für  $i \in I$ , fast alle  $\lambda_i = 0$ , so dass gilt

$$v = \sum_{i \in I} \lambda_i v_i$$

- (iii) Die Familie ist ein minimales Erzeugendensystem, d.h. jede echte Teilfamilie ist kein Erzeugendensystem mehr.
- (iv) Die Familie ist eine maximale linear unabhängige Familie, d.h. jede echt größere Familie ist linear abhängig.

**Bemerkung.** Aus Sicht der Physik ist die Charakterisierung (ii) entscheidend: Die Wahl einer Basis bedeutet die Einführung von *Koordinaten* durch die bijektive Abbildung

$$v \mapsto (\lambda_i)_{i \in I}$$

*Beweis:* (i)  $\Rightarrow$  (ii) Die Existenz der Darstellung ist die Erzeugendensystemeigenschaft. Die Eindeutigkeit folgt wie im endlichen Fall in Lemma 3.6.

(ii)  $\Rightarrow$  (i) Die Existenz der Darstellung ist wieder die Erzeugendensystemeigenschaft. Die Eindeutigkeit für  $v = 0$  ist genau die lineare Unabhängigkeit.

(i)  $\Rightarrow$  (iii) Sei  $v_i \in V$  für  $i \in I$  eine Basis, also ein Erzeugendensystem. z.z. ist die Minimalität. Angenommen, es gibt  $J \subsetneq I$  so dass die  $v_j$  für  $j \in J$  immer noch eine Erzeugendensystem bilden. Nach Voraussetzung gibt es  $j \in I \setminus J$ . Wir betrachten  $v_i$ . Es liegt im Erzeugnis der  $v_j$  für  $j \in J$ , d.h., es kann geschrieben werden als

$$v_i = \sum_{j \in J} a_j v_j \Rightarrow \sum_{j \in J} a_j v_j - 1v_i = 0$$

Dies ist ein Widerspruch zur linearen Unabhängigkeit.

(iii)  $\Rightarrow$  (i) Sei  $v_i$  für  $i \in I$  ein minimales Erzeugendensystem. z.z. ist die lineare Unabhängigkeit. Angenommen, es gibt eine nicht-triviale Relation

$$0 = \sum_{i \in I} \lambda_i v_i$$

Dann gibt es  $i \in I$ , so dass wie oben  $v_i$  durch die anderen ausgedrückt werden kann. Sei  $J = I \setminus \{i\}$ . Diese Teilfamilie ist immer noch ein Erzeugendensystem.

(i)  $\Rightarrow$  (iv) Sei  $v_i$  für  $i \in I$  eine Basis, also linear unabhängig. z.z. ist die Maximalität. Sei  $J \subsetneq I$  und  $v_j$  für  $j \in J$  eine linear unabhängige Familie, die die gegebene Familie als Teilfamilie hat. Nach Voraussetzung gibt es  $j \in J \setminus I$ . Dieses  $v_j$  ist linear unabhängig von  $v_i$  für  $i \in I$ , liegt also nicht in  $\langle v_i | i \in I \rangle$ . Dies ist ein Widerspruch zur Erzeugendensystemeigenschaft unserer Basis.

(iv)  $\Rightarrow$  (i) Sei  $v_i$  für  $i \in I$  eine maximale linear unabhängige Familie. z.z. ist Erzeugendensystem. Sei  $v \in V$  beliebig. Falls  $v \notin \langle v_i | i \in I \rangle$ , so ist  $v$  linear unabhängig von  $v_i$  für  $i \in I$ . Wir setzen  $J = I \cup \{v\}$  und erhalten eine echt größere linear unabhängige Familie, im Widerspruch zur Maximalität.  $\square$

**Theorem 3.10.** *Sei  $k$  ein Körper. Dann hat jeder  $k$ -Vektorraum eine Basis.*

*Beweis:* (1. Anlauf) Wir betrachten den endlich erzeugten Fall. Sei  $V = \langle v_1, \dots, v_n \rangle$ . Angenommen, die  $v_i$  sind nicht linear unabhängig, etwa  $v_n \in \langle v_1, \dots, v_{n-1} \rangle$ , also nach Lemma 3.8  $V = \langle v_1, \dots, v_{n-1} \rangle$ . Nun wiederholen wir das Verfahren, bis wir bei einer linear unabhängigen Teilfamilie von  $v_1, \dots, v_n$  angekommen sind, die immer noch ganz  $V$  erzeugt. Dies ist die gesuchte Basis.  $\square$

Wir betrachten immer noch den endlich erzeugten Fall, aber diesmal formal sauber mit vollständiger Induktion.

**Lemma 3.11** (vollständige Induktion). *Sei  $M \subset \mathbb{N}$  eine Teilmenge, für die gilt:*

(i) *Induktionsanfang:  $1 \in M$ .*

(ii) *Induktionsschluss: Für jedes  $m \in M$  gilt  $m + 1 \in M$ .*

*Dann ist  $M = \mathbb{N}$ .*



*Beweis:* Dies ist gar kein Lemma, sondern ein Axiom der Definition der natürlichen Zahlen.  $\square$

*Beweis von Theorem 3.10.* (2. Anlauf) Wir führen den Beweis durch vollständige Induktion nach der Anzahl der Erzeuger.

**Induktionsanfang:**  $n = 1$ ,  $V = \langle v_1 \rangle$ . Falls  $v_1 = 0$ , so ist  $V = 0$  und  $\emptyset$  ist die gesuchte Basis. Falls  $v_1 \neq 0$ , so ist die Familie linear unabhängig, und dies ist die gesuchte Basis.

**Induktionsvoraussetzung:** Sei  $n \in \mathbb{N}$  und jeder Vektorraum, der von einer  $n$ -elementigen Familie erzeugt wird, habe eine Basis.

**Induktionsschluss:** Sei nun  $V$  ein Vektorraum, der von einer  $n+1$ -elementigen Familie erzeugt wird. Ist diese linear unabhängig, so haben wir die gesuchte Basis gefunden. Ist sie nicht linear unabhängig, so wird nach Lemma 3.8  $V$  bereits von einer  $n$ -elementigen Familie erzeugt. Also hat  $V$  eine Basis nach Induktionsvoraussetzung.

Durch vollständige Induktion ist das Theorem für alle endlich erzeugten Vektorräume bewiesen.  $\square$

**Bemerkung.** In der fortgeschrittenen Literatur wird oft nur der Teil der vollständigen Induktion durchgeführt, der nicht-trivial ist. Je nach Situation kann dies der Induktionsanfang oder der Induktionsschluss sein.

*Beweis von Theorem 3.10.* (3. Anlauf) Sei  $V$  nun ein beliebiger Vektorraum. Dann gibt es ein Erzeugendensystem, nämlich ganz  $V$ . Die vollständige Induktion wird nun durch ein anderes Axiom der Mengenlehre ersetzt, nämlich die transfinite Induktion. Sie ist äquivalent zum "Zornschen Lemma" oder zum "Auswahlaxiom". Dieses formale Argument soll hier nicht vorgeführt werden.  $\square$

#### Exkurs für Interessierte

**Zornsches Lemma:** Sei  $(M, \leq)$  eine nichtleere partiell geordnete Menge, d.h.  $\leq$  ist eine transitive, reflexive Relation (Aus  $a \leq b$  und  $b \leq c$  folgt  $a \leq c$  und  $a \leq a$  gilt für alle  $a$ ) für die gilt: Für alle  $a, b \in M$  mit  $a \leq b$  und  $b \leq a$  folgt  $a = b$ . Für jedes total geordnete Teilmengen  $N \subset M$  (d.h. für alle  $a, b \in N$  gilt  $a \leq b$  oder  $b \leq a$ ) gebe es eine obere Schranke, d.h. ein  $a \in M$  mit  $a_i \leq a$  für alle  $i = 1, 2, \dots$ . Dann hat  $M$  ein maximales Element, d.h. ein Element  $m$  für das gilt:  $m \leq a \Rightarrow m = a$ .

Dies ist ein Axiom der Mengenlehre, der Name hat historische Gründe.

Wir wenden dies an auf die Existenz von Basen: Sei  $V$  ein Vektorraum,

$$M = \{(U, B) \mid U \subset V \text{ Untervektorraum, } B \text{ Basis von } U\}$$

Diese ist partiell geordnet durch  $\subset$ , d.h.  $(U_1, B_1) \leq (U_2, B_2)$  falls  $U_1 \subset U_2$  und  $B_1$  eine Teilfamilie von  $B_2$ . Die Menge  $M$  ist nicht leer, denn sie enthält das Paar  $(0, \emptyset)$ . Wir überprüfen die Voraussetzung des Zornschen Lemmas. Gegeben eine total geordnete Teilmenge  $N \subset M$ . Dann ist das Paar  $(U, B)$  mit  $U = \bigcup_{(U, B) \in N} U$  und  $B = \bigcup_{(U, B) \in N} B$  eine obere Schranke. (Man überprüft leicht, dass  $U$  ein Untervektorraum und  $B$  eine Basis von  $U$ .) Nach dem Zornschen Lemma gibt es ein maximales Element  $(U_0, B_0)$  für ganz  $M$ .

**Behauptung.** Es ist  $U_0 = V$  (und damit hat  $V$  eine Basis.)

Angenommen,  $U_0 \neq V$ . Dann gibt es einen Vektor  $v \in V \setminus U_0$ . Dieser ist nicht linear abhängig von  $U_0$ . Damit ist  $\langle U_0, v \rangle \subsetneq U_0$  und  $v, B_0$  ist eine Basis. Dies ist ein Widerspruch zur Maximalität von  $(U_0, B_0)$ . Dies beweist das Theorem.

**Endes des Exkurses**

**Definition 3.12.** Sei  $k$  ein Körper,  $V$  ein  $k$ -Vektorraum,  $v_i \in V$  für  $i \in I$  eine Basis. Dann heißt die Anzahl der Elemente von  $I$  Dimension von  $V$ .

$$\dim_k V = \dim V = |I|$$

Dabei heißt die Dimension unendlich, falls die Menge  $I$  unendlich ist.

**Beispiel.**  $k^n$  hat die Dimension  $n$ . Ist  $Ax = 0$  ein  $m \times n$ -Gleichungssystem, wobei wir  $A$  in die Normalform des Gauß-Algorithmus gebracht haben. Dann ist die Dimension des Lösungsraums  $n - r$ .

**Wichtige Frage:** Ist  $\dim V$  unabhängig von der Wahl der Basis? Physikalisch formuliert: Ist die Anzahl der Koordinaten von  $V$  unabhängig von der Wahl des Koordinatensystems?

**Bemerkung.** Dies ist genau die Frage nach der Wohldefiniertheit der Zahl  $r$  im 1. Kapitel!

**Lemma 3.13** (Austauschlemma). Sei  $V$  ein  $k$ -Vektorraum mit Basis  $v_1, \dots, v_n$  und  $w \in V$ . Ist in der Darstellung  $w = \lambda_1 v_1 + \dots + \lambda_n v_n$  der Koeffizient  $\lambda_i \neq 0$ , so ist auch  $v_1, \dots, v_{i-1}, w, v_{i+1}, \dots, v_n$  eine Basis von  $V$ . Wir können  $v_i$  gegen  $w$  austauschen.

*Beweis:* Ohne Einschränkung ist  $i = 1$ . Nach Voraussetzung gilt

$$v_1 = \lambda_1^{-1} w - \lambda_1^{-1} \lambda_2 v_2 + \dots - \lambda_1^{-1} \lambda_n v_n$$

Sei  $v \in V$  beliebig, also

$$\begin{aligned} v &= \sum_{i=1}^n a_i v_i = a_1 \left( \lambda_1^{-1} w - \sum_{i=2}^n \lambda_1^{-1} \lambda_i v_i \right) + \sum_{i=2}^n a_i v_i \\ &= a_1 \lambda_1^{-1} w - \sum_{i=2}^n (a_1 \lambda_1^{-1} \lambda_i + a_i) v_i \end{aligned}$$

Daher ist  $w, v_2, \dots, v_n$  ein Erzeugendensystem. Nun überprüfen wir die lineare Unabhängigkeit. Sei

$$0 = \mu_1 w + \sum_{i=2}^n \mu_i v_i = \mu_1 \sum_{i=1}^n \lambda_i v_i + \sum_{i=2}^n \mu_i v_i$$

Wegen der linearen Unabhängigkeit der  $v_i$  für  $i = 1, \dots, n$  folgt zunächst  $\mu_1 \lambda_1 = 0$ . Nach Voraussetzung ist  $\lambda_1 \neq 0$ , also  $\mu_1 = 0$  (Multiplikation mit  $\lambda_1^{-1}$ ). Damit reduziert sich die Gleichung zu

$$0 = \sum_{i=2}^n \mu_i v_i \Rightarrow \mu_i = 0 \text{ für alle } i = 2, \dots, n$$

denn  $v_1, \dots, v_n$  ist linear unabhängig. □

Insbesondere bleibt bei dieser Operation die Anzahl der Basisvektoren erhalten!

**Satz 3.14** (Basisaustauschsatz). *Sei  $k$  Körper,  $V$  ein  $k$ -Vektorraum, darin  $v_1, \dots, v_n$  eine Basis und  $w_1, \dots, w_m$  linear unabhängig. Dann gibt es (paarweise verschiedene) Indizes  $i_1, \dots, i_m \in \{1, \dots, n\}$ , so dass man durch Austausch der  $v_{i_l}$  gegen  $w_l$  für  $l = 1, \dots, m$  wieder eine Basis erhält.*

*Beweis:* Wir führen (für alle  $n$  gleichzeitig) vollständige Induktion nach  $m$ . Man beachte: Wenn die Aussage für ein Paar  $(m, n)$  gilt, so muss insbesondere  $m \leq n$  gelten.

Induktionsanfang: Für  $m = 0$  ist nichts zu zeigen. Für  $m = 1$  ist dies genau das Austauschlemma.

Induktionsvoraussetzung: Die Aussage gelte für  $m-1$ , insbesondere für  $w_1, \dots, w_{m-1}$ . Nach geeigneter Umm Nummerierung der  $v_i$  ist also nun  $w_1, \dots, w_{m-1}, v_m, \dots, v_n$  eine Basis. Nach dem Austauschlemma können wir einen der Basisvektoren gegen  $w_m$  austauschen. Genauer:

$$w_m = \sum_{i=1}^{m-1} \lambda_i w_i + \sum_{i=m}^n \lambda_i v_i$$

Gilt  $\lambda_i = 0$  für alle  $i \geq m$  (dies schließt den Fall  $n < m$  ein), so wäre dies ein Widerspruch zur linearen Unabhängigkeit der  $w_i$ . Also gibt es  $i \geq m$  mit  $\lambda_i \neq 0$ . Nach dem Austauschlemma können wir  $v_i$  gegen  $w_m$  austauschen. Dies beendet den Beweis des Induktionsschlusses.

Mit vollständiger Induktion ist die Aussage allgemein gezeigt.  $\square$

**Korollar 3.15.** *Im Austauschatz gilt  $m \leq n$ .*

**Korollar 3.16** (Invarianz der Dimension). *Ist ein  $k$ -Vektorraum endlich erzeugt, so hat jede Basis eine endliche Anzahl von Elementen und die Anzahl ist unabhängig von der Wahl der Basis.*

Mit anderen Worten: die Dimension ist wohldefiniert.

*Beweis:* Gibt es ein endliches Erzeugendensystem, so auch eine endliche Basis  $v_1, \dots, v_n$ . Sei  $w_i$  für  $i \in I$  eine andere Basis. Jede endliche Teilmenge von  $I$  hat nach dem vorherigen Korollar höchstens  $n$  Elemente. Damit hat auch  $I$  höchstens  $n$  Elemente. Nun können wir den Austauschatz auch in die Gegenrichtung anwenden und erhalten  $|I| = n$ .  $\square$

Dies war der Beweis in der Sprache der Mathematiker, so wie er sich im Buch von Fischer findet. Dasselbe Argument lässt sich auch in der Sprache der Physiker formulieren, siehe z.B. das Buch von Lorenz. Wir wollen wenigstens den Ansatz skizzieren:

**Definition 3.17.** *Sei  $V$  ein Vektorraum,  $v_1, \dots, v_n$  und  $w_1, \dots, w_m$  Basen von  $V$ . Wir definieren die Basiswechselmatrix  $(a_{ij})_{i=1, j=1}^{m, n}$  durch*

$$v_j = \sum_{i=1}^m a_{ij} w_i$$

Die Aussage über die Dimension ist  $m = n$ . Zu zeigen ist also, dass Basiswechselmatrizen quadratisch sind.

Wir gehen nun schrittweise von einem Koordinatensystem zum anderen über und kontrollieren jeweils die Basiswechselmatrizen.

- (i) Vertauschen der Basisvektoren  $v_j$  und  $v_{j'}$  vertauscht die Spalten  $j$  und  $j'$  der Basiswechselmatrix.
- (ii) Vertauschen der Basisvektoren  $w_i$  und  $w_{i'}$  vertauscht die Zeilen  $i$  und  $i'$  der Basiswechselmatrix.
- (iii) Tauschen wir  $v_j$  gegen  $v'_j = v_j + \alpha v_l$  für  $l \neq j$ ,  $\alpha \in k$ , so lautet die neue Spalte  $j$  der Basiswechselmatrix

$$v'_j = \sum_{i=1}^m a_{ij} w_i + \alpha \sum_{i=1}^m a_{il} w_i = \sum_{i=1}^m (a_{ij} + \alpha a_{il}) w_i$$

d.h. wir addieren das  $\alpha$ -fache der  $l$ -ten Spalte zur  $i$ -ten.

- (iv) Tauschen wir  $w_s$  gegen  $w'_s = w_s + \beta w_t$  für  $s \neq t$ ,  $\beta \in k$  so lautet die neue Basiswechselmatrix

$$\begin{aligned} v_j &= \sum_{i=1}^m a_{ij} w_i = \sum_{i \neq s} a_{ij} w_i + a_{sj} (w'_s - \beta w_t) \\ &= \sum_{i \neq s, t} a_{ij} w_i + a_{sj} w'_s + (a_{tj} - \beta a_{sj}) w_t \end{aligned}$$

d.h. wir addieren das  $-\beta$ -fache der Zeile  $s$  zur Zeile  $t$ .

Dies sind genau die elementaren Zeilen- und Spaltentransformationen, wie wir sie im 1. Kapitel verwendet haben. Der Gaußalgorithmus führt daher (bei unverändertem  $n, m$ ) auf eine Basiswechselmatrix der Normalform aus Satz 1.16. Dies bedeutet

$$v_1 = w_1, v_2 = w_2, \dots, v_r = w_r, v_{r+1} = 0, \dots, v_n = 0$$

Hieraus folgt  $r = n$ , da die  $v_j$  linear unabhängig sind. Damit ist  $v_1, \dots, v_n$  eine Teilfolge von  $w_1, \dots, w_m$ . Da es sich um eine Basis handelt, ist es eine maximale linear unabhängige Folge, also  $m = n$ .

## Kapitel 4

# Lineare Abbildungen und Dimensionsformel

Bisher haben wir uns mit einzelnen Vektorräumen beschäftigt, nun studieren wir Abbildungen zwischen ihnen.

**Definition 4.1.** Sei  $k$  ein Körper,  $V, W$  seien  $k$ -Vektorräume. Eine Abbildung  $f : V \rightarrow W$  heißt linear, falls sie mit Addition und skalarer Multiplikation verträglich ist, d.h. für alle  $a, b \in k$  und  $x, y \in V$  gilt

$$f(ax + by) = af(x) + bf(y)$$

Die Menge der linearen Abbildungen heißt  $\text{Hom}_k(V, W)$ .

**Bemerkung.** Ist  $f$  linear, so gilt  $f(-x) = -f(x)$  ( $a = -1, b = 0$ ) und  $f(0) = 0$  ( $a, b = 0$ ).

**Beispiel.** (i)  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = 2x$  ist linear, denn

$$f(ax + by) = 2(ax + by) = a2x + b2y = af(x) + bf(y)$$

(ii)  $g : \mathbb{R} \rightarrow \mathbb{R}$  mit  $g(x) = x^2$  ist nicht linear, denn

$$g(x + y)^2 = x^2 + 2xy + y^2 \neq x^2 + y^2 = g(x) + g(y)$$

(iii)  $h : \mathbb{R} \rightarrow \mathbb{R}$  mit  $h(x) = 2x + 1$  ist nicht linear, denn

$$h(x+y) = 2(x+y)+1 = 2x+2y+1 \neq h(x)+h(y) = 2x+1+2y+1 = 2x+2y+2$$

(iv) Ein besonders wichtiges Beispiel: Sei  $V = k^n$ ,  $W = k^m$  und  $A$  eine  $m \times n$ -Matrix. Dann ist

$$F_A : k^n \rightarrow k^m \quad x \mapsto Ax$$

eine lineare Abbildung nach den Rechenregeln der Matrixmultiplikation in Satz 1.11. Sei  $e_1, \dots, e_n$  die bereits vorher benutzte Standardbasis des  $k^n$ , also  $e_s$  der Spaltenvektor, mit Einträgen

$$\delta_{js} = \begin{cases} 1 & s = j \\ 0 & \text{sonst} \end{cases}$$

in Zeile  $j$

Dann gilt

$$F_A(e_s) = Ae_s = \left( \sum_{j=1}^m a_{ij} \delta_{js} \right) = (a_{is})$$

d.h. das Bild von  $e_s$  ist genau die Spalte  $s$  von  $A$ .

(v)  $V = \text{Diff}(I, \mathbb{R})$  der Raum der differenzierbaren Funktionen auf einem Intervall. Dann ist  $f \mapsto f'$  eine lineare Abbildung (Rechenregeln der Differentiation).

**Lemma 4.2.** *Eine lineare Abbildung  $f : V \rightarrow W$  ist eindeutig bestimmt durch die Werte auf einer Basis von  $V$ . Ist  $v_i \in V$  für  $i \in I$  eine Basis und  $w_i \in W$  für  $i \in I$  eine Familie, so gibt es genau eine lineare Abbildung  $f : V \rightarrow W$  mit  $f(v_i) = w_i$ .*

*Beweis:* Die erste Aussage ist in der zweiten enthalten. Sei  $v \in V$ , also  $v = \sum_{i \in I} \lambda_i v_i$ . Damit folgt

$$f(v) = f\left(\sum_{i \in I} \lambda_i v_i\right) = \sum_{i \in I} \lambda_i f(v_i)$$

Dies zeigt die Eindeutigkeit von  $f$ . Andererseits wird durch die Formel

$$f(v) = \sum_{i \in I} \lambda_i w_i$$

tatsächlich eine lineare Abbildung definiert. Sie ist wohldefiniert wegen der Eindeutigkeit der Darstellung von  $v$ .  $\square$

**Erinnerung:** Eine Abbildung  $f : X \rightarrow Y$  heißt *injektiv*, wenn es für jedes  $y \in Y$  höchstens ein  $x \in X$  mit  $f(x) = y$  gibt. Sie heißt *surjektiv*, wenn es für jedes  $y \in Y$  mindestens ein  $x \in X$  gibt mit  $f(x) = y$ . Sie heißt *bijektiv*, wenn sie injektiv und surjektiv ist, d.h. wenn es für jedes  $y \in Y$  genau ein  $x \in X$  gibt mit  $f(x) = y$ . In diesem Fall ist  $f^{-1}$  mit  $f^{-1}(y) = x$  die *Umkehrabbildung* von  $f$ .

**Definition 4.3.** *Seien  $V, W$  Vektorräume. Eine injektive lineare Abbildung  $f : V \rightarrow W$  heißt Monomorphismus. Eine surjektive lineare Abbildung heißt Epimorphismus. Eine bijektive lineare Abbildung heißt Isomorphismus. Eine lineare Abbildung  $g : V \rightarrow V$  heißt Endomorphismus. Die Menge der Endomorphismen wird  $\text{End}_k(V)$  bezeichnet. Ein bijektiver Endomorphismus heißt Automorphismus. Die Menge der Automorphismen wird mit  $\text{Aut}_k(V)$  bezeichnet.*

**Beispiel.**  $\text{id} : V \rightarrow V$  mit  $\text{id}(x) = x$  ist ein Automorphismus.

**Lemma 4.4.** Sei  $k$  Körper,  $U, V, W$  Vektorräume.

(i) Sind  $f : U \rightarrow V$  und  $g : V \rightarrow W$  lineare Abbildungen, dann auch  $g \circ f : U \rightarrow W$ .

(ii) Ist  $f : U \rightarrow V$  ein Isomorphismus, so auch  $f^{-1} : V \rightarrow U$ .

(iii) Sind  $f, g : U \rightarrow V$  lineare Abbildungen,  $a, b \in k$ , so auch  $af + bg$ . Mit anderen Worten:  $\text{Hom}_k(V, W)$  ist ein  $k$ -Vektorraum.

*Beweis:* (i) Nach Definition ist  $g \circ f(u) = g(f(u))$ . Wir rechnen nach:

$$g \circ f(ax + by) = g(f(ax + by)) = g(af(x) + bf(y)) = ag(f(x)) + bg(f(y))$$

(ii) Ist  $f$  bijektiv, so auch  $f^{-1}$ . Zu zeigen ist, dass  $f^{-1}$  eine lineare Abbildung ist. Sei  $a, b \in k$ ,  $x, y \in V$ . Nach Definition ist  $f^{-1}(x) = x' \in U$  wobei  $f(x') = x$ , ebenso  $f^{-1}(y) = y'$  wobei  $f(y') = y$ . Dann folgt

$$f(ax' + by') = af(x') + bf(y') = ax + by$$

Nach Definition von  $f^{-1}$  ist dann

$$f^{-1}(ax + by) = ax' + by' = af(x) + bf(y)$$

Seien schließlich

(iii) Wir rechnen wieder nach:

$$\begin{aligned} (af + bg)(\alpha x + \beta y) &= af(\alpha x + \beta y) + bg(\alpha x + \beta y) \\ &= a\alpha f(x) + a\beta f(y) + b\alpha g(x) + b\beta g(y) \end{aligned}$$

□

**Bemerkung.**  $(\text{Aut}_k(V), \circ)$  ist eine Gruppe.

**Beispiel.** Sei  $A$  eine  $m \times p$ -Matrix,  $B$  eine  $p \times n$ -Matrix. Dann betrachten wir wie oben  $F_A : k^p \rightarrow k^m$ ,  $F_B : k^n \rightarrow k^p$ . Dann gilt

$$F_{AB} = F_A \circ F_B$$

Es gilt nämlich

$$F_A \circ F_B(x) = F_A(F_B(x)) = F_A(Bx) = A(Bx) = (AB)x = F_{AB}(x)$$

Entscheidend ist die Assoziativität der Matrizenmultiplikation!

**Definition 4.5.** Sei  $f : V \rightarrow W$  eine lineare Abbildung. Dann heißt

$$\text{Ker } f = \{x \in V \mid f(x) = 0\}$$

Kern der Abbildung. Es heißt

$$\text{Im } f = \{y \in W \mid \text{es gibt } x \in V \text{ mit } f(x) = y\}$$

Bild der Abbildung.

**Beispiel.** Sei  $A$  eine  $m \times n$ -Matrix.  $F_A : k^n \rightarrow k^m$  die zugehörige lineare Abbildung. Dann ist

$$\text{Ker } F_A = \{x \in k^n \mid Ax = 0\}$$

genau die Lösungsmenge des homogenen Gleichungssystems.

**Lemma 4.6.** Sei  $f : V \rightarrow W$  linear. Dann ist  $f$  injektiv genau dann, wenn  $\text{Ker } f = 0$ .

*Beweis:* Es gilt nach Definition

$$\text{Ker } f = f^{-1}(0)$$

die Menge aller Urbilder von 0.

Sei  $f$  injektiv. Dann hat  $0 \in W$  höchstens ein Urbild in  $V$ . Wegen  $f(0) = 0$  hat 0 also genau ein Urbild und  $\text{Ker } f = 0$ .

Sei umgekehrt  $\text{Ker } f = 0$ . Seien  $x, y \in V$  mit  $f(x) = f(y)$ . Dann gilt

$$f(x - y) = f(x) - f(y) = 0$$

d.h.  $x - y \in \text{Ker } f = 0$ , also  $x - y = 0 \Leftrightarrow x = y$ . Die Abbildung ist injektiv.  $\square$

**Lemma 4.7.** Sei  $f : V \rightarrow W$  linear. Dann sind  $\text{Ker } f \subset V$  und  $\text{Im } f \subset W$  Untervektorräume. Ist  $v_i$  für  $i \in I$  ein Erzeugendensystem von  $V$ , so ist  $f(v_i)$  für  $i \in I$  ein Erzeugendensystem von  $\text{Im } f$ .

*Beweis:* Wir beginnen mit dem Kern. Sei  $a, b \in k$ ,  $x, y \in \text{Ker } f$ . Dann gilt

$$f(ax + by) = af(x) + bf(y) = a0 + b0 = 0 + 0 = 0$$

d.h.  $ax + by \in \text{Ker } f$ . Seien nun  $x', y' \in \text{Im } f$ . Dann gilt es  $x, y \in V$  mit  $f(x) = x'$  und  $f(y) = y'$ . Es folgt

$$f(ax + by) = af(x) + bf(y) = ax' + by'$$

d.h.  $ax' + by'$  liegt in  $\text{Im } f$ .

Sei  $x' \in \text{Im } f$ , also  $x' = f(x)$  für ein  $x \in V$ . Dann ist  $x = \sum_{i \in I} \lambda_i v_i$  für geeignete  $\lambda_i$ . Es folgt

$$x' = f(x) = f\left(\sum_{i \in I} \lambda_i v_i\right) = \sum_{i \in I} \lambda_i f(v_i)$$

d.h.  $x' \in \langle f(v_i) \mid i \in I \rangle$ . Dies beweist  $\text{Im } f \subset \langle f(v_i) \mid i \in I \rangle$ . Die Umkehrung folgt, da jedes  $f(v_i) \in \text{Im } f$ ,  $\text{Im } f$  ein Untervektorraum, und  $\langle f(v_i) \mid i \in I \rangle$  der kleinste Vektorraum, der alle  $f(v_i)$  enthält.  $\square$



**Definition 4.8.** (i) Sei  $f : V \rightarrow W$  linear. Der Rang von  $f$  ist

$$\operatorname{rg} f = \dim_k \operatorname{Im} f$$

(ii) Sei  $A$  eine  $m \times n$ -Matrix. Dann heißt die maximale Zahl von linear unabhängigen Spalten Spaltenrang oder kurz Rang der Matrix.

**Lemma 4.9.** Es gilt  $\operatorname{rg} A = \operatorname{rg} F_A$ . Insbesondere ist der Rang einer Matrix wohldefiniert.

*Beweis:* Das Bild  $\operatorname{Im} F_A$  wird erzeugt von den  $F_A(e_i)$ , also den Spalten der Matrix. Wir erhalten eine Basis, wenn wir hierin eine linear unabhängige Teilmenge auswählen, die weiterhin  $\operatorname{Im} F_A$  erzeugt. Dies ist eine maximale linear unabhängige Teilmenge. Deren Anzahl heißt einerseits Dimension, andererseits Rang.  $\square$

**Satz 4.10** (Dimensionsformel). Sei  $f : V \rightarrow W$  eine lineare Abbildung. Dann gilt

$$\dim_k \operatorname{Ker} f + \operatorname{rg} f = \dim V$$

**Bemerkung.** Als Matrizen formuliert: Ist  $A$  eine  $m \times n$ -Matrix vom Rang  $r$ , so ist die Dimension des Lösungsraums des homogenen Gleichungssystems  $n - r$ . Beim Gauß-Algorithmus blieb der Lösungsraum (bis auf Vertauschen von Zeilen) unverändert. Damit bleibt auch der Rang der Matrizen konstant. Für Matrizen in Normalform liest man also den Rang also einfach ab. Im Kapitel 1 wurde er mit  $r$  bezeichnet. Für Matrizen in Normalform sieht man auch die Dimensionsformel sofort.

*Beweis:* (1. Anlauf) Ist  $\operatorname{Ker} f$  unendlich dimensional, so auch  $V$  und die Aussage gilt. Sei also  $\dim_k \operatorname{Ker} f = n < \infty$  und  $v_1, \dots, v_n$  eine Basis. Sei zunächst  $\dim V < \infty$ . Dann ergänzen wir (Basisaustauschsatz)  $v_1, \dots, v_n$  zu einer Basis  $v_1, \dots, v_N$  von ganz  $V$ .

**Behauptung.**  $f(v_{n+1}), \dots, f(v_N)$  ist eine Basis von  $\operatorname{Im} f$ .

Sei  $x' \in \operatorname{Im} f$ , also gibt es  $x \in V$  mit  $f(x) = x'$ . Wir schreiben

$$x = \sum_{i=1}^N \lambda_i v_i \Rightarrow f(x) = f\left(\sum_{i=1}^N \lambda_i v_i\right) = \sum_{i=1}^N \lambda_i f(v_i) = \sum_{i=n+1}^N \lambda_i f(v_i)$$

da  $f(v_i) = 0$  für  $i \leq n$ . Wie gewünscht ist

$$x' \in \langle f(v_i) \mid i = n+1, \dots, N \rangle$$

Nun überprüfen wir die lineare Unabhängigkeit. Sei hierfür

$$0 = \sum_{i=n+1}^N \lambda_i f(v_i) = f\left(\sum_{i=n+1}^N \lambda_i v_i\right)$$

d.h.  $\sum_{i=1}^N \lambda_i x_i \in \text{Ker } f = \langle v_1, \dots, v_n \rangle$ . Damit gibt es  $\mu_1, \dots, \mu_n$  mit

$$\mu_1 v_1 + \dots + \mu_n v_n = \lambda_{n+1} v_{n+1} + \dots + \lambda_N v_N$$

Wegen der linearen Unabhängigkeit von  $v_1, \dots, v_N$  folgt  $\lambda_{n+1} = \dots = \lambda_N = 0$  wie gewünscht.

Ist  $\dim_k V = \infty$ , so wenden wir dies an auf allen endlich dimensionalen Teilräume. Diese werden beliebig groß. Damit wird auch der jeweilige Bildraum beliebig groß. Insgesamt folgt  $\text{rg } f = \infty$  und die Gleichung ist ebenfalls erfüllt.  $\square$

**Beispiel.** Sei  $A$  eine  $m \times n$ -Matrix und das homogene Gleichungssystem sei eindeutig lösbar, d.h. einzige Lösung ist 0. Dann folgt  $\text{rg } A = n$  und daher  $m \geq n$ . Ist zusätzlich  $m = n$ , so ist  $\text{Im } F_A \subset k^n$  Untervektorraum der Dimension  $n$ , also  $\text{Im } F_A = k^n$ . Mit anderen Worten: Jedes inhomogene Gleichungssystem

$$Ax = b$$

ist eindeutig lösbar.

**Bemerkung.** Die Rechnung im Beweis zeigt: Ist  $f : V \rightarrow W$  ein Isomorphismus, so ist das Bild einer Basis wieder eine Basis. Insbesondere gilt  $\dim V = \dim W$ .

Wir wollen eine zweite Version der Dimensionsformel betrachten, die auch zu einem konzeptionelleren Beweis führt.

**Lemma 4.11.** *Seien  $V, W$  Vektorräume. Dann gilt*

$$\dim_k V \oplus W = \dim_k V + \dim_k W$$

*Beweis:* Sei  $v_i \in V$  für  $i \in I$  eine Basis. Sei  $w_j \in W$  für  $j \in J$  eine Basis. Wir erhalten eine Basis für  $V \oplus W$  mit dem Indexsystem  $I \cup J$  durch

$$v'_i = (v_i, 0) \text{ für alle } i \in I, w'_j = (0, w_j) \text{ für alle } j \in J$$

Die Aussage über die Dimension lesen wir ab.  $\square$

**Lemma 4.12.** *Sei  $V$  ein Vektorraum  $U_1, U_2$  Untervektorräume. Sei*

$$\phi : U_1 \oplus U_2 \rightarrow V \quad (u_1, u_2) \mapsto u_1 + u_2$$

*Dann ist  $\phi$  injektiv genau dann, wenn  $U_1 \cap U_2 = 0$ . Die Abbildung  $\phi$  ist surjektiv, genau dann, wenn*

$$U_1 + U_2 = \{u_1 + u_2 \in V \mid u_1 \in U_1, u_2 \in U_2\} = V$$

*Insbesondere ist sie ein Isomorphismus, falls beides erfüllt ist. In diesem Fall ist  $\dim_k U_1 + \dim_k U_2 = \dim_k V$ .*

*Beweis:* Wir berechnen:

$$\text{Ker } \phi = \{(u_1, u_2) | u_1 + u_2 = 0\} = \{(u_1, u_2) | u_1 = -u_2\}$$

Damit ist  $u_1 \in U_1 \cap U_2$ . Der Kern ist genau dann trivial, wenn  $U_1 \cap U_2 = 0$ . Die Aussage über das Bild ist klar.  $\square$

**Lemma 4.13.** *Sei  $U \subset V$  ein Untervektorraum.*

(i)  $\pi : V \rightarrow V/U$  mit  $\pi(v) = \bar{v}$  ist eine lineare Abbildung mit Kern  $U$ .

(ii) Es gibt eine lineare Abbildung

$$s : V/U \rightarrow V$$

mit  $\pi \circ s = \text{id}_{V/U}$ .

(iii)  $V \cong U \oplus \text{Im}(s)$ .

**Bemerkung.** Vom formalen Standpunkt ist dies die entscheidende Eigenschaft von Vektorräumen. Eigentlich sind alle Aussagen über Basen Folgerungen!

*Beweis:* Die Linearität von  $\pi$  ist nichts als die Definition von  $+$ ,  $\cdot$  in  $V/U$ . Der Kern von  $p$  ist

$$\text{Ker } \pi = \{v \in V | \bar{v} = 0\} = \{v \in V | v + U = 0 + U\} = \{v \in V | v - 0 \in U\}$$

Sei  $x_i \in V/U$  für  $i \in I$  eine Basis. Sei  $v_i \in V$  mit  $\pi(v_i) = x_i$ . Wir definieren  $s$  auf den Basiselementen durch  $s(x_i) = v_i$ . Nach Konstruktion gilt  $\pi \circ s(x_i) = x_i = \text{id}(x_i)$ . Da eine lineare Abbildung durch ihre Werte auf einer Basis eindeutig festgelegt ist, gilt  $\pi \circ s = \text{id}$ .

Schließlich überprüfen wir die Aussage zur direkten Summe, indem wir das Kriterium überprüfen. Sei  $v \in U \cap \text{Im}(s)$ . Dann ist  $v = s(x)$  für ein  $x \in V/U$ . Es folgt

$$x = \pi(s(x)) = \pi(v) = 0$$

da  $v \in \text{Ker}(\pi)$ . Dann ist aber auch  $v = s(0) = 0$ . Sei  $y \in V$  beliebig. Wir zerlegen

$$y = (y - s\pi(y)) + s(\pi(y)).$$

Der erste Summand liegt im Kern von  $\pi$ , der zweite im Bild von  $s$ .  $\square$

**Bemerkung.** Mit den Notationen des Beweises sei  $p = s \circ \pi : V \rightarrow V$ . Es gilt  $p^2 = (s\pi)(s\pi) = s(\pi s)\pi = s \text{id } \pi = p$ . Endomorphismen mit dieser Eigenschaft heißen *Projektor*. Sie spielen in der statistischen Physik eine große Rolle!

**Satz 4.14** (Homomorphiesatz). *Sei  $f : V \rightarrow W$  eine lineare Abbildung,  $U \subset \text{Ker } f$  ein Untervektorraum. Dann gibt es eine eindeutige Abbildung  $\bar{f} : V/U \rightarrow W$ , so dass  $f$  faktorisiert als*

$$f : V \xrightarrow{\pi} V/U \xrightarrow{\bar{f}} W$$

Es ist  $\text{Ker } \bar{f} = \text{Ker } f/U \subset V/U$  und  $\text{Im } \bar{f} = \text{Im } f$ . Speziell für  $U = \text{Ker } f$  induziert  $\bar{f}$  einen Isomorphismus

$$\bar{f} : V/\text{Ker } f \rightarrow \text{Im } f$$

*Beweis:* Sei  $\bar{v} = v + U$ . Wir setzen

$$\bar{f}(\bar{v}) = f(v)$$

Zu überprüfen ist als erstes die Wohldefiniertheit. Seien also  $y, y' \in V$  mit  $\bar{y} = \bar{y}' \Leftrightarrow y - y' \in U$ . Dann gilt

$$\bar{f}(\bar{y}) = \bar{f}(\bar{y}') = f(y) - f(y') = f(y - y') = 0$$

denn  $y - y' \in U \subset \text{Ker } f$ . Die Linearität von  $\bar{f}$  folgt aus der Linearität von  $f$ . Die Faktorisierung  $f = \bar{f}\pi$  gilt nach Definition. Offensichtlich ist die Wahl von  $\bar{f}$  die einzige mögliche. Da  $\pi$  surjektiv ist, folgt die Aussage über die Bilder rein mengentheoretisch. Ist  $\bar{y} \in \text{Ker } \bar{f}$ , also  $f(y) = \bar{f}(\bar{y}) = 0$ , so liegt  $y \in \text{Ker } f$ .  $\square$

*Beweis von Satz 4.10.* Sei  $f : V \rightarrow W$  linear. Wir wenden Lemma 4.13 an mit  $U = \text{Ker}(f)$ . Dann gilt nach dem Gezeigten

$$\text{Im}(f) \cong V/\text{Ker}(f) \cong \text{Im}(s)$$

und

$$V \cong \text{Ker}(f) \oplus \text{Im}(s).$$

Wir lesen die Dimensionen ab.  $\square$

## Kapitel 5

# Darstellende Matrizen

Zwei Vektorräume heißen *isomorph*, wenn es einen Isomorphismus zwischen ihnen gibt. Wir kennen eine *Invariante* von Vektorräumen, nämlich die Dimension. Damit ist gemeint: isomorphe Vektorräume haben dieselbe Dimension. Die Invariante ist auch scharf: Sie bestimmt den Vektorraum bis auf Isomorphie.

**Lemma 5.1.** *Sei  $n \in \mathbb{N}_0$  und  $V$  ein Vektorraum der Dimension  $n$ . Dann ist  $V$  isomorph zu  $k^n$ . Genauer: Sei  $B = (v_1, \dots, v_n)$  eine Basis von  $V$ , so ist*

$$\Phi_B : V \rightarrow k^n \quad \sum_{i=1}^n \lambda_i v_i \mapsto (\lambda_1, \dots, \lambda_n)$$

*ein Isomorphismus. Die Wahl einer Basis von  $V$  ist äquivalent zur Wahl eines Isomorphismus*

$$\Phi : V \rightarrow k^n$$

**Bemerkung.** Die Aussage gilt auch für unendlich-dimensionale Vektorräume, wenn man den Dimensionsbegriff in diesem Fall verfeinert. Es kommt auch die Mächtigkeit der Indexmenge einer Basis an. Zwei Mengen heißen *gleichmächtig*, wenn es eine Bijektion zwischen ihnen gibt.

*Beweis:* Die Bijektivität der Abbildung haben wir uns bereits überlegt (Bemerkung nach Satz 3.9). Die Linearität sieht man sofort für die Umkehrabbildung:

$$k^n \rightarrow V \quad (\lambda_1, \dots, \lambda_n) \mapsto \sum_{i=1}^n \lambda_i v_i$$

Damit definiert jede Basis einen Isomorphismus. Sei umgekehrt  $\Phi : V \rightarrow k^n$  ein Isomorphismus. Wir betrachten die Standardbasis von  $k^n$ . Dann ist  $\Phi^{-1}(e_i)$  für  $i = 1, \dots, n$  ein Basis  $B$  mit  $\Phi_B = \Phi$ .  $\square$

Um lineare Abbildungen zwischen beliebigen Vektorräumen zu verstehen, genügt es also lineare Abbildungen zwischen  $k^n$ s zu verstehen. Sei

$$F : k^n \rightarrow k^m$$

eine lineare Abbildung. Zur besseren Unterscheidung sei  $e_j$  für  $j = 1, \dots, n$  die Standardbasis von  $k^n$  und  $e'_i$  für  $i = 1, \dots, m$  die Standardbasis des  $k^m$ . Sei

$$F(e_j) = \sum_{i=1}^m a_{ij} e'_i$$

Wir betrachten die Matrix  $A = (a_{ij})$ . Sie hat  $m$  Zeilen und  $n$  Spalten. In der  $j$ -ten Spalte steht das Bild des  $j$ -ten Basisvektors. Sei nun  $x = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}$ . Dann gilt

$$F(x) = F\left(\sum_{j=1}^n x_j e_j\right) = \sum_{j=1}^n x_j F(e_j) = \sum_{j=1}^n \sum_{i=1}^m a_{ij} x_j e_i = \sum_{i=1}^m \sum_{j=1}^n a_{ij} x_j e_i = Ax$$

Damit haben wir gezeigt:

**Satz 5.2.** Sei  $F : k^n \rightarrow k^m$  eine lineare Abbildung. Dann gibt es eine eindeutige  $m \times n$ -Matrix  $A$  mit  $F = F_A$ .

$A$  heißt Matrix von  $F$ .

**Definition 5.3.** Seien  $V$  und  $W$  Vektorräume der Dimension  $n$ ,  $m$ . Seien  $A = (v_1, \dots, v_n)$  und  $B = (w_1, \dots, w_m)$  Basen von  $V$  und  $W$ . Für jede lineare Abbildung  $f : V \rightarrow W$  sei  $M_B^A(f)$  die Matrix der linearen Abbildung

$$F : k^n \xrightarrow{\Phi_A^{-1}} V \xrightarrow{f} W \xrightarrow{\Phi_B} k^m$$

$M_B^A(f)$  heißt darstellende Matrix von  $f$  bezüglich  $A$  und  $B$ .

Sei

$$f(v_j) = \sum_{i=1}^m a_{ij} w_i$$

Dann gilt also

$$F(e_j) = \Phi_B \circ f \circ \Phi_A^{-1}(e_j) = \Phi_B(f(v_j)) = \Phi_B\left(\sum_{i=1}^m a_{ij} w_i\right) = \sum_{i=1}^m a_{ij} e'_i$$

d.h. in den Spalten von der darstellenden Matrix stehen die Bilder der Basisvektoren in  $A$  ausgedrückt in der Basis  $B$ .

**Beispiel.** Sei  $V = \mathbb{Q}^2$  mit der Basis  $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ . Sei  $W = \mathbb{Q}^2$  mit der Basis  $w_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, w_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ . Sei  $f$  die Multiplikation mit  $\begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix}$ . Dann

gilt also

$$f(v_1) = \begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 6w_1 - 3w_2$$

$$f(v_2) = \begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} -3 \\ -1 \end{pmatrix} = -4w_1 + 3w_2$$

Die darstellende Matrix ist  $\begin{pmatrix} 6 & -4 \\ -3 & 3 \end{pmatrix}$ .

In diesem Beispiel sieht man, dass die darstellende Matrix von den Basen abhängt! Natürlich lassen sich darstellende Matrizen für verschieden Basen ineinander umrechnen. Übersichtlicher wird es in der Sprache der kommutativen Diagramme: Ein Viereck

$$\begin{array}{ccc} V_1 & \xrightarrow{g} & V_2 \\ f_1 \uparrow & & \uparrow f_2 \\ W_1 & \xrightarrow{g'} & W_2 \end{array}$$

heißt *kommutativ*, falls  $g \circ f_1 = f_2 \circ g'$ , d.h. die beiden möglichen Abbildungen  $W_1 \rightarrow V_2$  stimmen überein.

Die darstellende Matrix steht dann im kommutativen Viereck

$$\begin{array}{ccc} k^n & \xrightarrow{M_B^A(f)} & k^m \\ \Phi_A \uparrow & & \uparrow \Phi_B \\ V & \xrightarrow{f} & W \end{array}$$

**Satz 5.4.** Seien  $V, W, U$  Vektorräume der Dimensionen  $n, m, p$  mit Basen  $A, B, C$ . Seien  $f : V \rightarrow W$  und  $g : W \rightarrow U$  linear. Dann gilt

$$M_C^A(g \circ f) = M_C^B(g)M_B^A(f)$$

*Beweis:* Wir setzen die beiden kommutativen Diagramm zusammen:

$$\begin{array}{ccccc} k^n & \xrightarrow{M_B^A(f)} & k^m & \xrightarrow{M_C^B(g)} & k^p \\ \Phi_A \uparrow & & \uparrow \Phi_B & & \uparrow \Phi_C \\ V & \xrightarrow{f} & W & \xrightarrow{g} & U \end{array}$$

und lesen die Aussage ab.  $\square$

**Korollar 5.5.** Seien  $V, W$  Vektorräume der Dimension  $n$  mit Basen  $A, B$ . Sei  $f : V \rightarrow W$  ein Isomorphismus. Dann gilt

$$M_A^B(f^{-1})M_B^A(f) = E_n, \quad M_B^A(f)M_A^B(f^{-1}) = E_n$$

*Beweis:* Wir setzen im Satz  $U = V$ ,  $C = A$  und  $g = f^{-1}$ .

Es ist also  $M_A^B(f^{-1})M_B^A(f) = M_A^A(\text{id})$  und dies ist die Einheitsmatrix. Ebenso behandelt man die zweite Komposition.  $\square$

**Definition 5.6.** Eine  $n \times n$ -Matrix  $M$  heißt invertierbar, wenn es eine  $n \times n$ -Matrix  $M'$  gibt mit

$$MM' = M'M = E_n$$

Wir schreiben  $M' = M^{-1}$ .

Wir haben also gerade gezeigt, dass  $M_A^B(f^{-1}) = M_B^A(f)^{-1}$ .

**Bemerkung.** Im ersten Kapitel haben wir gezeigt, dass für  $n = 2$  eine Matrix  $M$  genau dann invertierbar ist, wenn  $\det(M) \neq 0$ . Dazu später mehr.

Ein wichtiger Spezialfall tritt für  $f = \text{id}$  ein. Seien also  $A$  und  $B$  Basen von  $V$ . Dann ist  $M_B^A(\text{id})$  die *Basiswechselformel* aus Definition 3.17.

$$\begin{array}{ccc} k^n & \xrightarrow{M_B^A(\text{id})} & k^n \\ \Phi_A \uparrow & & \uparrow \Phi_B \\ V & \xrightarrow{\text{id}} & V \end{array}$$

Basiswechselformeln sind invertierbar, da sie einen Isomorphismus darstellen.

**Satz 5.7** (Basiswechselformel). Seien  $A$  und  $A'$  Basen des  $n$ -dimensionalen Vektorraums  $V$ ,  $B$  und  $B'$  Basen des  $m$ -dimensionalen Vektorraums  $W$  und  $f : V \rightarrow W$  eine lineare Abbildung. Dann gilt

$$M_{B'}^{A'}(f) = (M_{B'}^{B'}(\text{id}_W))^{-1} M_B^A(f) M_A^{A'}(\text{id}_V)$$

*Beweis:* Wir wenden Satz 5.4 an auf die Komposition  $\text{id}_W \circ f \circ \text{id}_V$ .  $\square$

**Definition 5.8.** Zwei  $m \times n$ -Matrizen  $M$  und  $M'$  heißen äquivalent, wenn es invertierbare Matrizen  $S \in M_m(k)$  und  $T \in M_n(k)$  gibt mit

$$M' = S^{-1}MT$$

**Bemerkung.** Die Basiswechselformel besagt, dass zwei darstellende Matrizen desselben Morphismus  $f : V \rightarrow W$  äquivalent sind.

**Lemma 5.9.** Zwei  $m \times n$ -Matrizen sind genau dann äquivalent, wenn sie denselben Rang haben. Jede Matrix  $M$  ist äquivalent zu einer Matrix der Normalform

$$N_r = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

wobei  $r = \text{rg}M$ .



*Beweis:* Wir fassen die Matrizen als lineare Abbildungen  $k^n \rightarrow k^m$  auf. Ist  $M' = S^{-1}MT$ , so haben wir es mit einem kommutativen Diagramm

$$\begin{array}{ccc} k^n & \xrightarrow{F_M} & k^m \\ F_T \uparrow & & \uparrow F_S \\ k^n & \xrightarrow{F_{M'}} & k^m \end{array}$$

zu tun, in dem die senkrechten Abbildungen Isomorphismen sind. Also gibt  $\text{rg}M = \text{rg}F_M = \text{rg}F_{M'} = \text{rg}M'$ .

Nun zeigen wir die Normalformenaussage. Wir suchen geeignete Basen des  $k^n$  und  $k^m$ , so dass die darstellende Matrix möglichst einfach wird. Genau dies haben wir im zweiten Beweis der Invarianz der Dimension (Korollar 3.16) bereits durchgeführt. Einfache Basiswechseloperationen führten aus elementare Zeilen- und Spaltenoperationen der Matrix. Nach dem Gauß-Algorithmus wird die Normalform aus Satz 1.16 erreicht.

Oder alternativ: Wir wählen die Basen wie im ersten Beweis der Dimensionsformel. Eine Basis  $v_1, \dots, v_s$  von  $\text{Ker} F_M$  wird zu einer Basis  $v_1, \dots, v_n$  von  $V$  ergänzt. Dann ist  $Mv_{s+1}, \dots, Mv_n$  eine Basis von  $\text{Im} F_M$ . Wir setzen  $w_1 = Mv_{s+1}, \dots, w_r = Mv_n$  (also  $r = n - s$ ) und ergänzen zu einer Basis von  $W$ . Die darstellende Matrix von  $F_M$  in dieser Basis hat die gesuchte Form. Sie ist nach der vorherigen Bemerkung äquivalent zu  $M$ .

Haben zwei Matrizen  $M$  und  $M'$  denselben Rang  $r$ , so sind sie beide äquivalent zur selben Matrix in Normalform  $N_r$ . Es gibt also invertierbare Matrizen  $S, S' \in M_n(k)$ ,  $T, T' \in M_m(k)$  mit

$$N_r = S^{-1}MT = (S')^{-1}M'T' \Rightarrow M = S(S')^{-1}M'T'T^{-1}$$

Produkte von invertierbaren Matrizen sind invertierbar, also sind  $M$  und  $M'$  äquivalent.  $\square$

## Dualräume

**Definition 5.10.** Sei  $V$  ein Vektorraum. Dann heißt

$$V^* = \text{Hom}_k(V, k)$$

Dualraum von  $V$ .

**Beispiel.** Sei  $V = k^n$ . Dann gilt mit Satz 5.2

$$(k^n)^* = \text{Hom}_k(k^n, k) = M_{1 \times n}(k)$$

d.h. dies ist der Vektorraum der Zeilenvektoren der Länge  $n$ . Er hat die Dimension  $n$ , ist also isomorph zu  $k^n$ , dem Vektorraum der Spaltenvektoren. Naheliegender ist der Isomorphismus, der den Zeilenvektor  $e_i^* = (0, \dots, 1, 0, \dots, 0)$  auf  $e_i$  abbildet. Oder als Abbildungen formuliert:

$$e_i^*(e_j) = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

**Definition 5.11.** Sei  $f : V \rightarrow W$  eine lineare Abbildung. Dann heißt  $f^* : W^* \rightarrow V^*$  mit  $f^*(x) = x \circ f$  duale Abbildung von  $f$ .

**Beispiel.** Sei  $A$  eine  $m \times n$ -Matrix. Wir betrachten  $F_A : k^n \rightarrow k^m$ . Dann ist  $(F_A)^* : (k^m)^* \rightarrow (k^n)^*$  eine Abbildung auf Zeilenvektoren. Sei also  $v = (v_1, \dots, v_m)$  ein Zeilenvektor. Matrixmultiplikation mit Elementen von  $k^m$  macht ihn zu einer linearen Abbildung  $x \mapsto vx$ . Nach Definition ist  $(F_A)^*(v) = v \circ F_A$ . D.h. für jedes  $y \in k^n$  gilt

$$(F_A)^*(v)(y) = v \circ F_A(y) = v(Ay) = vAy$$

D.h.  $(F_A)^*(v) = vA$ . Damit ist  $(F_A)^*$  die *Rechtsmultiplikation* mit  $A$ .

**Lemma 5.12** (Funktorialität). Seien  $f : V \rightarrow W$  und  $g : W \rightarrow U$  lineare Abbildungen.

(i)  $f^*$  ist linear.

(ii) Es gilt  $(g \circ f)^* = f^* \circ g^*$ .

(iii) Es gilt  $(\text{id}_V)^* = \text{id}_{V^*}$ .

*Beweis:* (i) Seien  $x, y \in W^*$ ,  $a, b \in k$ . Wir betrachten  $f^*(ax + by) = (ax + by) \circ f$  und  $af^*(x) + bf^*(y) = ax \circ f + by \circ f$ . Zu überprüfen ist die Gleichheit dieser beiden Elemente von  $V^* = \text{Hom}_k(V, k)$ . Eine Gleichheit von Abbildungen überprüft man, indem man jedes  $v \in V$  einsetzt. Für  $v \in V$  gilt

$$\begin{aligned} (ax + by) \circ f(v) &= (ax + by)(f(v)) = ax(f(v)) + by(f(v)) = \\ (ax \circ f + by \circ f)(v) &= ax \circ f(v) + by \circ f(v) = ax(f(v)) + by(f(v)) \end{aligned}$$

(ii) Sei  $x \in U^*$ . Dann ist nach Definition

$$(g \circ f)^*(x) = x \circ (g \circ f)$$

Andererseits gilt

$$(f^* \circ g^*)(x) = f^*(g^*(x)) = f^*(x \circ g) = (x \circ g) \circ f$$

(iii) Klar. □

**Korollar 5.13.** Sei  $f : V \rightarrow W$  ein Isomorphismus. Dann ist  $f^* : W^* \rightarrow V^*$  ein Isomorphismus.

*Beweis:* Da  $f$  ein Isomorphismus ist, ist  $g = f^{-1} : W \rightarrow V$  linear mit  $f \circ g = \text{id}_W$  und  $g \circ f = \text{id}_V$ . Anwenden des Lemmas ergibt also

$$g^* \circ f^* = (f \circ g)^* = (\text{id}_W)^* = \text{id}_{W^*}, f^* \circ g^* = \text{id}_{V^*}$$

Damit ist  $f^*$  ein Isomorphismus. □

Ist insbesondere  $V$  ein Vektorraum der Dimension  $n$  mit Basis  $A$ , so können wir dies anwenden auf den Isomorphismus  $\Phi_A : V \rightarrow k^n$ .

**Definition 5.14.** Sei  $V$  ein Vektorraum mit Basis  $v_1, \dots, v_n$ . Dann heißt  $v_1^*, \dots, v_n^* \in V^*$  mit

$$v_i^*(v_j) = \delta_{ij}$$

duale Basis von  $V$ .

**Lemma 5.15.** Die duale Basis ist tatsächlich eine Basis von  $V^*$ .

*Beweis:* Sei  $f : V \rightarrow k$  ein Element von  $V^*$ . Sei  $a_i = f(v_i)$  für  $i = 1, \dots, n$ .

**Behauptung.**  $\sum_{i=1}^n a_i v_i^* = f$

Es genügt Gleichheit für alle  $v_j$  zu zeigen, da eine lineare Abbildung durch die Werte auf den Basisvektoren bestimmt ist. Es gilt

$$\left(\sum_{i=1}^n a_i v_i^*\right)(v_j) = \sum_{i=1}^n a_i v_i^*(v_j) = \sum_{i=1}^n a_i \delta_{ij} = a_j = f(v_j)$$

Damit bilden die  $v_i^*$  ein Erzeugendensystem von  $V^*$ .

Seien nun  $\lambda_1, \dots, \lambda_n \in k$  mit  $\sum_{i=1}^n \lambda_i v_i^* = 0$ . Wir setzen  $v_j$  ein und erhalten wie eben gesehen  $\lambda_j = 0$ . Dies gilt für alle  $j$ , die  $v_j^*$  sind linear unabhängig.  $\square$

Wir lesen insbesondere ab:

**Korollar 5.16.** Ist  $V$  ein Vektorraum der Dimension  $n < \infty$ , so gilt  $\dim_k V^* = n$ .

**Bemerkung.** Diese Aussage ist falsch für unendlich dimensionale Vektorräume!

Wir wollen nun die darstellende Matrix der dualen Abbildung in der dualen Basis bestimmen. Sei also  $f : V \rightarrow W$  eine lineare Abbildung,  $A = (v_1, \dots, v_n)$  und  $B = (w_1, \dots, w_m)$  seien Basen von  $V$  und  $W$ . Dann hat  $f$  die darstellende Matrix  $M_B^A(f) = (a_{ij})$ . Es gilt also

$$f(v_j) = \sum_{i=1}^m a_{ij} w_i$$

Wie sieht die darstellende Matrix  $M_{A^*}^{B^*}(f^*)$  für  $f^* : W^* \rightarrow V^*$  aus? Wir rechnen es nach. In den Spalten stehen die Bilder der Basisvektoren  $w_j^*$  ausgedrückt in der Basis  $v_i^*$ .

$$f^*(w_j^*) = \sum_{i=1}^n b_{ij} v_i^*$$

Um die  $b_{ij}$  zu bestimmen, setzen wir  $v_s$  ein und erhalten

$$\begin{aligned} f^*(w_j^*)(v_s) &= (w_j^* \circ f)(v_s) = w_j^*(f(v_s)) = w_j^*\left(\sum_{i=1}^m a_{is}w_i\right) \\ &= \sum_{i=1}^m a_{is}w_j^*(w_i) = \sum_{i=1}^m a_{is}\delta_{ij} = a_{js} \end{aligned}$$

Andererseits ist

$$\sum_{i=1}^n b_{ij}v_i^*(v_s) = \sum_{i=1}^n b_{ij}\delta_{is} = b_{sj}$$

D.h.  $a_{js} = b_{sj}$ .

**Definition 5.17.** Sei  $A = (a_{ij})_{i=1, j=1}^{m, n}$  eine  $m \times n$ -Matrix. Dann heißt  $A^t = (b_{st})_{s=1, t=1}^{n, m}$  mit  $b_{st} = a_{ts}$  die transponierte Matrix zu  $A$ .

Die Einträge werden also an der Diagonale gespiegelt. Wir haben also gezeigt:

**Satz 5.18.** Seien  $V, W$  endlichdimensionale Vektorräume mit Basen  $A$  und  $B$ . Sei  $f : V \rightarrow W$  linear. Dann gilt für die darstellenden Matrizen

$$(M_B^A(f))^t = M_{A^*}^{B^*}(f^*)$$

Schließlich vergleichen wir die Ränge von  $f$  und  $f^*$ .

**Korollar 5.19.** Sei  $f : V \rightarrow W$  lineare Abbildung von endlich-dimensionalen Vektorräumen. Dann gilt  $\text{rg}f = \text{rg}f^*$ .

*Beweis:* Wir lesen die Ränge an darstellenden Matrizen ab. Dabei wählen wir die Basis von  $V$  und  $W$  so, dass  $M_B^A(f)$  in Normalform ist. Dann ist automatisch  $M_{A^*}^{B^*}(f^*)$  in Normalform und hat denselben Rang.  $\square$

**Bemerkung.** In Matrizen formuliert bedeutet dies: Die maximale Zahl von linear unabhängigen Spalten von  $A$  und  $A^t$  stimmt überein. Bzw. Die maximale Zahl von linear unabhängigen Spalten von  $A$  ist gleich der maximalen Zahl von linear unabhängigen Zeilen von  $A$ . Die letzte Zahl heißt *Zeilenrang*. Es gilt also *Zeilenrang=Spaltenrang*.

Um dies zu zeigen, hätten wir keine Dualräume gebraucht. Es genügt zu beobachten, dass alle elementaren Zeilen- und Spaltentransformationen den Spaltenrang unverändert lassen. Aus Symmetriegründen gilt dies dann auch für den Zeilenrang. Sobald die Matrix in Normalform ist, lesen wir die Gleichheit ab.

## Kapitel 6

# Determinanten

Wir erinnern uns, dass eine  $2 \times 2$ -Matrix genau dann invertierbar ist, wenn ihre Determinante ungleich 0 ist. Dies wollen wir verallgemeinern auf beliebiges  $n$ . Wir identifizieren im Folgenden oft quadratische Matrizen in  $A \in M_n(k)$  mit  $n$ -Tupeln von Spaltenvektoren  $A = (v_1, \dots, v_n)$  für  $v_i \in V = k^n$ . Für  $A = (a_{ij})_{i,j=1}^n$  und  $1 \leq i_0, j_0 \leq n$  definieren wir die *Streichmatrix*  $A_{i_0 j_0}$ , in der wir aus  $A$  die Zeile  $i_0$  und die Spalte  $j_0$  streichen.

**Definition 6.1.** Sei  $k$  ein Körper,  $n \in \mathbb{N}$ . Die Determinante

$$\det = \det_n : M_n(k) \rightarrow k$$

ist definiert als

$$\det_1 = \text{id}$$

und

$$A = (a_{ij})_{i,j=1}^n \mapsto \sum_{i=1}^n (-1)^{i+1} a_{i1} \det_{n-1} A_{i1}.$$

**Beispiel.** Für  $n = 2$  und  $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$  erhalten wir die Streichmatrizen  $A_{11} = (a_{22})$ ,  $A_{12} = (a_{21})$  und damit

$$\det_2(A) = a_{11}a_{22} - a_{12}a_{21}.$$

Dies ist die Formel aus Kapitel 1.

**Satz 6.2.** (i) Die Abbildung  $\det_n$  ist multi-linear, d.h. für jedes  $1 \leq j \leq n$  und  $v_1, \dots, v_n, v'_j \in k^n$ ,  $a, b \in k$  gilt

$$\det(v_1 \ \dots \ v_{j-1} \ av_j + bv'_j \ v_{j+1} \ \dots \ v_n) = \\ a \det(v_1 \ \dots \ v_j \ \dots \ v_n) + b \det(v_1 \ \dots \ v_{j-1} \ v'_j \ v_{j+1} \ \dots \ v_n)$$

(ii) Die Abbildung  $\det_n$  ist alternierend, d.h. ist  $M \in M_n(K)$  eine Matrix, in der zwei Spalten übereinstimmen, d.h.  $v_1, \dots, v_n \in k^n$  mit  $v_i = v_j$  für ein Paar  $i \neq j$  von Indizes gilt

$$\det(v_1 \ \dots \ v_n) = 0$$

**Bemerkung.** Für alternierende, multilineare Abbildungen folgt

$$d(\dots, v, \dots, w, \dots) = -d(\dots, w, \dots, v, \dots),$$

beim Vertauschen von zwei Spalten ändert sich das Vorzeichen. Es ist nämlich:

$$\begin{aligned} 0 &= d(\dots, v+w, \dots, v+w, \dots) \\ &= d(\dots, v, \dots, v, \dots) + d(\dots, v, \dots, w, \dots) + d(\dots, w, \dots, v, \dots) + d(\dots, w, \dots, w, \dots) \\ &\quad d(\dots, v, \dots, w, \dots) + d(\dots, w, \dots, v, \dots) + \end{aligned}$$

*Beweis:* Wir beginnen mit der Multilinearität. Wir argumentieren mit vollständiger Induktion. Die Aussage ist wahr für  $n = 1$ . Angenommen, sie ist wahr für Matrizen der Größe  $n - 1$ ,  $n \geq 2$ . Sei

$$\begin{aligned} A &= (v_1, \dots, v_n) = (a_{ij}) \\ A' &= (v_1, \dots, v'_j, \dots, v_n) = (a'_{ij}) \\ B &= (v_1, \dots, av_j + bv'_j, \dots, v_n) \end{aligned}$$

Wir behandeln zunächst  $j = 1$ . Dann stimmen die Streichmatrizen von  $A$ ,  $A'$  und  $B$  überein. Die Linearität folgt direkt aus der Definition.

Sei nun  $j > 1$ . Die Streichmatrizen  $A_{1i}$ ,  $A'_{1i}$  und  $B_{i1}$  stehen in der selben Relation wie  $A$ ,  $A'$  und  $B$ , während alle drei dieselbe erste Spalte haben. Nach Induktionsvoraussetzung gilt

$$\det_{n-1}(B_{i1}) = a \det_{n-1} A_{i1} + b \det_{n-1} A'_{i1}.$$

Durch Aufsummieren erhalten wir die Behauptung.

Wir wenden uns nun der zweiten Behauptung zu, wieder mit vollständiger Induktion. Für  $n = 1$  ist nichts zu zeigen. Für  $n = 2$  gilt die Behauptung direkt aus der Formel. Sei nun  $n \geq 2$ . Seien  $j < j'$ , so dass die Spalten  $j$  und  $j'$  übereinstimmen. Ist  $j > 1$ , so hat jede Streichmatrix zwei gleiche Spalten, also Determinante 0 nach Induktionsvoraussetzung. Sei also jetzt  $j = 1$ . Durch Vertauschen von Spalten (bereits erlaubt für die Spalten ab 2) ändert sich das Vorzeichen der Determinante. Ohne Beschränkung der Allgemeinheit reicht es,  $j' = 2$  zu betrachten.

Wir benutzen nun die ersten zwei Schritte der rekursiven Definition der Determinante von  $A = (a_{ij})$ . Wir erhalten also eine Doppelsumme

$$\sum_{i_1 \neq i_2} \pm a_{i_1 1} a_{i_2 2} \det(A^{i_1, i_2})$$

wobei  $A^{i_1, i_2}$  die Streichmatrix ohne die Spalten 1, 2 und die Zeilen  $i_1, i_2$  ist. Nach Voraussetzung ist  $a_{i_1} = a_{i_2}$ , daher fassen wir die Summanden  $i_1 = s, i_2 = t$  und  $i_1 = t, i_2 = s$  zusammen. Ohne Einschränkung ist  $s < t$ . Wir erhalten beide Male den Summanden  $\pm a_s a_t \det(A^{st})$ . Nun wenden wir uns dem Vorzeichen zu. Beim Streichen von  $i_1 = s$  erhalten wir das Vorzeichen  $(-1)^{s+1}$ . Dabei rutscht  $i_2 = t$  in die Zeile  $t-1$  und trägt dann das Vorzeichen  $(-1)^t$  bei. Insgesamt also  $(-1)^{s+t+1}$ . Beim Streichen von  $i_1 = t$  erhalten wir das Vorzeichen  $(-1)^{t+1}$ . Wegen  $s < t$  bleibt  $s$  in der Zeile  $s$ . Im zweiten Schritt erhalten wir das Vorzeichen  $(-1)^{s+1}$ , zusammen also  $(-1)^{s+t+2}$ . Da die Vorzeichen verschieden sind, heben sich die Summanden weg. Insgesamt ist die Summe 0.  $\square$

Eine multilineare Abbildung ist eindeutig bestimmt durch die Werte auf den Basisvektoren, also auf Matrizen der Form

$$E_\sigma = (e_{\sigma(1)}, \dots, e_{\sigma(n)})$$

für eine Abbildung  $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ .

**Beispiel.** Für  $\sigma = \text{id}$ , also die Einheitsmatrix erhalten wir rekursiv

$$\det(E_n) = 1.$$

**Lemma 6.3.** Wenn es  $j \neq j'$  gibt mit  $\sigma(j) = \sigma(j')$ , dann folgt

$$\det(E_\sigma) = 0.$$

Wenn  $\sigma$  injektiv (und damit bijektiv) ist, so ist  $\det(E_\sigma) \in \{\pm 1\}$  unabhängig vom Körper  $k$ .

*Beweis:* Die erste Aussage ist ein Spezialfall des Satzes: det ist alternierend. Im zweiten Fall sortieren wir die Spalten der Matrix durch Vertauschen. Dies trägt jedesmal einen Vorzeichenwechsel bei. Sobald wir bei der Einheitsmatrix angekommen sind, ist der Wert 1.  $\square$

Wir erinnern uns: Die Menge der bijektiven Abbildungen von  $\{1, \dots, n\}$  nach  $\{1, \dots, n\}$  heißt *symmetrische Gruppe*  $S_n$ . Solche Abbildungen heißen *Permutationen*. Eine Abbildung die genau zwei Elemente vertauscht, heißt *Transposition*.

**Lemma 6.4.** Seien  $\sigma, \sigma' \in S_n$ . Dann ist

$$E_\sigma E_{\sigma'} = E_{\sigma\sigma'}.$$

*Beweis:* Wir überprüfen auf Basisvektoren. Sei  $\sigma : e_i \mapsto e_{\sigma(i)}$ ,  $\sigma' : e_j \mapsto e_{\sigma'(j)}$ , also

$$e_i \mapsto e_{\sigma(i)} \mapsto e_{\sigma'(\sigma(i))}.$$

$\square$

**Definition 6.5.** Sei  $\sigma \in S_n$ . Wir definieren

$$\text{sgn}(\sigma) = \det E_\sigma.$$

**Bemerkung.** Das Vorzeichen von  $\sigma$  zählt also, ob wir eine gerade oder ungerade Anzahl von Vertauschungen von benachbarten Einträgen benötigen, um das Tupel  $(\sigma(1), \dots, \sigma(n))$  zu sortieren. Wir haben nebenbei gezeigt, dass das Vorzeichen wohldefiniert ist.

Nützlich zu wissen:

**Satz 6.6** (Leibniz-Formel). *Sei  $A \in M_n(k)$ . Dann gilt*

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)}$$

*Beweis:* Multilinearität von  $\det$  plus Definition von  $\operatorname{sgn}(\sigma)$ . □

**Satz 6.7** (Eindeutigkeit von  $\det$ ). *Sei  $d : M_n(k) \rightarrow k$  multilinear und alternierend. Dann gilt*

$$d(A) = d(E_n) \det(A).$$

*Die Determinante ist die eindeutige multilineare, alternierende Abbildung mit  $\det(E_n) = 1$ .*

*Beweis:* Wegen der Multilinearität gilt

$$d(A) = \sum_{\sigma} d(E_{\sigma}) a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)}$$

wobei  $\sigma$  alle Abbildungen  $\{1, \dots, n\} \rightarrow \{1, \dots, n\}$  durchläuft. Mit dem Argument aus Lemma 6.3 ist  $d(E_{\sigma}) = 0$ , wenn  $\sigma$  nicht injektiv ist und  $d(E_{\sigma}) = \operatorname{sgn}(\sigma) d(E_n)$ , wenn doch. Dies stimmt bis auf den Faktor  $d(E_n)$  mit der Leibniz-Formel überein. □

**Theorem 6.8** (Produktformel). *Seien  $A, B \in M_n(k)$ . Dann gilt*

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

*Beweis:* Wir halten  $A$  fest. Beide Seiten definieren multilineare, alternierende Abbildungen  $M_n(k) \rightarrow k$  mit Wert  $\det(A)$  auf  $B = E_n$ . Man beachte nämlich, dass  $A(v_1, \dots, v_n) = (Av_1, \dots, Av_n)$ . Nach dem Eindeutigkeitsatz stimmen die beiden Seiten überein. □

**Korollar 6.9.** *Ist  $A$  invertierbar, so ist  $\det(A) \in k^* = k \setminus \{0\}$ .*

*Beweis:* Sei  $AB = E_n$ , so folgt  $\det(A) \det(B) = \det(E_n) = 1$ , also  $\det(A)$  invertierbar. □

Die Umkehrung gilt ebenfalls. Tatsächlich können wir die inverse Matrix angeben. Wir holen noch etwas aus.

**Korollar 6.10.** *Sei  $A \in M_n(k)$ . Dann ist  $\det A = \det A^t$ .*



*Beweis:* Die Leibniz-Formel für  $\det(A^t)$  liefert

$$\det(A^t) = \sum_{\sigma \in S_n} a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}.$$

Wir vertauschen die Faktoren, so dass der Faktor mit  $\sigma(1) = 1$  an der ersten Stelle steht, der mit  $\sigma(2) = 2$  an der zweiten. Wir erhalten den Summanden in der Leibniz-Formel für  $\det(A)$  zur Permutation  $\sigma^{-1}$ . Zu zeigen bleibt noch  $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$ . Die gilt wegen

$$\operatorname{sgn}(\sigma) \operatorname{sgn}(\sigma^{-1}) = \det(E_\sigma) \det(E_{\sigma'}) = \det(E_n) = 1.$$

□

**Bemerkung.** Damit gelten alle Eigenschaften der Determinante, die wir für Spalten formuliert haben, auch für die Zeilen der Matrizen.

**Satz 6.11** (Entwicklungsformel). *Sei  $A = (a_{ij})$  eine  $n \times n$ -Matrix. Für jedes Paar  $(i, j)$  sei  $A_{ij}$  die  $(n-1) \times (n-1)$ -Matrix, die aus  $A$  durch Streichen der  $i$ -ten Zeile und  $j$ -ten Spalte entsteht. Sei  $1 \leq k \leq n$  ein Index. Dann gilt*

$$\det A = \sum_{j=1}^n (-1)^{k+j} a_{kj} \det A_{kj}$$

(Entwicklung nach der  $k$ -ten Zeile) und

$$\det A = \sum_{i=1}^n (-1)^{k+i} a_{ik} \det A_{ik}$$

(Entwicklung nach der  $k$ -ten Spalte)

*Beweis:* Die Entwicklung nach der ersten Spalte war unsere Definition. Die Formel für die Entwicklung nach der  $k$ -Spalte entsteht durch Vertauschen der Spalten 1 und  $k$  und Entwicklung nach der neuen ersten Spalten.

Die Zeilenentwicklungsformeln folgen durch Betrachten der transponierten Matrix. □

**Satz 6.12** (Cramersche Regel). *Sei  $A = (a_{ij})_{i,j} \in M_n(k)$  mit  $\det A \neq 0$ . Dann gilt*

$$A^{-1} = \frac{1}{\det A} \left( (-1)^{i+j} \det A_{ji} \right)_{i,j}$$

wobei  $A_{ij}$  die Streichmatrix ist, die aus  $A$  durch Weglassen der Zeile  $i$  und der Spalte  $j$  entsteht.

*Beweis:* Nachrechnen! Sei  $B = (b_{ij})$  das  $\det A$ -fache der obigen Matrix, also  $b_{ij} = (-1)^{i+j} \det A_{ji}$ . Wir betrachten die Einträge

$$(c_{ik})_{i,k} = AB = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots \\ a_{21} & a_{22} & a_{23} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix} \begin{pmatrix} \det A_{11} & -\det A_{21} & \det A_{31} & \cdots \\ -\det A_{12} & \det A_{22} & -\det A_{32} & \cdots \\ \det A_{13} & -\det A_{23} & \det A_{33} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

Es gilt

$$c_{11} = a_{11}\det A_{11} - a_{12}\det A_{12} + a_{13}\det A_{13} \pm \dots = \det A$$

nach der Entwicklungsformel für die erste Zeile. Weiter gilt

$$c_{12} = -a_{11}\det A_{21} + a_{12}\det A_{22} - a_{13}\det A_{23} \pm \dots$$

Dies kann aufgefasst werden als Entwicklung nach der zweiten Zeile einer Matrix, deren Zeilen mit denen von  $A$  übereinstimmen, aber deren zweite Zeile  $(a_{11}, a_{12}, a_{13}, \dots)$  lautet. In dieser Matrix stimmen die ersten beiden Zeilen überein, also hat sie Determinante 0. Dasselbe Argument funktioniert für alle übrigen Einträge. Wir erhalten

$$AB = \det A E_n$$

Dies beweist die Behauptung. □

**Beispiel.** Sei  $n = 2$ ,  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ . Dann ist

$$A_{11} = (d), A_{12} = (c), A_{21} = (b), A_{22} = (a)$$

Die Formel besagt also

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

**Bemerkung.** Für praktische Zwecke ist die Formel meist nutzlos. Es müssen  $n^2$  Determinante von Streichmatrizen berechnet werden, das ist viel aufwändiger als die direkte Berechnung der inversen Matrix.

**Bemerkung.** Wir haben nirgends benutzt, dass wir über einem Körper arbeiten. Ist  $R$  ein kommutativer Ring mit 1, so ist  $A \in M_n(R)$  genau dann invertierbar, wenn  $\det(A) \in R$  ein multiplikatives Inverses hat. Für  $R = \mathbb{Z}$  erhalten wir also die Bedingung  $\det(A) = \pm 1$ .

## Berechnung von Determinanten

Wie berechnet man Determinanten in der Praxis?

- Für  $n = 2$  mit der expliziten Formel.
- Für  $n = 3$  ebenfalls mit der expliziten Formel, meist formuliert als Sarrusregel:

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

Über alle Diagonalen der Länge drei ist zu Multiplizieren und dann mit Vorzeichen zu summieren. Dabei erhalten die die Parallelen der Hauptdiagonalen  $a_{11}a_{22}a_{33}$  das Vorzeichen  $+$ , die anderen das Vorzeichen  $-$ .

- Für allgemeines  $n$ : *Nicht* mit der Leibnizformel. Für eine  $n \times n$ -Matrix sind hier nämlich  $n!$  Summanden mit jeweils  $n$  Faktoren zu betrachten. Dies ist sehr schnell sehr kompliziert, wie ein Selbstversuch mit  $n = 5$  zeigt. Auch für Computer ist es viel zu aufwändig.
- Bei dünn besetzten Matrizen ist Zeilen- oder Spaltenentwicklung oft hilfreich.

$$\begin{vmatrix} 1 & 0 & 0 & 3 \\ 2 & 2 & 1 & 0 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 1 & 1 \end{vmatrix} = (-1)^{3+4} 4 \begin{vmatrix} 1 & 0 & 0 \\ 2 & 2 & 1 \\ 0 & 0 & 1 \end{vmatrix} = -4 \begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix} = -8$$

(Entwicklung nach der dritten Zeile, dann nach der ersten)

- Am schnellsten geht es stets mit dem Gauß-Algorithmus. Dieser bringt sehr schnell die Matrix in Diagonalform - und dann lesen wir die Determinante ab.

$$\begin{vmatrix} 1 & 2 & 3 & 2 \\ 2 & 3 & 0 & 1 \\ 0 & 1 & 1 & 5 \\ 4 & 2 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 & 2 \\ 0 & -1 & -6 & -3 \\ 0 & 1 & 1 & 5 \\ 0 & -6 & -11 & -8 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 & 2 \\ 0 & -1 & -6 & -3 \\ 0 & 0 & -5 & 2 \\ 0 & 0 & 25 & 10 \end{vmatrix} = 1(-1)(-5 \cdot 10 - 2 \cdot 25) = 100$$

Hier haben wir verwendet:

**Lemma 6.13.** Seien  $A, B \in M_n(k)$ ,  $\lambda \in k$ . Entsteht  $B$  aus  $A$  durch Addieren des  $\lambda$ -Fachen einer Zeile oder Spalte zu einer anderen, so ist

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

*Beweis:* Sei  $A = (v_1, \dots, v_n)$ ,  $B = (v'_1, \dots, v'_n)$  mit  $v_i = v'_i$  für  $i \neq s$  und  $v'_s = v_s + \lambda v_t$  für  $t \neq s$ . Dann ist

$$\begin{aligned} \det(B) &= \det(A) + \det(v_1, \dots, \lambda v_t, \dots, v_t) \\ &= \det(A) + \lambda(\det(v_1, \dots, v_t, \dots, v_t, \dots)) = \det(A). \end{aligned}$$

Das Argument für die Zeilen geht genauso. □

Wir verstehen also, wie sich Determinanten unter beliebigen elementaren Zeilen- und Spaltentransformationen verändern.

## Berechnen der inversen Matrix

Der theoretische Nutzen der Determinante ist so viel höher als der praktische. Ob die Matrix invertierbar ist, stellt man am schnellsten fest, in dem man versucht, ihr Inverses auszurechnen. Theoretische Anwendungen werden wir in der Eigenwerttheorie kennenlernen.

Das Berechnen der inversen Matrix ist ein lineares Gleichungssystem

$$AX = E_n$$

aus  $n^2$  Gleichungen für die  $n^2$  Einträge von  $X$ . Tatsächlich zerfällt es in  $n$  lineare Gleichungssysteme

$$A \begin{pmatrix} x_{1j} \\ x_{2j} \\ \dots \end{pmatrix} = e_j$$

bestehend aus  $n$  Gleichungen mit  $n$  Unbekannten. Da es jedesmal dieselbe linke Seite ist, lassen wir den Gauß-Algorithmus für alle Gleichungen parallel ablaufen. D.h. wir versuchen  $A$  in Diagonalgestalt zu bringen und bestimmen dadurch die Einträge von  $X$ . Statt den Algorithmus zu formulieren, rechnen wir ein Beispiel mit  $k = \mathbb{Q}$ .

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -1 \\ -1 & 1 & 0 \end{pmatrix}$$

Dazu ist die erweiterte Matrix

$$\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 0 & 1 \end{array}$$

Addition der ersten Zeile zur dritten:

$$\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 3 & 3 & 1 & 0 & 1 \end{array}$$

Addition des Dreifachen der zweiten Zeile zur dritten

$$\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 6 & 1 & -3 & 1 \end{array}$$

Division der dritten Zeile durch 6

$$\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & \frac{1}{6} & -\frac{1}{2} & \frac{1}{6} \end{array}$$

Nun addiert man wieder umgekehrt die dritte zur zweiten etc.

$$\begin{array}{ccc|ccc} 1 & 2 & 0 & \frac{1}{2} & \frac{3}{2} & -\frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{6} & \frac{1}{2} & \frac{1}{6} \\ 0 & 0 & 1 & \frac{1}{6} & -\frac{1}{2} & \frac{1}{6} \end{array}$$

$$\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{6} & \frac{1}{2} & -\frac{5}{6} \\ 0 & 1 & 0 & \frac{1}{6} & \frac{1}{2} & \frac{1}{6} \\ 0 & 0 & 1 & \frac{1}{6} & -\frac{1}{2} & \frac{1}{6} \end{array}$$

Nun steht rechts die inverse Matrix. (Wenn man sich nicht verrechnet.)

**Bemerkung.** Es dürfen *nur* elementare Zeilentransformationen verwendet werden. (Oder *nur* elementare Spaltentransformationen.)

*Beweis:* Wir betrachten ein Tableau  $A|E_n$ . Dann multiplizieren wir beiden Matrizen von links mit Elementarmatrizen, Permutationsmatrizen und invertierbaren Diagonalmatrizen. Wir erhalten ein Tableau

$$S_1 \dots S_N A | S_1 \dots S_n E_n$$

Sei die linke Matrix gleich  $E_n$ . Dann gilt also

$$E_n = S_1 \dots S_N A \Rightarrow S_N^{-1} S_{N-1}^{-1} \dots S_1^{-1} = A$$

(Multiplikation von links mit den Inversen der  $S_i$  und  $T_j$ .) Durch Invertieren der Relation erhalten wir

$$S_1 \dots S_N = A$$

also die rechte Matrix des Tableaus.  $\square$

## Determinanten von linearen Abbildungen

**Definition 6.14.** Sei  $V$  ein endlich-dimensionaler  $k$ -Vektorraum,  $f : V \rightarrow V$  ein Endomorphismus mit darstellender Matrix  $M(f)$ . Dann heißt

$$\det(f) = \det M(f)$$

Determinante von  $f$ .

**Lemma 6.15.** Die Determinante eines Endomorphismus ist wohldefiniert.

*Beweis:* Sind  $M$  und  $M'$  darstellende Matrizen bezüglich zweier Basen, so ist

$$M' = S M S^{-1}$$

mit einer Basiswechselmatrix  $S$ . Es folgt mit der Produktformel

$$\det(M') = \det(S) \det(M) \det(S^{-1}) = \det(M).$$

$\square$

**Satz 6.16.** Sei  $V$  ein endlich-dimensionaler Vektorraum,  $f : V \rightarrow V$  ein Endomorphismus. Dann sind äquivalent:

- (i)  $f$  ist injektiv.
- (ii)  $\text{rg}(f) = \dim V$ .
- (iii)  $f$  ist ein Automorphismus.
- (iv)  $\det(f) \neq 0$ .

*Beweis:* Nach der Dimensionsformel ist

$$\dim \text{Ker}(f) + \text{rg}(f) = \dim V,$$

daher sind die beiden ersten Aussagen äquivalent. Insbesondere ist  $f$  dann auch ein Isomorphismus. Dies äquivalent dazu, dass die darstellende Matrix invertierbar ist, also die Determinante ungleich 0.  $\square$



## Kapitel 7

# Eigenwerte und Eigenvektoren

Von nun an studieren wir Endomorphismen von Vektorräumen.

**Definition 7.1.** Sei  $k$  ein Körper,  $V$  ein  $k$ -Vektorraum,  $f : V \rightarrow V$  ein Endomorphismus. Ein Vektor  $v \in V \setminus \{0\}$  heißt Eigenvektor von  $f$  zum Eigenwert  $\lambda \in k$ , falls gilt

$$f(v) = \lambda v$$

**Bemerkung.** 0 ist kein Eigenvektor, da sonst jedes  $\lambda \in k$  Eigenwert wäre. Der Eigenwert 0 ist erlaubt. Er tritt auf, wenn  $f$  nicht invertierbar ist.

**Beispiel.** (i) Sei  $f : \mathbb{Q}^2 \rightarrow \mathbb{Q}^2$  gegeben durch Multiplikation mit  $\begin{pmatrix} 3 & 1 \\ -4 & -2 \end{pmatrix}$ .

Dann gilt

$$f \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ -4 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \end{pmatrix}$$

Damit ist  $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$  Eigenvektor zum Eigenwert 2.

- (ii) Sei  $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  die Spiegelung an einer Nullpunktsgerade. Ein Vektor in Richtung der Spiegelachse ist dann Eigenvektor zum Eigenwert 1. Ein senkrecht dazu stehender Vektor ist Eigenvektor zum Eigenwert  $-1$ .
- (iii) Sei  $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  eine Drehung um 0 um den Winkel  $\alpha$ . Dann ist kein Vektor Eigenvektor (Ausnahmen:  $\alpha = 0$ , alle Vektoren ungleich 0 sind Eigenvektoren zum Eigenwert 1; oder  $\alpha = \pi$ , alle Vektoren ungleich 0 sind Eigenvektoren zum Eigenwert  $-1$ .)

In der Physik wimmelt es von Eigenwertproblemen.

**Beispiel.** (i) (Kreiselgleichungen) Wir betrachten die Bewegungsgleichung eines starren Körpers im Raum. Die Bewegung des Schwerpunkts kann abgespalten werden, ohne Einschränkung ist dieser fest im Mittelpunkt des Koordinatensystems. Dann kann der Körper sich immer noch drehen. Die Bewegungsgleichung lautet

$$\vec{L} = I \vec{\omega}$$

wobei  $\vec{\omega} \in \mathbb{R}^3$  die Winkelgeschwindigkeit ist,  $\vec{L} \in \mathbb{R}^3$  der Drehimpuls und  $I \in M_3(\mathbb{R})$  der *Trägheitstensor*. Er wird aus der Masseverteilung des starren Körpers berechnet. Ein Eigenvektor des Trägheitstensors ist eine (meta)-stabile Drehachse des Körpers. Es ist eine physikalische oder mathematische Tatsache, dass jeder starre Körper genau drei, auf einander senkrecht stehende Eigenachsen hat (LA 2: wegen  $I = I^t$ ).

- (ii) Wir betrachten die Schwingungen einer eingespannten Saite. Ihre Zustände bilden einen (unendlichdimensionalen) Vektorraum. Als Operator betrachten wir die Schwingungsgleichung. Die "Grundschiwungen" und "Overtöne" der Saite sind die Eigenvektoren. Man spricht ja auch von Eigenfrequenzen.
- (iii) Mathematisch sehr ähnlich sind die Gleichungen der Quantenmechanik: Wir betrachten das quantenmechanische Modell des Wasserstoffatoms. Seine Zustände werden durch Wellengleichungen beschrieben. Der Endomorphismus ist der Schrödingeroperator bzw. die Schrödingergleichung. Dann sind die Eigenvektoren die aus Bildern bekannten Elektronenwolken. Die zugehörigen Energieniveaus sind die Eigenwerte.
- (iv) Googles Algorithmus zur Bewertung von Webseiten ist ebenfalls eine Eigenwertaufgabe mit einem riesigen linearen Gleichungssystem.

Wie kommt man auf den Eigenwert 2 im ersten Beispiel? Sei wieder  $f : \mathbb{Q}^2 \rightarrow \mathbb{Q}^2$ , wobei  $f$  Multiplikation mit  $\begin{pmatrix} 3 & 1 \\ -4 & -2 \end{pmatrix}$ .

Gesucht ist ein  $\lambda \in \mathbb{Q}$ , so dass es ein  $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \neq 0$  gibt mit

$$f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 3 & 1 \\ -4 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Dies ist (für festes  $\lambda$ ) ein lineares Gleichungssystem. Wir bringen es in vertraute Form:

$$\begin{pmatrix} 3 & 1 \\ -4 & -2 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 - \lambda & 1 \\ -4 & -2 - \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$$

Dies ist ein homogenes Gleichungssystem, und wir suchen nach einem nichttrivialen Element des Kerns der linearen Abbildung gegeben durch Multiplikation



mit der Matrix  $\begin{pmatrix} 3-\lambda & 1 \\ -4 & -2-\lambda \end{pmatrix}$ . Diese existiert genau dann, wenn die Determinante verschwindet, also

$$\begin{vmatrix} 3-\lambda & 1 \\ -4 & -2-\lambda \end{vmatrix} = (3-\lambda)(-2-\lambda) + 4 = \lambda^2 - \lambda - 2 = 0$$

Das nichttriviale  $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$  existiert also genau dann, wenn  $\lambda$  eine Nullstelle des quadratischen Polynoms ist. Diese sieht man sofort:

$$\lambda^2 - \lambda - 2 = (\lambda - 2)(\lambda + 1)$$

Die Eigenwerte sind 2 und  $-1$ . Die zugehörigen Eigenvektoren zu berechnen ist dann die Lösung des linearen Gleichungssystems. Auf jeden Fall gibt es sie! Offensichtlich funktioniert diese Argumentation für alle endlichdimensionalen Vektorräume. Wir müssen uns daher etwas mit Polynomen beschäftigen.

## Polynomringe

Sei  $k$  ein Körper. Ein *Polynom* über  $k$  ist ein formaler Ausdruck der Form  $\sum_{i=0}^n a_i X^i$  mit  $n \in \mathbb{N}_0$ ,  $a_i \in k$  und einer *Unbestimmten*  $X$ . Die Menge der Polynome heißt *Polynomring*  $k[X]$ . Sie wird mit Addition und Multiplikation von Polynomen zu einem Ring.

### Beispiel.

$$(1+2X+X^2)(X^4-X^5) = X^4+2X^5+X^6-X^5-2X^6-X^7 = X^4+X^5-X^6-X^7$$

Formal sauber:

**Definition 7.2.** Sei  $k$  ein Körper. Der Polynomring  $k[X]$  über  $k$  ist die Menge

$$k[X] = \bigoplus_{i=0}^{\infty} k = \{(a_i)_{i \in \mathbb{N}_0} \mid a_i = 0 \text{ für fast alle } i \in \mathbb{N}_0\}$$

mit der komponentenweisen Addition und dem Cauchyprodukt

$$(a_i)_{i=0}^{\infty} (b_j)_{j=0}^{\infty} = \left( \sum_{i+j=r} a_i b_j \right)_{r=0}^{\infty}$$

Wir schreiben suggestiver  $X = (0, 1, 0, \dots)$  und

$$(a_i)_{i=0}^{\infty} = \sum_{i=0}^n a_i X^i$$

mit  $n$  genügend groß. Der Grad  $\deg P$  eines Polynoms ist das Maximum der  $i \in \mathbb{N}_0$  mit  $a_i \neq 0$ . (Dabei setzen wir  $\deg 0 = -\infty$ ).

**Lemma 7.3.**  $k[X]$  ist ein kommutativer Ring mit Eins. Für  $P, Q \in k[X]$  gilt

$$\deg(PQ) = \deg P + \deg Q, \deg(P + Q) \leq \max(\deg P, \deg Q)$$

Insbesondere gilt  $P, Q \neq 0 \Rightarrow PQ \neq 0$ .

*Beweis:* Bezüglich der Addition ist dies ein Spezialfall der direkten Summe von Vektorräumen.  $(k[X], +)$  ist insbesondere eine abelsche Gruppe.  $k[X]$  hat die Basis  $e_i = (\delta_{ij})_{j=0}^{\infty}$  mit Eins an der Stelle  $i$  und 0 sonst. Man überprüft leicht, dass  $e_0$  das neutrale Element der Multiplikation ist. Wir schreiben daher ab sofort  $e_0 = 1$  und  $ae_0 = a$  für alle  $a \in k$ . Dann überprüft man leicht, dass  $aP$  für alle  $a \in k$  und  $P \in k[X]$  den gleichen Wert ergibt, egal ob wir es als skalare Multiplikation des Vektorraums oder als Cauchyprodukt auffassen.

Wir schreiben  $e_1 = X$ . Man sieht sofort, dass  $e_i = X^i$ . Daher hat jedes Polynom eine eindeutige Darstellung  $\sum_{i=0}^{\infty} a_i X^i$ . Die Produktformel in dieser Schreibweise lautet

$$\left( \sum_{i=0}^{\infty} a_i X^i \right) \left( \sum_{j=0}^{\infty} b_j X^j \right) = \sum_{r=0}^{\infty} \left( \sum_{i+j=r} a_i b_j \right) X^r$$

Die Gültigkeit der Assoziativ- und Distributivgesetze rechnet man ebenfalls leicht nach.

Nun betrachten wir die Gradaussagen. Falls  $P = 0$  oder  $Q = 0$ , so gelten die Aussagen. Sei also  $n = \deg P \geq 0$ ,  $m = \deg Q \geq 0$ , also

$$P = \sum_{i=0}^n a_i X^i, Q = \sum_{j=0}^m b_j X^j$$

mit  $a_n, b_m \neq 0$ . Sei ohne Einschränkung  $n \leq m$ . Dann ist

$$PQ = \sum_{r=0}^{n+m} \sum_{i+j=r} a_i b_j X^r$$

denn der maximale Werte von  $i + j$ , so dass  $a_i b_j \neq 0$  ist  $n + m$ . Gleichzeitig hat für  $r = n + m$  die Summe nur einen Summanden  $a_n b_m$ , der ungleich Null ist. Damit ist  $\deg PQ = n + m$ . Für die Summe gilt

$$P + Q = \sum_{r=0}^m (a_r + b_r) X^r$$

denn für  $i > m$  ist  $a_i = b_i = 0$ . Damit ist  $\deg(P + Q) \leq m$ . (Gleichheit folgt nicht. Es könnte nämlich  $n = m$ , und  $a_n = -b_n$  sein.)  $\square$

Jedes Polynom definiert eine Abbildung  $k \rightarrow k$ , die zugehörige Polynomfunktion.

**Definition 7.4.** Sei  $P = \sum_{i=0}^n a_i X^i$  ein Polynom,  $\lambda \in k$ . Dann heißt

$$P(\lambda) = \sum_{i=0}^n a_i \lambda^i$$

Wert von  $P$  in  $\lambda$ . Insbesondere heißt  $\lambda$  Nullstelle von  $P$ , falls  $P(\lambda) = 0$ . Die Abbildung

$$\text{ev} : k[X] \rightarrow \text{Abb}(k, k) \quad P \mapsto (\lambda \mapsto P(\lambda))$$

heißt Auswertungsabbildung (*Evaluierung, englisch Evaluation*). Die Abbildungen im Bild heißen polynomiale Funktionen.

**Lemma 7.5.**  $\text{ev}$  ist eine lineare Abbildung von  $k$ -Vektorräumen und verträglich mit Multiplikation.

*Beweis:* Hinschreiben. □

**Satz 7.6.** Gegeben seien paarweise verschiedene  $\lambda_0, \dots, \lambda_n \in k$  und beliebige  $b_0, \dots, b_n \in k$ . Dann gibt es genau ein Polynom vom Grad höchstens  $n$  mit

$$P(\lambda_i) = b_i \quad \text{für alle } i = 0, \dots, n$$

Inbesondere ist ein Polynom vom Grad  $n$  durch die Werte in  $n+1$  Elementen von  $k$  eindeutig bestimmt.

*Beweis:* Wir stellen die Gleichung auf. Gesucht sind  $a_0, \dots, a_n \in k$  mit

$$\begin{aligned} a_0 + \lambda_0 a_1 + \lambda_0^2 a_2 + \dots + \lambda_0^n a_n &= b_0 \\ a_0 + \lambda_1 a_1 + \lambda_1^2 a_2 + \dots + \lambda_1^n a_n &= b_1 \\ &\dots \\ a_0 + \lambda_n a_1 + \lambda_n^2 a_2 + \dots + \lambda_n^n a_n &= b_n \end{aligned}$$

Dies sind  $n+1$  inhomogene lineare Gleichungen für  $n+1$  Unbekannte  $a_0, \dots, a_n$ . Damit haben wir eine Gelegenheit unser Wissen über lineare Gleichungssystem zu nutzen. Die Koeffizientenmatrix lautet

$$L = \begin{pmatrix} 1 & \lambda_0 & \lambda_0^2 & \dots & \lambda_0^n \\ 1 & \lambda_1 & \lambda_1^2 & \dots & \lambda_1^n \\ \dots & \dots & \dots & \dots & \dots \\ 1 & \lambda_n & \lambda_n^2 & \dots & \lambda_n^n \end{pmatrix}$$

Dies ist eine Vandermondsche Matrix. Ihre Determinante ist (Übungsaufgabe)

$$\det L = \prod_{j>i} (\lambda_j - \lambda_i)$$

Sie ist ungleich null, da die  $\lambda_i$  paarweise verschieden sind. Damit ist die Koeffizientenmatrix invertierbar, und das Gleichungssystem ist eindeutig lösbar. □

**Korollar 7.7.** Sei  $P$  ein Polynom vom Grad höchstens  $n$  mit  $n+1$  Nullstellen. Dann ist  $P = 0$ .

*Beweis:* Dies ist der Spezialfall  $b_0 = \dots = b_n = 0$  der Eindeutigkeit.  $\square$

**Korollar 7.8.** Hat  $k$  unendlich viele Elemente, so ist die Auswertungsabbildung  $\text{ev} : k[X] \rightarrow \text{Abb}(k, k)$  injektiv.

*Beweis:* Sei  $P \in \text{Ker ev}$ . Falls  $P \neq 0$ , so wenden wir das vorherige Korollar mit  $n = \deg P$  an. Da  $P$  in allen Elementen von  $k$  verschwindet, hat es mehr als  $n$  Nullstellen.  $\square$

**Bemerkung.** Wer in erster Linie an  $\mathbb{R}$  und  $\mathbb{C}$  interessiert ist, braucht als den Unterschied zwischen formalen Polynomen und polynomialen Funktionen nicht zu machen. Aber es gibt ja auch endliche Körper!

## Charakteristische Polynome

Nun kehren wir zurück zur Eigenwertfrage.

**Definition 7.9.** Sei  $V$  ein endlichdimensionaler Vektorraum mit Basis  $A$ . Sei  $f : V \rightarrow V$  ein Endomorphismus. Dann heißt

$$\chi_f = \det(f - X \text{id}) = \det(M_A^A(f) - X E_n) \in k[X]$$

charakteristisches Polynom von  $f$ .

Hier fassen wir  $M_A^A(f) - X E_n$  auf als Matrix in  $M_n(k[X])$ . Unsere Theorie von Determinanten funktionierte ja für kommutative Ringe mit 1. Beim Einsetzen von  $\lambda \in k$  erhalten wir dann

$$\chi_f(\lambda) = \det(M_A^A(f) - \lambda E_n) \in k.$$

**Lemma 7.10.** Das charakteristische Polynom ist unabhängig von der Wahl der Basis.

*Beweis:* Im allgemeinen sei  $B$  eine zweite Basis. Dann gilt nach der Basiswechselformel

$$M_B^B(f) = S^{-1} M_A^A(f) S$$

für eine invertierbare Matrix  $S \in M_n(k)$ . Es folgt

$$M_B^B(f) - X E_n = S^{-1} (M_A^A(f) - X E_n) S \in M_n(k[X])$$

Nach der Produktformel gilt

$$\det(M_B^B(f) - X E_n) = \det(S)^{-1} \det(M_A^A(f) - X E_n) \det(S) = \det(M_A^A(f) - X E_n)$$

$\square$

Wie sieht das charakteristische Polynom aus? Sei  $(a_{ij})_{i,j=1}^n$  darstellende Matrix von  $f$ . Dann ist mit der Leibniz-Formel

$$\begin{aligned}\chi_f &= \det((a_{ij} - X\delta_{ij})) \\ &= \sum_{\sigma \in S_n} (a_{1\sigma(1)} - X\delta_{1\sigma(1)})(a_{2\sigma(2)} - X\delta_{2\sigma(2)}) \dots (a_{n\sigma(n)} - X\delta_{n\sigma(n)})\end{aligned}$$

Wir multiplizieren aus und fassen zusammen. Es treten maximal  $n$  Faktoren  $X$  auf, nämlich genau für den Summanden mit  $1 = \sigma(1), 2 = \sigma(2), \dots, n = \sigma(n)$  bzw.  $\sigma = \text{id}$ . Der Vorfaktor lautet  $(-1)^n$ . Einfach ist auch der konstante Term. Man erhält ihn als  $\chi_f(0) = \det(f)$ . Also:

$$\chi_f = (-1)^n X^n + \dots + \det(f)$$

Insbesondere ist der Grad von  $\chi_f$  die Dimension von  $V$ .

**Satz 7.11.** Sei  $k$  ein Körper,  $V$  ein endlichdimensionaler  $k$ -Vektorraum,  $f : V \rightarrow V$  ein Endomorphismus,  $\lambda \in k$ . Dann sind äquivalent:

- (i)  $\lambda \in k$  ist ein Eigenwert von  $f$ .
- (ii)  $\lambda$  ist Nullstelle des charakteristischen Polynoms, d.h.  $\chi_f(\lambda) = 0$ .

*Beweis:* Wir formen die Aussagen äquivalent um:

$$\begin{aligned}\lambda \text{ Eigenwert} &\Leftrightarrow \\ f(v) = \lambda v \text{ für ein } v \in V \setminus \{0\} &\Leftrightarrow \\ f(v) - \lambda v = 0 \text{ für ein } v \in V \setminus \{0\} &\Leftrightarrow \\ (f - \lambda \text{id})(v) = 0 \text{ für ein } v \in V \setminus \{0\} &\Leftrightarrow \\ \text{Ker}(f - \lambda \text{id}) \neq 0 &\Leftrightarrow \\ \det(f - \lambda \text{id}) = 0 &\Leftrightarrow \\ \chi_f(\lambda) = 0 &\end{aligned}$$

□

**Korollar 7.12.** Sei  $\dim V = n$ ,  $f : V \rightarrow V$  Endomorphismus. Dann hat  $f$  höchstens  $n$  Eigenwerte.

*Beweis:* Es ist  $\deg \chi_f = n$ . Damit hat es höchstens  $n$  Nullstellen. □

**Bemerkung.** Nullstellen von quadratischen Polynomen findet man durch quadratische Ergänzung und die zugehörige Lösungsformel. Sie gilt über allen Körpern (soweit man die auftretenden Quadratwurzeln ziehen kann!). Auch für Gleichungen vom Grad drei und vier gibt es die Cardanoschen Formeln, die seit dem 16. Jahrhundert bekannt sind. Für die Nullstellen von allgemeinen Polynomen von höherem Grad gibt es solche Formeln nicht. Dies ist ein tiefes Resultat der Algebra, 1824 erstmals bewiesen von Abel. Der Beweis dieses Satzes ist zentraler Gegenstand der Veranstaltung Algebra und Zahlentheorie im 3. Semester.

Hat jeder Endomorphismus Eigenwerte?

**Beispiel.** Wir betrachten  $k = \mathbb{R}$ ,  $V = \mathbb{R}^2$  und  $f = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ . (Drehung um  $\pi/2$ ). Dann ist

$$\chi_f = \begin{vmatrix} -X & 1 \\ -1 & -X \end{vmatrix} = X^2 + 1$$

Dieses Polynom hat keine Nullstellen in  $\mathbb{R}$ , also hat  $f$  keine Eigenwerte. Über  $\mathbb{C}$  hat es jedoch die Nullstellen  $\pm i$ , also sehr wohl Eigenwerte.

**Definition 7.13.** Ein Körper heißt algebraisch abgeschlossen, wenn jedes nicht-konstante Polynom eine Nullstelle hat.

**Theorem 7.14** (Fundamentalsatz der Algebra).  $\mathbb{C}$  ist algebraisch abgeschlossen.

*Beweis:* Tief. Am leichtesten geht es mit den Methoden der Funktionentheorie.  $\square$

**Korollar 7.15.** Sei  $k$  algebraisch abgeschlossen. Dann hat jeder Endomorphismus eines endlichdimensionalen Vektorraums einen Eigenwert.

*Beweis:* Das charakteristische Polynom hat eine Nullstelle.  $\square$

Die Voraussetzung an die Dimension ist nötig.

**Beispiel.** Sei  $V = k[X]$ ,  $f$  die Multiplikation mit  $X$ . Dann hat  $f$  keinen Eigenwert, denn

$$X \sum_{i=0}^n a_i X^i = \sum_{i=0}^n a_i X^{i+1} = \lambda \sum_{i=0}^n a_i X^i$$

ist äquivalent zu

$$0 = \lambda a_0, a_0 = \lambda a_1, a_1 = \lambda a_2, \dots, a_{n-1} = \lambda a_n, a_n = 0$$

Hieraus folgt rekursiv von oben  $a_i = 0$  für alle  $i$ . Das Nullpolynom ist aber als Eigenvektor nicht erlaubt.

Bisher haben wir uns auf die Eigenwerte konzentriert. Nun wollen wir die Eigenvektoren noch etwas studieren.

**Definition 7.16.** Sei  $V$  ein  $k$ -Vektorraum,  $f : V \rightarrow V$  ein Endomorphismus. Sei  $\lambda \in k$  ein Eigenwert. Dann heißt

$$V_\lambda = \{v \in V \mid f(v) = \lambda v\}$$

Eigenraum zu  $\lambda$ .

Der Eigenraum enthält die Eigenvektoren zum Eigenwert  $\lambda$  und außerdem 0.

**Lemma 7.17.** Der Eigenraum ist ein Untervektorraum von  $V$ .

*Beweis:* Seien  $a, b \in k$ ,  $x, y \in V_\lambda$ . Dann gilt

$$f(ax + by) = af(x) + bf(y) = a\lambda x + b\lambda y = \lambda(ax + by)$$

d.h.  $ax + by \in V_\lambda$ . □

Wie steht es mit Eigenvektoren zu verschiedenen Eigenwerten?

**Satz 7.18.** *Sei  $V$  ein Vektorraum,  $f : V \rightarrow V$  ein Endomorphismus. Seien  $v_1, \dots, v_n$  Eigenvektoren zu paarweise verschiedenen Eigenwerten  $\lambda_1, \dots, \lambda_n$ . Dann ist die Familie  $v_1, \dots, v_n$  linear unabhängig.*

*Beweis:* Wir beweisen die Aussage mit vollständiger Induktion nach  $n$ . Für  $n = 1$  gilt die Aussage, denn Eigenvektoren sind ungleich 0. Die Aussage sei nun wahr für  $n - 1$ . Wir betrachten ein  $n$ -Tupel von Eigenvektoren wie im Satz. Seien  $a_1, \dots, a_n \in k$  mit

$$a_1 v_1 + \dots + a_n v_n = 0. \tag{I}$$

Angenommen dies ist eine nicht-triviale Relation. Ohne Einschränkung ist  $a_1 \neq 0$ . Durch Anwenden von  $f$  erhalten wir neue Relationen:

$$0 = f(a_1 v_1 + \dots + a_n v_n) = a_1 f(v_1) + \dots + a_n f(v_n) = a_1 \lambda_1 v_1 + \dots + a_n \lambda_n v_n \tag{II}$$

Wir bilden  $\text{II} - \lambda_1 \text{I}$  und erhalten die neue Relation

$$(\lambda_1 a_1 - a_1 \lambda_1) v_1 + (\lambda_1 a_2 - \lambda_2 a_2) v_2 + \dots + (\lambda_1 a_n - \lambda_n a_n) = 0.$$

In dieser Relation ist der erste Koeffizient 0. Es handelt sich also um eine Linearkombination von  $n - 1$  der  $v_i$ . Nach Induktionsvoraussetzung sind diese linear unabhängig, es folgt also

$$(\lambda_1 - \lambda_i) a_i = 0 \quad \text{für } i = 2, \dots, n.$$

Da die  $\lambda_i$  paarweise verschieden ist, müssen die  $a_i$  für  $i \geq 2$  verschwinden. Die Relation (I) ist also nur

$$a_1 v_1 = 0.$$

Das ist aber (Induktionsanfang) unmöglich. Die Relation (I) ist also trivial, die Vektoren sind linear unabhängig. □

**Korollar 7.19.** *Sei  $V$  ein  $n$ -dimensionaler Vektorraum,  $f : V \rightarrow V$  ein Endomorphismus mit  $n$  verschiedenen Eigenwerten. Dann hat  $f$  eine Basis aus Eigenvektoren. Bezüglich dieser Basis  $A$  lautet die darstellene Matrix*

$$M_A^A(f) = [\lambda_1, \dots, \lambda_n] := \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ & & \dots & & \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

wobei die  $\lambda_i$  die Eigenwerte sind. Alle Eigenräume sind eindimensional.

*Beweis:* Seien  $v_1, \dots, v_n$  Eigenvektoren zu den verschiedenen Eigenwerten  $\lambda_1, \dots, \lambda_n$ . Nach dem Satz sind sie linear unabhängig, wegen  $\dim V = n$  also eine Basis. Wir berechnen die darstellende Matrix:

$$f(v_i) = \lambda_i v_i = \sum_{j=1}^n \delta_{ij} \lambda_i v_j$$

Zuletzt berechnen wir den Eigenraum: Sei  $v = \sum_{j=1}^n a_j v_j$  ein Eigenvektor zu Eigenwerte  $\lambda_i$ . Es gilt also  $f(v) = \lambda_i v$ , bzw.

$$f\left(\sum_{j=1}^n a_j v_j\right) = \sum_{j=1}^n a_j \lambda_j v_j = \sum_{j=1}^n \lambda_i a_j v_j \Leftrightarrow a_j \lambda_j = a_j \lambda_i \text{ für alle } j$$

Wegen  $\lambda_i \neq \lambda_j$  für  $i \neq j$  folgt hieraus  $a_j = 0$  für  $j \neq i$ . Damit hat  $v$  die Form  $a_i v_i$ . Der Eigenraum hat den Basisvektor  $v_i$ .  $\square$

Dies ist insbesondere der Fall, wenn die  $\chi_f$   $n$  verschiedene Nullstellen hat.

**Bemerkung.** Physiker nennen solche Endomorphismen *nicht-degeneriert*, andernfalls *degeneriert*. Da sich durch kleine Störungen es Systems die Eigenwerte ändern, verschwinden auch zufällige Gleichheiten in der realen Welt.

Nicht jeder Endomorphismus hat eine Basis aus Eigenvektoren! Einen Hinderungsgrund haben wir schon gesehen, nämlich wenn der Körper nicht algebraisch abgeschlossen ist. Aber selbst über  $\mathbb{C}$  ist es falsch.

**Beispiel.** Sei  $V = \mathbb{C}^2$ ,  $f = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ . Das charakteristische Polynom ist  $X^2$ . Einzige Nullstelle (egal in welchem Körper) ist 0. Wir berechnen den Eigenraum  $V_0$ , also den Kern:

$$0 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ 0 \end{pmatrix}$$

Hieraus folgt  $y = 0$ , aber  $x$  beliebig. Der Eigenraum ist eindimensional. Einziger Eigenvektor von  $f$  (bis auf Vielfache) ist  $e_1$ .

Dies systematisch zu verstehen, wird eine der Aufgaben der LA 2 sein.



## Kapitel 8

# Etwas Gruppentheorie

Nebenbei haben wir verschiedene Gruppen getroffen und Eigenschaften verifiziert. Wir fassen das nun explizit zusammen.

Wir kennen bereits einige Gruppen, insbesondere  $S_n$  oder  $\text{Aut}_k(V)$  für einen Vektorraum  $V$ . Ist  $k$  ein Körper, so haben wir die *additive Gruppe*  $\mathbb{G}_a(k) = (k, +)$  und die *multiplikative Gruppe*  $\mathbb{G}_m(k) = (k^*, \cdot)$ .

**Sprechweise:** Sei  $A \in M_{m \times n}(k)$ . Wir sagen  $A$  hat vollen Rang, falls  $\text{rg}A$  maximal ist, also das Minimum von  $n$  und  $m$ .

Invertierbare Matrizen sind also genau die quadratischen Matrizen von vollem Rang.

**Definition 8.1.** Die Menge  $\text{GL}_n(k)$  der invertierbaren Matrizen in  $M_n(k)$  heißt allgemeine lineare Gruppe (*general linear group*).

Speziell für  $n = 1$  gilt  $\mathbb{G}_m(k) := \text{GL}_1(k) = (k^*, \cdot)$ .

**Lemma 8.2.**  $\text{GL}_n(k)$  ist eine Gruppe bezüglich der Multiplikation von Matrizen.

*Beweis:* Matrizenmultiplikation ist assoziativ. Das neutrale Element ist die Einheitsmatrix  $E_n$ . Nach Definition hat  $A \in \text{GL}_n(k)$  eine inverse Matrix. Schließlich sind Produkte von invertierbaren Matrizen wieder invertierbar, denn für  $S, T \in \text{GL}_N(k)$  ist

$$T^{-1}S^{-1}ST = E_n \Rightarrow (ST)^{-1} = T^{-1}S^{-1}$$

□

**Bemerkung.** Dies ist einfach die Matrizenversion der Aussage, dass  $\text{Aut}(V)$  eine Gruppe bezüglich der Komposition ist.

### Die symmetrische Gruppe

Es gibt verschiedene Arten, Permutationen hinzuschreiben. Am einfachsten ist, alle Werte hinzuschreiben. Die Permutation  $\sigma \in S_n$  wird dann angegeben als

$$\begin{pmatrix} 1 & 2 & 3 & \dots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \dots & \sigma(n) \end{pmatrix}.$$

**Lemma 8.3.** Die symmetrische Gruppe  $S_n$  hat  $n! = n \cdot (n-1) \dots 2 \cdot 1$  Elemente.

*Beweis:* Für  $\sigma(1)$  gibt es  $n$  Möglichkeiten. Für  $\sigma(2)$  scheidet der Wert  $\sigma(1)$  aus, also bleiben  $n-1$  Möglichkeiten usw.  $\square$

**Definition 8.4.** Für  $1 \leq i, j \leq n$ ,  $i \neq j$  sei  $\sigma = (ij) \in S_n$  definiert durch

$$\sigma(s) = \begin{cases} s & s \neq i, j \\ j & s = i \\ i & s = j \end{cases}$$

Elemente dieser Form heißen Transpositionen.

Für  $i = j$  ist die Abbildung ebenfalls definiert, aber dann die Identität. Diese ist keine Transposition. Transpositionen sind zu sich selbst invers,  $(ij)(ij) = \text{id}$ . Die folgenden Eigenschaft haben wir bereits mehrfach benutzt:

**Satz 8.5.** Jedes Element von  $S_n$  ist ein Produkt von Transpositionen. Hierbei genügen die Transpositionen der Form  $(i \ i+1)$  für  $i = 1, \dots, n-1$ .

*Beweis:* Wir betrachten eine Permutation  $\sigma$  und geben sie wie folgt an:

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

Diese Permutation können wir durch Vertauschungen "in die richtige Reihenfolge bringen". Das geht wie folgt: Wir betrachten  $\tau_1 = (1\sigma(1))$ ,  $\sigma_1 = \tau_1\sigma$ . Es ist

$$\tau_1\sigma(1) = (1\sigma(1))(\sigma(1)) = 1$$

In  $\sigma_1$  steht 1 an der richtigen Stelle. Sei  $\tau_2 = (2\sigma_1(2))$ ,  $\sigma_2 = \tau_2\sigma_1$ . Es ist

$$\tau_2\sigma_1(2) = (2\sigma_1(2))(\sigma_1(2)) = 2$$

d.h. in  $\sigma_2$  steht 2 an der richtigen Stelle. Gleichzeitig ist

$$\tau_2\sigma_1(1) = \tau_2(1) = 1$$

denn  $\sigma_1(2) \neq \sigma_1(1) = 1$ . Dieses Verfahren wiederholen wir: Sei induktiv  $\tau_i = (i\sigma_{i-1}(i))$  und  $\sigma_i = \tau_i\sigma_{i-1}$ . Dann stehen in  $\sigma_i$  die Zahlen  $1, \dots, i$  an der richtigen Stelle. Wir erhalten also  $\sigma_n = \text{id}$ . Es gilt

$$\text{id} = \tau_n\tau_{n-1}\dots\tau_1\sigma \Rightarrow \tau_1\tau_2\dots\tau_n = \sigma$$

da  $\tau_i^{-1} = \tau_i$ . Eventuell sind einige der  $\tau_i = \text{id}$ , diese Faktoren lassen wir weg. Dann haben wir  $\sigma$  als Produkt von Transpositionen geschrieben. Um den den Zusatz zu beweisen, betrachten wir jetzt noch ein  $(ij)$  mit  $i < j$ . Es gilt

$$(ij) = (i \ i+1) \dots (j-2 \ j-1)(j-1 \ j) \dots (i+1 \ i+2)(i \ i+1)$$

$\square$

**Definition 8.6.** Sei  $G$  eine Gruppe. Eine Teilmenge  $S \subset G$  heißt Erzeugendensystem, wenn jedes Element aus  $G$  endliches Produkt von Elementen aus  $S$  und  $S^{-1} = \{s^{-1} \mid s \in S\}$  ist.

Wir haben also gezeigt:  $S_n$  wird von Transpositionen erzeugt. Es genügen die Transpositionen der Form  $(ii+1)$ .

## Die allgemeine lineare Gruppe

Als nächstes wollen wir ein Erzeugendensystem für die  $GL_n(k)$  beschreiben. Dafür geben wir einige besonders einfache Elemente an.

**Definition 8.7.** (i) Seien  $\lambda_1, \dots, \lambda_n \in k^*$ . Dann schreiben wir  $[\lambda_1, \dots, \lambda_n]$  für die invertierbare Diagonalmatrix

$$[\lambda_1, \dots, \lambda_n] = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ & & \dots & & \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

(ii) Für jedes  $1 \leq i, j \leq n$  mit  $i \neq j$  und jede  $\lambda \in k$  sei

$$E_{ij}(\lambda) = (a_{st})$$

wobei alle  $a_{st} = 0$  außer die Diagonaleinträge  $a_{ss} = 1$  für alle  $s$  und  $a_{ij} = \lambda$ . Der Eintrag  $\lambda$  steht also in Zeile  $i$ , Spalte  $j$ . Matrizen vom Typ  $E_{ij}(\lambda)$  heißen Elementarmatrizen.

(iii) Für jedes  $1 \leq i, j \leq n$  mit  $i \neq j$  sei

$$P_{(ij)} = (b_{st})$$

wobei alle Einträge  $b_{st} = 0$  sind außer  $b_{ss} = 1$  für  $s \neq i, j$ , und  $b_{ij} = b_{ji} = 1$ . Matrizen vom Typ  $P_{ij}$  heißen Vertauschungsmatrizen.

Diese Matrizen sind invertierbar. Es gilt

$$[\lambda_1, \dots, \lambda_n]^{-1} = [\lambda_1^{-1}, \dots, \lambda_n^{-1}] \quad E_{ij}(\lambda)^{-1} = E_{ij}(-\lambda) \quad P_{(ij)}^{-1} = P_{(ij)}$$

Durch diese Matrizen werden die Basiswechselmatrizen zu elementaren Transformationen realisiert.

**Lemma 8.8.** Sei  $A = (v_1 \ \dots \ v_n)$  mit  $v_i \in k^n$ . Sei  $B = \begin{pmatrix} w_1 \\ \dots \\ w_n \end{pmatrix}$  für Zeilenvektoren  $w_j \in (k^n)^*$ .

(i) Seien  $\lambda_1, \dots, \lambda_n \in k$ . Dann ist

$$A[\lambda_1, \dots, \lambda_n] = (\lambda_1 v_1 \quad \dots \quad \lambda_n v_n)$$

$$[\lambda_1, \dots, \lambda_n]B = \begin{pmatrix} \lambda_1 w_1 \\ \dots \\ \lambda_n w_m \end{pmatrix}$$

(ii) Sei  $i \neq j$ ,  $\lambda \in k$ . Dann ist

$$AE_{ij}(\lambda) = (v_1 \quad \dots \quad v_j + \lambda v_i \quad \dots \quad v_n)$$

$$E_{ij}(\lambda)B = \begin{pmatrix} w_1 \\ \dots \\ w_i + \lambda w_j \\ \dots \\ w_n \end{pmatrix}$$

Zur  $i$ -ten Spalte von  $A$  wird das  $\lambda$ -Fache der  $j$ -ten addiert. Zur  $j$ -ten Zeile von  $B$  wird das  $\lambda$ -Fache der  $i$ -ten addiert.

(iii) Sei  $i \neq j$ . Dann entsteht  $AP_{ij}$  aus  $A$  durch Vertauschen der Spalten  $i$  und  $j$ . Ebenso entsteht  $P_{ij}B$  aus  $B$  durch Vertauschen der Zeilen  $i$  und  $j$ .

*Beweis:* Ausmultiplizieren. □

Jede invertierbare Matrix (also voller Rang!) kann mit Hilfe des Gauß-Algorithmus mit elementaren Zeilen- und Spaltentransformationen in die Einheitsmatrix überführt werden. Mit anderen Worten:

**Satz 8.9.** Die  $GL_n(k)$  wird von den Elementarmatrizen und Diagonalmatrizen erzeugt. Genauer:

Sei  $A \in GL_n(k)$ . Dann gibt es endlich viele Matrizen  $S_1, \dots, S_N$ , die entweder invertierbare Diagonalmatrizen oder Elementarmatrizen sind, so dass

$$A = S_1 \dots S_N$$

Dabei genügt es, eines der  $S_i$  als Diagonalmatrix zu wählen.

*Beweis:* Wir vereinfachen  $A$  mit dem Gauß-Algorithmus. Wir wollen dabei jedoch ohne Vertauschungen oder Multiplikation von Zeilen oder Spalten auskommen. Dies entspricht der Multiplikation mit Elementarmatrizen von rechts oder von links.

**Behauptung.** Durch Multiplikation von  $A$  durch Elementarmatrizen von rechts und links wird die Normalform  $[\lambda_1, \dots, \lambda_n]$  für geeignete  $\lambda_i \in k^*$  erreicht.

Wir verfahren wie folgt: Ist  $a_{11} = 0$ , so addieren wir eine andere Zeile zur ersten, so dass dieser Eintrag ungleich 0 ist. Dies ist möglich, da  $A$  vollen Rang hat. Die erste Spalte kann nicht der Nullvektor sein. Danach addieren wir geeignete Vielfache der ersten Zeile zu allen anderen und erreichen so, dass  $a_{i1} = 0$  wird für

$i > 1$ . Wir addieren Vielfache der ersten Spalte zu allen anderen und erreichen so, dass  $a_{1j} = 0$  wird für  $j > 1$ . Danach betrachten wir  $a_{22}$ . Ist  $a_{22} = 0$ , so gibt es ein  $a_{i2} \neq 0$ , wobei  $i > 2$ . Wir addieren diese Zeile zur zweiten, etc. Damit wird die Matrix in Diagonalgestalt gebracht.

Es gilt also

$$[\lambda_1, \dots, \lambda_n] = S_1 \dots S_N A T_1 \dots T_M$$

mit Elementarmatrizen  $S_i, T_j$ . Wir multiplizieren von links mit dem Inversen von  $S_1$ , dann  $S_2$  etc. Danach multiplizieren wir von rechts mit dem Inversen von  $T_M$ , dann  $T_{M-1}$  etc. Wir erhalten die Gleichheit

$$S_N^{-1} \dots S_1^{-1} [\lambda_1, \dots, \lambda_n] T_M^{-1} \dots T_1^{-1} = A$$

Diese ist die gewünschte Form. □

Wir verstehen die Determinanten unserer Erzeuger:

- (i)  $\det(E_{ij}(\lambda)) = 1$ ;
- (ii)  $\det([\lambda_1, \dots, \lambda_n]) = \lambda_1 \dots \lambda_n$ ;
- (iii)  $\det(P_{ij}) = -1$ .

Wir können also die Determinante von  $A$  leicht ausrechnen, wenn wir sie wie im Satz schreiben. Genau dies ist der Algorithmus, den wir im Determinantenkapitel kennengelernt haben.

## Abbildungen zwischen Gruppen

**Definition 8.10.** (i) Seien  $G, H$  Gruppen. Eine Abbildung  $f : G \rightarrow H$  heißt Gruppenhomomorphismus, wenn für alle  $g, g' \in G$  gilt

$$f(gg') = f(g)f(g').$$

(ii) Ein bijektiver Gruppenhomomorphismus heißt Isomorphismus (von Gruppen).

(iii) Sei  $f : G \rightarrow H$  ein Gruppenhomomorphismus. Der Kern von  $f$  ist

$$\text{Ker}(f) = \{g \in G \mid f(g) = e\}.$$

(iv) Sei  $G$  eine Gruppe. Eine Teilmenge  $H \subset G$  heißt Untergruppe, wenn sie bezüglich der Einschränkung der Gruppenmultiplikation auf  $G$  zu einer Gruppe wird. Insbesondere liegen Produkte von Elementen aus  $H$  wieder in  $H$ .

**Lemma 8.11.** Sei  $f : G \rightarrow H$  ein Gruppenhomomorphismus.

- (i) Dann gilt  $f(e) = e$ ,  $f(g^{-1}) = f(g)^{-1}$ .

(ii)  $\text{Ker}(f) \subset G$  und  $\text{Im}(f) \subset H$  sind Untergruppen.

(iii)  $\text{Ker}(f) = \{e\}$  genau dann, wenn  $f$  injektiv ist.

*Beweis:* Wie für lineare Abbildungen:

$$f(e) = f(e^2) = f(e)f(e)$$

Wir multiplizieren mit dem Inversen von  $f(e)$  und erhalten

$$e = f(e).$$

Genauso:

$$e = f(e) = f(gg^{-1}) = f(g)f(g^{-1}) \Rightarrow f(g)^{-1} = f(g^{-1}).$$

Wie im Fall von linearen Abbildungen überprüft man, dass  $\text{Ker}(f)$  und  $\text{Im}(f)$  abgeschlossen sind unter Multiplikation und Inversenbildung. Die Axiome gelten automatisch.

Seien  $g, g' \in G$  mit  $f(g) = f(g')$ . Dann folgt

$$f(g^{-1}g') = f(g)^{-1}f(g') = e \Rightarrow g^{-1}g' \in \text{Ker}(f).$$

Ist  $\text{Ker}(f) = \{e\}$ , so folgt  $g^{-1}g' = e$ , also  $g = g'$ . Die Abbildung ist injektiv. Für die Rückrichtung betrachten wir das Urbild von  $e$ .  $\square$

Wir kennen schon eine ganze Reihe von Gruppenhomomorphismen.

**Beispiel.** (i) Die Abbildung

$$E : S_n \rightarrow \text{Gl}_n(k); \quad \sigma \mapsto E_\sigma$$

ist nach Lemma 6.4 ein Gruppenhomomorphismus. Er ist injektiv.

(ii)  $\det : \text{Gl}_n(k) \rightarrow k^*$  ist ein Gruppenhomomorphismus nach der Produktformel Theorem 6.8.

(iii) Sei  $V$  endlich dimensionaler Vektorraum mit Basis  $A$ . Die Zuordnung  $\text{Aut}_k(V) \rightarrow M_n(K)$ , die einem Automorphismus seine darstellende Matrix zuordnet ist nach Satz 5.4 ein Gruppenhomomorphismus. Da es eine Bijektion ist, haben wir also einen Isomorphismus.

**Definition 8.12.** Der Kern von  $\det : \text{Gl}_n(k) \rightarrow k^*$  heißt spezielle lineare Gruppe  $\text{Sl}_n(k)$ .

Ein besonders wichtiges Beispiel eines Gruppenhomomorphismus:

**Satz 8.13.**  $\text{sgn} : S_n \rightarrow \{\pm 1\}$  ist ein Gruppenhomomorphismus, wobei  $\{\pm 1\}$  mit der Multiplikation in  $\mathbb{Q}$  zu einer Gruppe wird. In anderen Worten: für  $\sigma, \sigma' \in S_n$  gilt

$$\text{sgn}(\sigma\sigma') = \text{sgn}(\sigma) \text{sgn}(\sigma').$$

*Beweis:* Nach Definition ist  $\text{sgn} = \det \circ E$ . Man sieht leicht, dass die Komposition von Gruppenhomomorphismen ein Gruppenhomomorphismus wird.  $\square$

Wir nennen eine Permutation *gerade*, wenn  $\text{sgn}(\sigma) = 1$ , *ungerade* andernfalls.

**Korollar 8.14.** *Eine Permutation ist gerade bzw. ungerade, wenn sie Produkt einer geraden bzw. ungeraden Anzahl von Transpositionen ist.*

*Beweis:* Da wir bereits wissen, dass  $\text{sgn}$  ein Gruppenhomomorphismus ist, reicht es zu zeigen, dass  $\text{sgn}(ij) = -1$ . Die Matrix  $E_{(ij)}$  ist die Permutationsmatrix, die aus  $E_n$  durch Vertauschen der Spalten  $i$  und  $j$  entsteht. Also ist ihre Determinante  $-1$ .  $\square$

**Bemerkung.** Die Wohldefiniertheit des Vorzeichens einer Permutation kann man auf verschiedene Weisen beweisen, auch ohne über Determinanten zu sprechen. In unserem Kontext ging es so besonders schnell und einfach.

**Beispiel.** (i) Sei  $\mathbb{G}_m(k) = (k^*, \cdot)$  die multiplikative Gruppe von  $k$ . Dann ist

$$f : \mathbb{G}_m(k) \rightarrow \text{GL}_n(k) \quad \lambda \mapsto [\lambda, 1, \dots, 1]$$

ein Gruppenhomomorphismus. (Ebenso natürlich für  $\lambda$  in der Spalte  $i \neq 1$ )

(ii) Sei  $\mathbb{G}_a(k) = (k, +)$  die additive Gruppe von  $k$ . Dann ist für jedes  $i \neq j$

$$f : \mathbb{G}_a(k) \rightarrow \text{GL}_n(k) \quad \lambda \mapsto E_{ij}(\lambda)$$

ein Gruppenhomomorphismus. Wir rechnen dies für  $n = 2, i = j = 1$  nach:

$$f(\lambda)f(\mu) = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \mu \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \lambda + \mu \\ 0 & 1 \end{pmatrix} = E_{ij}(\lambda + \mu) = f(\lambda + \mu)$$

Der allgemeine Fall geht genauso.

## Exkurs

**Theorem 8.15.**  $\text{SL}_n(k)$  wird erzeugt von den Elementarmatrizen.

*Beweis:* Die Elementarmatrizen liegen in  $\text{SL}_n(k)$ . Sei nun  $A \in \text{SL}_n(k)$  beliebig. Nach Satz 8.9 ist  $A$  von der Form  $S_1 S_2 \dots S_N$ , wobei für ein  $i$  die Matrix  $L = S_i = [\lambda_1, \dots, \lambda_n]$  und für alle  $j \neq i$  die Matrix  $S_j$  eine Elementarmatrix ist. Wegen  $\det(S_j)$  für die Elementarmatrizen und  $\det(A) = 1$ , folgt

$$\det(L) = \lambda_1 \dots \lambda_n = 1$$

Es bleibt zu zeigen, dass  $L$  Produkt von Elementarmatrizen ist. Sei  $\mu_i = \lambda_1 \lambda_2 \dots \lambda_i$ . Dann ist

$$\mu_i \mu_{i-1}^{-1} = \lambda_i \quad \mu_n = 1 \Rightarrow \mu_{n-1}^{-1} = \lambda_n$$

Damit zerlegen wir

$$L = [\mu_1, \mu_1^{-1}, 1, \dots, 1][1, \mu_2, \mu_2^{-1}, 1, \dots, 1] \dots [1, \dots, 1, \mu_{n-1}, \mu_{n-1}^{-1}]$$

Ohne Einschränkung hat  $L$  also die Form  $[1, \dots, 1, \lambda, \lambda^{-1}, 1, \dots, 1]$ . Wir betrachten nur den Fall  $n = 2$ , der allgemeine geht genauso.

$$L = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}$$

Wir multiplizieren  $L$  von rechts und links mit Elementarmatrizen, d.h. wir addieren Vielfache einer Zeile/Spalte zur anderen. Durch Addition des  $\lambda$ -Fachen der rechten Spalte zur linken erhalten wir:

$$\begin{pmatrix} \lambda & 0 \\ 1 & \lambda^{-1} \end{pmatrix}$$

Durch Addition des  $(1 - \lambda)$ -Fachen der zweiten Zeile zur ersten erhalten wir:

$$\begin{pmatrix} \lambda + 1 - \lambda & \lambda^{-1}(1 - \lambda) \\ 1 & \lambda^{-1} \end{pmatrix} = \begin{pmatrix} 1 & \lambda^{-1}(1 - \lambda) \\ 1 & \lambda^{-1} \end{pmatrix}$$

Durch Subtraktion des  $\lambda^{-1}(1 - \lambda)$ -Fachen der linken Spalte von der rechten erhalten wir:

$$\begin{pmatrix} 1 & 0 \\ 1 & \lambda^{-1} - \lambda^{-1}(1 - \lambda) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

Durch Subtraktion der rechten Spalte von der linken erhalten wir

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

□

**Bemerkung.** Für jeden kommutativen Ring  $R$  mit 1 betrachtet man die Gruppe  $\text{GL}(R)$  der invertierbaren Matrizen  $(a_{ij})_{i,j=1}^{\infty, \infty}$  mit  $a_{ij} \in R$ , für die es  $N$  gibt mit  $a_{ii} = 1$  für  $i > N$ ,  $a_{ij} = 0$  für  $i \neq j$  und  $i > N$  oder  $j > N$ . Die Elemente haben also die Form

$$\begin{pmatrix} A & 0 \\ 0 & E_{\infty} \end{pmatrix}$$

wobei  $E_{\infty}$  die Diagonalmatrix mit allen Einträgen 1 ist und  $A \in \text{GL}_N(R)$  eine invertierbare  $N \times N$ -Matrix. Hierin gibt es die Untergruppe  $E(R)$  der Matrizen, die von Elementarmatrizen erzeugt werden. Dann heißt

$$K_1(R) = \text{GL}(R)/E(R)$$

erste algebraische  $K$ -Gruppe von  $R$ . Es ist eine wichtige Invariante des Rings. Das Theorem besagt nun, dass

$$K_1(k) = k^*$$



## Kapitel 9

# Etwas affine und projektive Geometrie

Affine Geometrie ist Geometrie ohne Abstandsbegriff. Es geht also um Strukturen, die unter Translationen, Scherungen usw. erhalten bleiben. Man kann dies rein axiomatisch angehen, wie Euklid es tat. Wir arbeiten von vornherein mit Koordinaten. Man spricht dann auch von analytischer Geometrie.

### Affine Geometrie der Ebene

Sei  $k$  ein Körper. Wir nennen jetzt  $\mathbb{A}^2 = k^2$  Ebene und die Elemente Punkte. Wir wollen dabei zwischen der Menge ( $\mathbb{A}^2$ ) und dem Vektorraum ( $k^2$ ) unterscheiden. Für  $P, Q \in \mathbb{A}^2$  schreiben wir  $\overrightarrow{PQ} \in k^2$  für den Vektor von  $P$  nach  $Q$ , also

$$\overrightarrow{PQ} = Q - P$$

**Definition 9.1.** Eine affine Gerade in der Ebene ist eine Teilmenge von  $\mathbb{A}^2$  von der Form

$$L = P_0 + V$$

wobei  $P_0 \in \mathbb{A}^2$  ein Punkt und  $V \subset k^2$  ein Untervektorraum der Dimension 1.

**Lemma 9.2.** Sei  $L \subset \mathbb{A}^2$  eine affine Gerade. Dann gilt für jedes  $P \in L$

$$L = P + V$$

wobei  $V = \{\overrightarrow{RS} \mid R, S \in L\}$ . Insbesondere ist  $V$  unabhängig von  $P$ .

Die Elemente von  $V$  heißen Richtungsvektoren von  $L$ .

*Beweis:* Nach Definition ist  $L = P_0 + V$  für einen eindimensionalen Untervektorraum  $V \subset k^2$ . Seien  $R, S \in L$ . Dann gibt es  $v, w \in V$  mit  $R = P_0 + v, S = P_0 + w$ . Es folgt

$$\overrightarrow{RS} = S - R = (P_0 + w) - (P_0 + v) = w - v \in V$$

Umgekehrt hat jedes  $v \in V$  die Form  $\overrightarrow{P_0S}$  mit  $S = P_0 + v$ . Dies beweist die intrinsische Charakterisierung von  $V$  als Raum der Richtungsvektoren.

Sei nun  $P \in L$  beliebig. Wir zeigen  $P + V = P_0 + V$ . Nach Voraussetzung ist  $P = P_0 + v$  für ein  $v \in V$ . Ein beliebiges Element  $Q$  von  $P_0 + V$  hat die Form  $P_0 + w$ . Es folgt

$$Q = P_0 + w = P_0 + v + (w - v) = P + (w - v)$$

Wegen  $w - v \in V$  folgt  $Q \in P + V$ . Die andere Inklusion sieht man genauso.  $\square$

**Lemma 9.3.** *Seien  $P, Q \in \mathbb{A}^2$ ,  $P \neq Q$ . Dann gibt es eine eindeutige affine Gerade, die  $P$  und  $Q$  enthält.*

*Beweis:* Sei  $V = \langle \overrightarrow{PQ} \rangle$ . Dann ist  $L = P + V$  die gesuchte Gerade. Ist umgekehrt  $L$  eine Gerade durch  $P, Q$ , so enthält der Raum der Richtungsvektoren  $\overrightarrow{PQ}$ , also auch  $V$ . Wegen der Eindimensionalität ist er gleich  $V$ .  $\square$

**Definition 9.4.** *Zwei affine Geraden heißen parallel, wenn ihre Räume von Richtungsvektoren übereinstimmen.*

**Satz 9.5** (Parallelenaxiom). *Sei  $L$  eine affine Gerade,  $P \in \mathbb{A}^2$  ein Punkt. Dann gibt es genau eine zu  $L$  parallele Gerade durch  $P$ .*

*Beweis:* Sei  $L = P_0 + V$ . Die gesuchte Gerade ist  $P + V$ . Offensichtlich ist dies eindeutig.  $\square$

**Lemma 9.6.** *Zwei nicht-parallele Gerade schneiden sich in genau einem Punkt.*

*Beweis:* Sei  $L = P + V$  und  $L' = P' + V'$  mit  $V \neq V'$ . Es ist

$$V \cap V' \subset V, V' \subset V + V'$$

Wegen  $V \neq V'$  gilt an beiden Stellen Ungleichheit. Aus Dimensionsgründen folgt  $V + V' = k^2$  und  $V \cap V' = 0$ . Wir betrachten  $P - P'$ . Es lässt sich nun schreiben als  $v - v'$  mit  $v \in V, v' \in V'$ . Dann ist  $P + v = P' + v'$  der gesuchte Schnittpunkt.

Seien umgekehrt  $Q, Q'$  Schnittpunkte. Dann liegt  $\overrightarrow{QQ'}$  in  $V$  und in  $V'$ . Wegen  $V \cap V' = 0$  gilt  $\overrightarrow{QQ'} = 0$ , also  $Q = Q'$ .  $\square$

Viele Sätze der Schulgeometrie lassen sich mit den Mitteln der affinen Geometrie beweisen.

**Satz 9.7.** *Sei  $0 \neq 2, 3$  in  $k$ . Dann gilt: Die Seitenhalbierenden eines Dreiecks schneiden sich in einem Punkt.*

*Beweis:* Seien  $A, B, C$  die Ecken des Dreiecks. Sei  $a = \overrightarrow{BC}$ ,  $b = \overrightarrow{CA}$  und damit  $a + b = \overrightarrow{BA}$ . Wir setzen voraus, dass es sich um ein echtes Dreieck handelt, also  $a$  und  $b$  linear unabhängig sind.

Die Seitenmittelpunkte sind

$$P_a = B + \frac{1}{2}a, P_b = C + \frac{1}{2}b, P_c = B + \frac{1}{2}(a + b)$$

Die Gerade durch  $A$  und  $P_a$  (bzw.  $B$  und  $P_b$ ,  $C$  und  $P_c$ ) sind

$$L_a = A + \langle \overrightarrow{AP_a} \rangle = A + \langle -\frac{1}{2}a - b \rangle$$

$$L_b = B + \langle \overrightarrow{BP_b} \rangle = B + \langle a + \frac{1}{2}b \rangle$$

$$L_c = C + \langle \overrightarrow{CP_c} \rangle = C + \langle \frac{1}{2}(b - a) \rangle$$

Wir betrachten  $M = A - \frac{1}{3}a - \frac{2}{3}b$ . Es gilt

$$\begin{aligned} M &= A - \frac{2}{3}\left(\frac{1}{2}a + b\right) \\ &= B + a + b - \frac{1}{3}a - \frac{2}{3}b = B + \frac{2}{3}a + \frac{1}{3}b = B + \frac{2}{3}\left(a + \frac{1}{2}b\right) \\ &= C + b - \frac{1}{3}a - \frac{2}{3}b = C - \frac{1}{3}a + \frac{1}{3}b = C + \frac{1}{3}(b - a) \end{aligned}$$

Dies ist also der gemeinsame Schnittpunkt der Seitenhalbierenden.  $\square$

**Bemerkung.** Der Satz über den Schnittpunkt der Winkelhalbierenden oder der Höhen lässt sich nicht mit diesen Methoden beweisen - wir haben noch keinen Winkelbegriff eingeführt.

## Affine Räume

Zu einer affinen Gerade gehört also eine Punktmenge und ein eindimensionaler Raum von Richtungsvektoren. Dies lässt sich abstrahieren.

**Definition 9.8.** Sei  $k$  Körper. Ein affiner Raum über  $k$  der Dimension  $n$  ist ein Paar  $(A, V)$  wobei  $A$  eine Menge ist und  $V$  ein  $k$ -Vektorraum der Dimension  $n$ , zusammen mit einer Abbildung

$$A \times A \rightarrow V \quad (P, Q) \mapsto \overrightarrow{PQ}$$

für die gilt:

(i)  $\overrightarrow{PP} = 0$  für alle  $P \in A$ .

(ii)  $\overrightarrow{PQ} = 0$  für  $P, Q \in A \Rightarrow P = Q$ .

(iii)  $\overrightarrow{PQ} = \overrightarrow{PS} + \overrightarrow{SQ}$  für alle  $P, Q, S \in A$ .

(iv) Für jedes  $v \in V$  und jedes  $P \in A$  gibt es  $Q \in A$  mit  $v = \overrightarrow{PQ}$ .

Affine Räume der Dimensionen 0, 1, 2 heißen Punkte, Geraden, Ebenen.

**Beispiel.** (i) Jeder Vektorraum kann als affiner Raum aufgefasst werden, wobei  $A = V$  mit  $\overrightarrow{PQ} = Q - P$ . Speziell für  $V = k^n$  nennen wir diesen affinen Raum  $\mathbb{A}^n$ .

(ii) Sei  $M \in M_{m \times n}(k)$ ,  $b \in k^m$ . Dann ist nach Lemma 1.14 die Menge  $A = \{x \in k^n \mid Ax = b\}$  (falls nicht leer) ein affiner Raum zum Vektorraum  $V = \{x \in k^n \mid Ax = 0\}$ .

**Bemerkung.** Es ist  $\overrightarrow{PQ} = -\overrightarrow{QP}$ , denn nach Axiom (iii) und (i) gilt

$$\overrightarrow{PQ} + \overrightarrow{QP} = \overrightarrow{PP} = 0$$

**Lemma 9.9.** Sei  $(A, V)$  ein affiner Raum,  $P_0 \in A$ . Dann ist

$$A \rightarrow V \quad P \mapsto \overrightarrow{P_0P}$$

bijektiv.

Man sagt: Ein affiner Raum ist ein Vektorraum ohne Ursprung.

*Beweis:* Die Injektivität ist Eigenschaft (ii). Die Surjektivität ist (iv). □

**Bemerkung.** Ist  $(A, V)$  ein affiner Raum, so erhalten wir eine Abbildung

$$V \times A \rightarrow A$$

wobei  $(v, P)$  auf den eindeutigen Punkt  $Q$  mit  $v = \overrightarrow{PQ}$  abgebildet wird. Wir schreiben  $Q = P + v$ . Alternativ kann man  $(A, V)$  auch definieren als ein Paar von Menge und Vektorraum und einer Abbildung

$$+ : V \times A \rightarrow A \quad (v, P) \mapsto P + v$$

mit geeigneten Axiomen. Dieser Standpunkt ordnet sich besser in die allgemeine Theorie ein. Es handelt sich um einfach transitive Operation von  $(V, +)$  auf der Menge  $A$ .

**Definition 9.10.** Sei  $(A, V)$  ein affiner Raum. Ein affiner Teilraum ist ein Paar  $(A', V')$  wobei  $A' \subset A$  Teilmenge und  $V' \subset V$  ein Untervektorraum, so dass  $(A', V')$  mit der Restriktion der Strukturabbildung von  $(A, V)$  ein affiner Raum wird.

Dies bedeutet  $\overrightarrow{PQ} \in V'$  für alle  $P, Q \in A'$  und Axiom (iv) ist für  $(A', V')$  erfüllt.

**Beispiel.** Eine affine Gerade in  $\mathbb{A}^2$  ist ein eindimensionaler Unterraum des affinen Raums  $k^2$  mit der Standardstruktur.

**Lemma 9.11** (Parameterdarstellung von affinen Räumen). Sei  $(A, V) \subset \mathbb{A}^n$  ein affiner Teilraum der Dimension  $d$ ,  $P \in A$ .

- $A$  die Form  $P + V \subset \mathbb{A}^n$  für ein  $P \in A$ .
- Es gibt Vektoren  $v_1, \dots, v_d \in k^n$ , so dass

$$A = \left\{ P + \sum_{i=1}^d \lambda_i v_i \mid \lambda_1, \dots, \lambda_d \in k \right\}.$$

*Beweis:* Für jedes  $Q \in A$  ist  $v = \overrightarrow{PQ} \in V$ . Da es ein Unterraum von  $\mathbb{A}^n$  ist, bedeutet dies  $Q = P + v$  mit  $v \in V$ . Umgekehrt liegt jeder Punkt des  $\mathbb{A}^n$  von dieser Form in  $A$ .

Sei  $v_1, \dots, v_d$  eine Basis von  $V$ . Dann hat jedes  $v \in V$  eindeutig die Form  $v = \sum \lambda_i v_i$ .  $\square$

Neben dieser expliziten Charakterisierung gibt es auch eine implizite als Lösung eines Gleichungssystems.

**Lemma 9.12** (Implizite Gleichung für affine Teilräume). *Sei  $(A, V) \subset \mathbb{A}^n$  affiner Teilraum der Dimension  $d$ . Dann gibt es  $B \in M_{(n-d) \times n}(k)$  und  $b \in k^{n-d}$ , so dass*

$$A = \{x \in k^n \mid Bx = b\}.$$

*Beweis:* Sei  $P \in A$ . Wenn  $P \in V$ , setzen wir  $b$ , andernfalls  $b = P$ . Wir wollen eine lineare Abbildung  $\beta : k^n \rightarrow k^n$  definieren mit  $\text{Ker}(\beta) = V$ . Sei  $v_1, \dots, v_d$  eine Basis von  $V$ . Wir setzen sie durch  $v_{d+1}, \dots, v_n$  zur einer Basis von  $k^n$  fort. Sei  $w_1, \dots, w_{n-d}$  eine Basis von  $k^{n-d}$ . In dieser Basis definieren wir die lineare Abbildung

$$\beta : v_i \mapsto \begin{cases} 0 & i \leq d \\ w_{i-d} & i > d \end{cases}$$

Sie hat die gewünschten Eigenschaft.

Sei  $B$  eine darstellende Matrix von  $\beta$  in der Standardbasis und  $b' = \beta(P)$ . Nach Konstruktion

$$\{x \in k^n \mid Bx = 0\} = \text{Ker}(\beta) = V$$

und es gibt  $x_0 \in k^n$  mit  $Bx_0 = b$ . Die Lösungsmenge des inhomogenen Systems ist genau  $A$ .  $\square$

**Beispiel.** Eine Hyperebene (d.h.  $\dim A = n - 1$ ) ist gegeben durch eine einzige Gleichung

$$a_1 x_1 + a_2 x_2 + \dots + a_n x_n = b.$$

Für  $n = 2$  erhält man eine Gerade in der Ebene als

$$L = \{(x, y) \in k^2 \mid ax + by = c\}$$

für  $a, b, c \in k$ ,  $(a, b) \neq (0, 0)$ .

**Korollar 9.13.** *Ein affiner Teilraum des  $\mathbb{A}^n$  der Dimension  $d$  ist der Schnitt von  $n - d$  Hyperebenen.*

*Beweis:* Jede der  $n - d$  Gleichungen beschreibt eine Hyperebene.  $\square$

**Definition 9.14.** Sei  $(A, V)$  ein affiner Raum. Zwei affine Teilräume  $(A_1, V_1)$  und  $(A_2, V_2)$  heißen parallel, falls  $V_1 \subset V_2$  oder  $V_2 \subset V_1$ .

**Lemma 9.15.** Parallele Teilräume sind disjunkt oder einer ist im anderen enthalten.

*Beweis:* Sei  $P \in A_1 \cap A_2$ . Ohne Einschränkung ist  $V_1 \subset V_2$ . Dann ist jedes Element von  $A_1$  von der Form  $P + v$  für  $v \in V_1$ . Wegen  $v \in V_2$  ist dann  $P + v \in A_2$ .  $\square$

**Definition 9.16.** Sei  $(A, V)$  ein affiner Raum,  $(A_1, V_1)$  und  $(A_2, V_2)$  affine Teilräume. Sei  $A_1 + A_2$  der von  $A_1, A_2$  aufgespannte affine Teilraum, also der kleinste affine Teilraum von  $A$ , der  $A_1$  und  $A_2$  enthält.

Der Raum der Richtungsvektoren von  $A_1 + A_2$  wird aufgespannt von den  $\overrightarrow{PQ}$  mit  $P, Q \in A_1 \cup A_2$ . Er enthält also  $V_1 + V_2$ , ist im allgemeinen aber größer.

**Satz 9.17** (Dimensionsformel). Sei  $(A, V)$  ein affiner Raum,  $(A_1, V_1)$  und  $(A_2, V_2)$  affine Teilräume. Dann gilt:

$$\dim A_1 + A_2 = \dim A_1 + \dim A_2 + \begin{cases} -\dim A_1 \cap A_2 & A_1 \cap A_2 \neq \emptyset \\ -\dim V_1 \cap V_2 + 1 & A_1 \cap A_2 = \emptyset \end{cases}$$

*Beweis:* Wir behandeln zunächst den Fall  $A_1 \cap A_2 \neq \emptyset$ . Sei  $P$  ein Schnittpunkt. Dann gilt

$$A_1 = P + V_1, A_2 = P + V_2, A_1 \cap A_2 = P + V_1 \cap V_2, A_1 + A_2 = P + (V_1 + V_2)$$

denn  $P + (V_1 + V_2)$  ist ein affiner Teilraum, der  $A_1, A_2$  enthält und alle diese Punkte sind in  $A_1 + A_2$  enthalten. Die Dimensionsformel für affine Räume folgt also aus der Dimensionsformel für Vektorräume. Im Detail: Sei  $v_1, \dots, v_n$  eine Basis von  $V_1 \cap V_2$ . Diese ergänzen wir durch  $w_1, \dots, w_m$  zu einer Basis von  $V_1$  und durch  $u_1, \dots, u_l$  zu einer Basis von  $V_2$ . Dann ist  $v_1, \dots, v_n, w_1, \dots, w_m, u_1, \dots, u_l$  eine Basis von  $V_1 + V_2$ .

Nun behandeln wir den Fall  $A_1 \cap A_2 = \emptyset$ . Sei  $A_1 = P_1 + V_1, A_2 = P_2 + V_2, v = \overrightarrow{P_1 P_2}$ . Dann gilt

$$A_1 + A_2 = P + (V_1 + V_2 + \langle v \rangle)$$

denn die rechte Seite enthält  $P, Q$  und dann auch ganz  $A_1$  und  $A_2$ . Andererseits liegen alle Elemente der rechten Seite offensichtlich in  $A_1 + A_2$ .

**Behauptung.**  $v \notin V_1 + V_2$

Angenommen,  $v = v_1 + v_2$  mit  $v_i \in V_i$ . Dann liegt  $P + v_1 = P + v - v_2 = Q - v_2$  in  $A_1 \cap A_2$ . Widerspruch.

Wir lesen ab:

$$\dim A_1 + A_2 = \dim(V_1 + V_2) + 1$$

Die Formel für die Dimension der affinen Räume folgt wieder aus der Formel für die Dimension der Vektorräume.  $\square$

**Beispiel.** Sei  $\dim A = 2$ ,  $\dim V_i = 1$ . Wir betrachten also ebene Geraden. Dann ist  $\dim A_1 + A_2 = 1, 2$ .

Ist der aufgespannte Raum eindimensional, so sind die beiden Geraden gleich. Auch der Schnitt ist eindimensional.

Sei nun  $\dim A_1 + A_2 = 2$ .

1. Fall:  $A_1 \cap A_2 \neq \emptyset$ . Die Formel lautet:

$$2 = 1 + 1 - \dim A_1 \cap A_2$$

Der Schnitt ist ein affiner Raum der Dimension 0, also ein Punkt.

2. Fall:  $A_1 \cap A_2 = \emptyset$ . Die Formel lautet

$$2 = 1 + 1 - \dim V_1 \cap V_2 + 1$$

Hierausfolgt  $\dim V_1 \cap V_2 = 1$ , also  $V_1 = V_2$ . Die Geraden sind parallel.

Sei nun  $\dim A = 3$ ,  $\dim V_i = 1$ . Wenn die Geraden nicht in einer Ebene liegen, also  $\dim A_1 + A_2 = 3$ , so muss der Schnitt leer sein:

$$3 = 1 + 1 - 0 + 1$$

Dies ist der Fall von windschiefen Geraden im Raum.

## Die projektive Raum

In der affinen Ebene gibt es Geraden, die sich nicht schneiden: Parallelen. Höherdimensional führt dies zu einer häßlichen Fallunterscheidung in der Dimensionsformel. Wir gehen daher zur *projektiven Ebene* über. Sie enthält für jede Schar von parallelen Geraden (also für jeden eindimensionalen Unterraum von  $k^2$ ) einen weiteren Punkt, in dem sich die Parallelen schneiden. Dies ist der *unendlich ferne Punkt* zur gegebenen Richtung. Alle unendlich fernen Punkte liegen auf einer Geraden: der *unendlich fernen Geraden*. Also:

$$\mathbb{P}^2 = \mathbb{A}^2 \cup \{V \subset k^2 \mid \dim(V) = 1\}.$$

Wir definieren eleganter.

**Definition 9.18.** Sei  $k$  ein Körper,  $V$  ein  $k$ -Vektorraum der Dimension  $n + 1$ . Der projektive Raum  $\mathbb{P}(V)$  über  $k$  ist die Menge der Nullpunktsgersten in  $V$ .

Wir setzen  $\dim \mathbb{P}(V) = \dim V - 1$ .

Speziell für  $V = k^{n+1}$  schreiben wir

$$\mathbb{P}^n = \mathbb{P}(k^{n+1})$$

der  $n$ -dimensionale projektive Raum über  $k$ .

**Bemerkung.** In der algebraischen Geometrie schreibt man meist  $\mathbb{P}^n(k)$  oder  $\mathbb{P}_k^n$ . In der Differentialgeometrie  $\mathbb{R}P^n$  ( $k = \mathbb{R}$ ) und  $\mathbb{C}P^n$  ( $k = \mathbb{C}$ ).

Sei  $x = (x_0, \dots, x_n) \in k^{n+1} \setminus \{0\}$ . Dann definieren 0 und  $x$  eine eindeutige Gerade in  $k^{n+1}$ , also einen Punkt in  $\mathbb{P}^n$ . Wir schreiben  $[x] = [x_0 : \dots : x_n]$  für diesen Punkt. Es gilt

$$[x_0 : \dots : x_n] = [y_0 : \dots : y_n] \Leftrightarrow \exists \lambda \in k^* : x_i = \lambda_i \forall i.$$

Dies ist eine Äquivalenzrelation auf  $k^{n+1} \setminus \{0\}$ .

**Beispiel.** (i)  $n = 0$ : Es gibt genau eine Nullpunktsgerade in  $k^1$ , also hat  $\mathbb{P}^0$  ein einziges Element. Nulldimensionale projektive Räume sind einfach Punkte.

(ii)  $n = 1$ . Die Elemente von  $\mathbb{P}^1$  sind repräsentiert durch  $[x_0 : x_1]$  mit  $(x_0, x_1) \in k^2 \setminus \{(0, 0)\}$ . Ist  $x_0 \neq 0$ , so ist  $[x_0 : x_1] = [1 : a]$  mit  $a = x_1/x_0 \in k$ . Diese Punkte stehen also in Bijektion mit einer affinen Geraden. Ist  $x_0 = 0$ , so ist  $x_1 \neq 0$  und  $[x_0 : x_1] = [0 : 1]$ . Dies ist der unendlich ferne Punkt.

**Definition 9.19.** Sei  $0 \leq i \leq n$ . Sei  $U_i = \{[x_0 : \dots : x_n] \in \mathbb{P}^n \mid x_i \neq 0\}$ ,

$$\phi_i : U_i \rightarrow k^n, \quad [x_0 : \dots : x_n] \mapsto \left( \frac{x_0}{x_i}, \dots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \dots, \frac{x_n}{x_i} \right)$$

Wir nennen  $U_i$   $i$ -te standardaffine Teilmenge und  $\phi_i$  die  $i$ -te standardaffine Karte von  $\mathbb{P}^n$ .

**Lemma 9.20.** (i)  $\phi_i$  ist bijektiv mit Umkehrabbildung

$$\psi_i : k^n \rightarrow \mathbb{P}^n, (y_1, \dots, y_n) \mapsto [y_1 : \dots : y_i : 1 : y_{i+1} : \dots : y_n]$$

(ii)  $\mathbb{P}^n \setminus U_i = \mathbb{P}(V_i)$  mit  $V_i = \{(x_0, \dots, x_n) \in k^{n+1} \mid x_i = 0\}$ .

(iii) Es gilt  $\mathbb{P}^n = \bigcup_{i=0}^n U_i$ .

*Beweis:* Offensichtlich hat  $\psi_i$  Werte in  $U_i$  und  $\phi_i \circ \psi_i = \text{id}$ . Umgekehrt ist

$$\begin{aligned} \psi_i \phi_i([x_0 : \dots : x_n]) &= [x_0/x_i, \dots, x_{i-1}/x_i : 1 : x_{i+1}/x_i : \dots : x_n/x_i] \\ &= [x_0 : \dots : x_n]. \end{aligned}$$

Die Elemente im Komplement haben  $x_i = 0$ , wie behauptet. Für jedes  $[x_0 : \dots : x_n] \in \mathbb{P}^n$  gibt es  $i$  mit  $x_i \neq 0$ , also  $x_i \in U_i$ . □

**Bemerkung.** Für  $k = \mathbb{R}$  oder  $\mathbb{C}$  macht dies  $\mathbb{P}^n$  zu einer Mannigfaltigkeit, also zu einem Objekt der Analysis. Mit der offensichtlichen Topologie erhält man eine Kompaktifizierung von  $\mathbb{R}^n$  bzw.  $\mathbb{C}^n$ . (Das ist nicht ganz offensichtlich.)

Ab jetzt fassen wir  $\mathbb{A}^n$  via  $\psi_0$  als Teilmenge von  $\mathbb{P}^n$  auf. Wir nennen

$$H_\infty = \mathbb{P}^n \setminus \mathbb{A}^n = \mathbb{P}(0 \times k^n)$$

die *unendlich ferne Hyperebene*.

Sei  $U \subset V$  ein Untervektorraum. Jede Nullpunktsgerade in  $U$  ist eine Nullpunktsgerade in  $V$ , also  $\mathbb{P}(U) \subset \mathbb{P}(V)$ .



**Definition 9.21.** Für einen Untervektorraum  $U \subset V$  heißt  $\mathbb{P}(U)$  projektiver Teilraum von  $\mathbb{P}(V)$ . Projektive Teilräume der Dimensionen 0, 1, 2,  $n-1$  heißen Punkte bzw. projektive Geraden bzw. projektive Ebenen bzw. Hyperebenen. Seien  $H_1, H_2$  projektive Teilräume. Der von  $H_1$  und  $H_2$  aufgespannte Teilraum ist der kleinste projektive Teilraum, der  $H_1$  und  $H_2$  enthält. Wir schreiben  $H_1 + H_2$ .

**Theorem 9.22** (Dimensionsformel). Seien  $H_1, H_2 \subset \mathbb{P}^n$  projektive Teilräume. Dann gilt

$$\dim H_1 + \dim H_2 = \dim(H_1 + H_2) + \dim H_1 \cap H_2.$$

*Beweis:* Sei  $H_i = \mathbb{P}(U_i)$ . Dann ist  $H_1 \cap H_2 = \mathbb{P}(U_1 \cap U_2)$  und  $H_1 + H_2 = \mathbb{P}(U_1 + U_2)$ . Es gilt

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

und jeweils  $\dim \mathbb{P}(U) = \dim U - 1$ . □

**Beispiel.** Sei  $n = 2$ ,  $H_1 \neq H_2$  projektive Geraden. Es folgt  $\dim(H_1 + H_2) > 1$ , also  $H_1 + H_2 = \mathbb{P}^2$ . Laut Dimensionsformel gilt  $\dim(H_1 \cap H_2) = 0$ , also ist  $H_1 \cap H_2$  ein Punkt. Ungleiche projektive Geraden schneiden sich in genau einem Punkt.

Zum Abschluss wollen wir verstehen, wie die affine und die projektive Situation zusammenhängen.

**Satz 9.23.** Sei  $H \subset \mathbb{A}^n$  ein affiner Teilraum zum Vektorraum  $V$ . Dann gibt es einen eindeutig bestimmten projektiven Unterraum  $\overline{H} \subset \mathbb{P}^n$  mit  $\overline{H} \cap \mathbb{A}^n = H$ . Es gilt  $\overline{H} \cap H_\infty = \mathbb{P}(V)$ .

*Beweis:* Sei  $H = \{x \in k^n \mid Mx = b\}$  für festes  $b \in k^r$ ,  $M \in M_{r \times n}(k)$ . Wir setzen  $M' = (-bM)$ ,  $x' = (x_0, \dots, x_n)^t \in k^{n+1}$  und damit

$$\overline{H} = \{[x'] \mid M'x' = 0\}.$$

Wir berechnen  $\overline{H} \cap \mathbb{A}^n$ , also  $[x'] \in \overline{H} \mid x'_0 \neq 0$ , indem wir auf dem Bild von  $\psi_0$  auswerten. Es gilt

$$(-b \quad M) \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = 0 \Leftrightarrow M \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = b,$$

wie behauptet. Es ist außerdem

$$\overline{H} \cap H_\infty = \{[x'] = [0 : x_1 : \dots : x_n] \mid M'x' = 0\} = \{[0 : x] \mid Mx = 0\} = V.$$

Wir zeigen die Eindeutigkeit von  $\overline{H}$ . Sei  $\mathbb{P}(U) \cap \mathbb{A}^n = H$  für  $U \subset k^{n+1}$ . Die Voraussetzung  $H \subset \mathbb{P}(U)$  bedeutet  $(1, x_1, \dots, x_n)^t \in U$  für alle  $(x_1, \dots, x_n)^t$  mit  $Mx = 0$ . Damit enthält  $U$  alle  $x' \in k^{n+1}$  mit  $M'x' = 0$ . Mit anderen Worten,  $\overline{H} \subset \mathbb{P}(U)$ .

Sei umgekehrt  $(u_0, \dots, u_n)^t \in U$ . Sei zunächst  $u_0 \neq 0$ . Dann ist  $u = [u_0 : \dots : u_n] = [1 : u_1/u_0 : \dots : u_n/u_0] \in \mathbb{P}(U) \cap \mathbb{A}^n = H \subset \overline{H}$ . Insbesondere gilt  $M'u = 0$ .

Es bleibt der Fall  $u_0 = 0$ . Wir wählen einen Hilfspunkt  $x \in U$  so dass  $[x] \in H$  (so ein Punkt existiert, da  $H \neq \emptyset$ ). Dann liegt auch  $[x + u] \in H$ . Beide Punkte erfüllen nach dem ersten Fall die Gleichung

$$M'x = 0, M'(x + u) = 0 \Rightarrow M'u = 0.$$

□

**Korollar 9.24.** *Seien  $L_1, L_2 \subset \mathbb{A}^n$  parallele Geraden. Dann schneiden sich  $\overline{L}_1$  und  $\overline{L}_2$  auf der unendlich fernen Hyperebene. Jeder Punkt der unendlichen fernen Hyperebene entspricht genau einer Menge von parallelen Geraden in  $\mathbb{A}^n$ .*

*Beweis:* Dies ist der Spezialfall  $\dim V = 1$  des Satzes: Da  $L_1$  und  $L_2$  parallel sind, haben sie denselben Vektorraum von Richtungsvektoren und schneiden  $H_\infty$  im Punkt  $\mathbb{P}(V)$ . Jeder Punkt von  $H_\infty$  hat diese Form. □

Damit haben wir unsere erste Definitions-idee der projektiven Ebene verifiziert.

**Lemma 9.25.** *Seien  $[x], [y], [z] \in \mathbb{P}^2$ . Dann sind äquivalent:*

(i) *Die drei Punkte liegen auf einer Geraden.*

(ii)  *$x, y, z$  sind linear abhängig.*

*Beweis:* Die drei Punkte entsprechen Ursprungsgeraden  $kx, ky, kz$  in  $k^3$ . Der von ihnen aufgespannte projektive Teilraum ist  $\mathbb{P}(V)$  mit  $V = \langle x, y, z \rangle$ . Dies ist eine projektive Gerade, wenn  $\dim V = 2$ , also die drei Punkte linear abhängig. □

Wir sagen, die Punkte sind *kollinear*.

**Theorem 9.26** (Satz von Desargues). *Seien zwei Dreiecke  $ABC$  und  $A'B'C'$  in der projektiven Ebene gegeben. Gehen die Verbindungsgeraden der entsprechenden Punkte durch einen gemeinsamen Schnittpunkt, so liegen die Schnittpunkte einander entsprechender Seiten auf einer Geraden.*

*Beweis:* Wir schreiben  $A = [a]$  mit  $a \in k^2 \setminus \{0\}$  etc. Jeder Punkt auf der Geraden durch  $A$  und  $A'$  hat die Form  $\lambda_1 a + \lambda_2 a'$ . Der gemeinsame Schnittpunkt  $S$  hat also die Darstellungen

$$S = [\lambda_1 a + \mu_1 a'] = [\lambda_2 b + \mu_2 b'] = [\lambda_3 c + \mu_3 c'].$$

Wir skalieren die Koeffizienten, so dass Gleichheit gilt:

$$\lambda_1 a + \mu_1 a' = \lambda_2 b + \mu_2 b' = \lambda_3 c + \mu_3 c'.$$

Wir stellen um zu

$$\begin{aligned}\lambda_1 a - \lambda_2 b &= -\mu_1 a' + \mu_2 b' =: r_1 \\ \lambda_1 a - \lambda_3 c &= -\mu_1 a' + \mu_3 c' =: r_2 \\ \lambda_2 b - \lambda_3 c &= -\mu_2 b' + \mu_3 c' =: r_3\end{aligned}$$

Die  $r_i$  sind nicht 0, da sonst  $A = B$  etc. Sei  $R_i = [r_i]$ . Dann gilt wegen der obigen Gleichungen

$$R_1 \in AB \cap A'B', R_2 \in AC \cap A'C', R_3 \in BC \cap B'C'.$$

Dies sind also die drei Schnittpunkte. Es gilt

$$r_1 - r_2 + r_3 = 0$$

also liegen die drei Punkte auf einer Geraden.  $\square$

**Bemerkung.** Der Satz von Desargues wird klar, wenn man die Konstellation als Projektion einer dreidimensionalen Konstellation auffasst, in der die drei Geraden  $AA'$ ,  $BB'$  und  $CC'$  nicht in einer Ebene liegen. Sie schneiden sich in einem Punkt. Die beiden Dreiecke  $ABC$  und  $A'B'C'$  definieren Ebenen. Die Schnittpunkte entsprechender Seiten liegen in der Schnittgeraden der beiden Ebenen.

**Bemerkung.** Im axiomatischen Zugang zur projektiven Geometrie gilt der Satz von Desargues nicht automatisch, sondern genau dann wenn die projektive Ebene von einem Schiefkörper herkommt. Addition und Multiplikation von Punkten auf einer Gerade kann man rein geometrisch definieren. Der Satz von Desargues ist das Assoziativgesetz der Multiplikation.

## Affine Bewegungen

**Definition 9.27.** Seien  $(A_1, V_1)$  und  $(A_2, V_2)$  affine Räume. Eine affine Abbildung  $\phi$  ist eine Abbildung  $\phi_A : A_1 \rightarrow A_2$  zusammen mit einer linearen Abbildung  $\phi_V : V_1 \rightarrow V_2$ , so dass  $\phi_V(\overrightarrow{PQ}) = \overrightarrow{\phi_A(P)\phi_A(Q)}$  für alle  $P, Q \in A$ . Bijektive affine Abbildungen  $A \rightarrow A$  heißen affine Bewegungen.

Eine affine Abbildung ist durch  $\phi_A$  bereits eindeutig bestimmt. Wir schreiben meist kurz  $\phi = \phi_A$ .

**Beispiel.** Sei  $A = V = k$ . Dann ist jede Abbildung der Form  $x \mapsto ax + b$  (für  $a, b \in k$ ) eine affine Abbildung. Die zugehörige lineare Abbildung auf Richtungsvektoren ist  $v \mapsto av$ . Die Abbildung ist genau dann bijektiv, wenn  $a \neq 0$ .

Affine Bewegungen  $\phi : A \rightarrow A$  bilden eine Gruppe. Bereits in Dimension 1 ist diese nichtkommutativ.

**Satz 9.28** (Koordinaten). *Sei  $(A, V)$  ein  $n$ -dimensionaler affiner Raum. Dann gibt es eine bijektive affine Abbildung  $(A, V) \rightarrow (\mathbb{A}^n, k^n)$*

*Beweis:* Sei  $P_0 \in A$  ein Punkt,  $\phi : V \rightarrow k^n$  ein Isomorphismus. Für jedes  $P \in A$  gibt es ein eindeutiges  $v \in V$  mit  $P = P_0 + v$ . Wir setzen  $\phi_A(P) = \phi(v)$ . Für  $Q = P_0 + w$  gilt dann

$$\overrightarrow{\phi_A(P)\phi_A(Q)} = \overrightarrow{\phi(v)\phi(w)} = \phi(w) - \phi(v) = \phi(w - v)$$

Daher ist  $\phi_A$  injektiv. Die auf  $V$  induzierte Abbildung ist  $\phi$ , also linear. Jedes Element von  $k^n$  hat die Form  $\phi(v)$ , ist also von der Form  $\phi_A(P_0 + v)$ . Damit ist  $\phi_A$  bijektiv.  $\square$

Will man also affine Bewegungen verstehen, so genügt es, den Fall  $\mathbb{A}^n$  zu betrachten.

**Satz 9.29.** *Die Gruppe der affinen Bewegungen des  $\mathbb{A}^n$  ist gegeben durch Abbildungen der Form  $x \mapsto Ax + b$  mit  $A \in \text{GL}_n(k)$  und  $b \in k^n$ .*

*Beweis:* Sei  $\phi(x) = Ax + b$ . Dann gilt

$$y - x = \overrightarrow{x\bar{y}} \mapsto \overrightarrow{(Ax + b)(Ay + b)} = Ay + b - Ax - b = A(y - x)$$

Dies ist eine bijektive lineare Abbildung. Damit ist  $\phi$  eine Bewegung.

Sei umgekehrt  $\phi$  eine Bewegung. Sei  $b = \phi(0)$  und  $A$  die Matrix zur linearen Abbildung

$$x - y = \overrightarrow{x\bar{y}} \mapsto \overrightarrow{\phi(x)\phi(y)} = \phi(x) - \phi(y)$$

Diese ist invertierbar. Für jedes  $x \in \mathbb{A}^n$  ist

$$A(x - 0) = \phi(x) - \phi(0) = \phi(x) - b \Rightarrow \phi(x) = Ax + b$$

$\square$

**Bemerkung.** Noch besser: Wir betten  $\mathbb{A}^n$  als Vektoren der Form  $\begin{pmatrix} x \\ 1 \end{pmatrix}$  nach  $k^{n+1}$  ein. Matrizen der Form

$$M(A, b) = \begin{pmatrix} A & b \\ 0_n & 1 \end{pmatrix} \quad M \in \text{GL}_n(k), b \in k^n$$

respektieren den affinen Teilraum. Wir haben gerade gezeigt, dass jede affine Bewegung diese Form hat. Die Komposition der affinen Abbildungen ist nun einfach die Komposition der Matrizen  $M(A, b)$ .

Wir haben gesehen, dass dieser Standpunkt natürlich ist: Er benutzt einfach die Einbettung von  $\mathbb{A}^n$  nach  $\mathbb{P}^n$ .