Skript zur Vorlesung „Lineare Algebra I“

Alexander Schmitt

Berlin, Wintersemester 2013/2014

Inhalt

Inhalt . . . . . **i**

Einleitung . . . . . **iii**

I Lineare Gleichungssysteme **1**
I.1 (2×2)-Gleichungssysteme 1
I.2 $(m \times n)$-Gleichungssysteme 8
I.3 Zeilenumformungen von Matrizen 14
I.4 Das Gaußsche Lösungsverfahren 19

II Grundbegriffe **31**
II.1 Mengen und Abbildungen 31
II.2 Gruppen 43
II.3 Ringe und Körper 48
II.4 Äquivalenzrelationen 52

III Vektorräume und lineare Abbildungen **59**
III.1 Vektorräume 59
III.2 Erzeugendensysteme 63
III.3 Linear unabhängige Teilmengen 66
III.4 Basen 68
III.5 Lineare Abbildungen 76
III.6 Quotientenvektorräume 94

IV Matrizenrechnung **99**
IV.1 Das Matrixprodukt 100
IV.2 Dualräume 119
IV.3 Die symmetrische Gruppe 123
IV.4 Determinanten 132

V Eigenwerte und -vektoren 155
V.1 Äquivalenzrelationen auf Matrizen 155
V.2 Polynome 159
V.3 Charakteristische Polynome 166
V.4 Diagonalisierbare Endomorphismen 173
V.5 Der Satz von Cayley-Hamilton 178
V.6 Trigonalisierbare Endomorphismen 184
V.7 Nilpotente Endomorphismen 186
V.8 Die Jordansche Normalform 195

Literaturhinweise 207

Stichwortverzeichnis 209

Einleitung

Der vorliegende Text ist ein Skript zur Vorlesung „Lineare Algebra I“, die ich an der Universität Essen im Sommersemester 2004 und an der Freien Universität Berlin im Wintersemester 2013/2014 gehalten habe. Das Skript enthält die Theorie linearer Gleichungssysteme, endlichdimensionaler Vektorräume und linearer Abbildung bis hin zum Satz über die Jordansche Normalform.

Wir beginnen mit dem Lösungsverfahren linearer Gleichungssysteme nach Gauß. Dies vermittelt zum einen ein Gefühl für die Objekte, mit denen wir uns später beschäftigen werden, zum anderen wirft es einige grundlegende Fragen auf, die bei der weiteren Entwicklung der Theorie beantwortet werden. Der Gauß-Algorithmus wird uns als unerlässliches Hilfsmittel für explizite Berechnungen während der gesamten Vorlesung begleiten.

Das zweite Kapitel beschäftigt sich mit den Grundbegriffen der Mathematik. Wir gehen kurz auf die Mengenlehre ein. Insbesondere weisen wir darauf hin, dass alle Konzepte, mit denen wir uns in der Mathematik beschäftigen werden, mit Hilfe von Mengen definiert werden (müssen). Wir reden weiter über Gruppen, Ringe und Körper. Ohne Gruppen und Körper können wir nicht einmal den Begriff des Vektorraums erklären. Auch Ringe treten im späteren Verlauf als Polyomringe, Matrizenringe usw. auf. Ein weiterer zentraler Begriff ist der der Äquivalenzrelation. Er wird zum Beispiel bei der Konstruktion von Quotientenvektorräumen benötigt.

Kapitel drei ist der Entwicklung der Grundlagen der Linearen Algebra gewidmet. Die Lineare Algebra trifft Aussagen über Vektorräume und lineare Abbildungen zwischen ihnen. Daher werden diese Begriff eingeführt, an verschiedenen Beispielen illustriert und ihre fundamentalen Eigenschaften studiert. Oftmals werden wir uns auf endlichdimensionale Vektorräume beschränken. In der Tat werden an vielen Stellen Basen benutzt werden. Basen für unendlichdimensionale Vektorräume sind in den meisten Fällen nutzlos. (Es gibt natürlich Ausnahmen.) Im Unendlichdimensionalen sollte man auch unendliche Linearkombinationen zulassen. Dazu muss man Lineare Algebra und Analysis mit-

einander verbinden. Dies geschieht in der linearen Funktionalanalysis.

Das vierte Kapitel erklärt, wie man mit Matrizen arbeiten kann. Matrizen sind zunächst ein Hilfsmittel, um explizit mit linearen Abbildungen zu rechnen. Die Verknüofung von linearen Abbildungen, die sich elegant definieren und untersuchen lässt, ergibt eine sehr interessante mathematische Operation, das Produkt zweier Matrizen. Nun lassen sich die Techniken aus dem ersten Kapitel in die Theorie der Matrizen übertragen und führen zu nützlichen Anwendungen. Als weiteren beeindruckenden Formalismus besprechen wir die Theorie der Determinanten quadratischer Matrizen. In diesem Zusammenhang müssen wir uns auch zum ersten Mal gründlich mit der symmetrischen Gruppe auseinandersetzen.

In den ersten vier Kapiteln werden sehr viele Begriffe eingeführt. Die bewiesenen Sätze sind dabei meist recht elementar und intuitiv leicht greifbar. Im fünften Kapitel werden wir ein tiefer liegendes Problem untersuchen, und zwar die Klassifikation quadratischer Matrizen bis auf Ähnlichkeit. Ein äquivalentes Problem ist, zu einem gegebenen Endomorphismus eines endlichdimensionalen Vektorraums eine Basis zu finden, bzgl. der die darstellende Matrix besonders einfach ist. Das soll nicht nur bedeuten, dass sie schön anzusehen ist, sie soll uns auch möglichst viel über die Struktur des Endomorphismus verraten. Um das Problem zu lösen, haben wir also zu überlegen, wie wir die Struktur eines Endomorphismus überhaupt beschreiben möchten. Die zentralen Begriffe sind hier diejenigen des Eigenvektors und -werts. Eigenvektoren geben Richtungen im betrachteten Vektorraum an, die invariant unter dem gegebenen Endomorphismus sind. Eigenvektoren und -werte gibt es im Allgemeinen nur, wenn der zugrundeliegende Körper algebraisch abgeschlossen ist. An dieser Stelle gehen zum ersten Mal die Eigenschaften des Grundkörpers wesentlich in die Theorie ein. Dies ist ein Indiz für die Schwierigkeit des gestellten Problems. In speziellen Situation besitzen Endomorphismen nicht genügend Eigenvektoren. Um diese Situationen zu begreifen, muss man sogenannte nilpotente Endomorphismen betrachten. Unsere Bemühungen gipfeln schließlich im Satz über die Jordansche Normalform.

Ich danke Frau Anna Wißdorf und Herrn Ángel Muñoz für die Durchsicht des Manuskripts und zahlreiche Korrekturen. Die biographischen Angaben zu den MathematikerInnen stammen aus WIKIPEDIA. Das Skript basiert hauptsächlich auf dem Buch [6]. Weitere Standardwerke sind [2] und [3]. Das Buch [1] enthält verschiedene Anwendungen der Linearen Algebra.

Alexander Schmitt
Berlin, im April 2014

I

Lineare Gleichungssysteme

Lineare Gleichungssysteme treten in vielen Bereichen der Mathematik, Natur- und Wirtschaftswissenschaften auf. Daher benötigt man zunächst ein effektives Verfahren, um solche Gleichungssysteme zu lösen. Dieses Verfahren wird durch den sogenannten Gauß[1]-Algorithmus bereitgestellt. Da es sich dabei um ein sehr konkretes Verfahren handelt, bietet es sich an, die Vorlesung damit zu eröffnen. Wir werden sehen, dass bereits dieses Verfahren natürliche Fragestellungen theoretischer Art aufwirft (s. Problem I.4.6). Diese Probleme mögen als eine Motivation dienen, den abstrakten und umfangreichen Begriffsapparat der Linearen Algebra zu entwickeln. Diese Arbeit wird in Kapitel II begonnen. Es wird sich im weiteren Verlauf der Vorlesung herausstellen, dass der Gauß-Algorithmus an vielen Stellen das Verfahren ist, das konkrete Berechnungen möglich macht.

I.1 (2×2)-Gleichungssysteme

In diesem Abschnitt untersuchen wir den einfachsten Fall linearer Gleichungssysteme, an dem schon einige Phänomene studiert werden können. Zunächst müssen wir den Zahlbereich fixieren, in dem wir rechnen wollen. Dies muss ein sogenannter Körper K sein. Die formale Definition eines Körpers werden wir erst in Kapitel II, Definition II.3.1, kennenlernen. An dieser Stelle gehen wir davon aus, dass der Leser zumindest mit dem Körper $\mathbb{Q}$ der rationalen Zahlen und dem Körper $\mathbb{R}$ der reellen Zahlen vertraut ist. Vielleicht ist dem einen oder

[1] Johann Carl Friedrich Gauß (1777 - 1855), deutscher Mathematiker, Astronom und Physiker

anderen auch schon der Körper $\mathbb{C}$ der komplexen Zahlen begegnet, in dem man die Wurzel aus -1 ziehen kann.

Wir fixieren nun den Körper K, also z.B. $K = \mathbb{Q}$ oder $K = \mathbb{R}$. Wir benötigen auch das folgende Objekt:

$$K^2 := \left\{ \begin{pmatrix} s \\ t \end{pmatrix} \,\middle|\, s, t \in K \right\}.$$

Den Raum $\mathbb{R}^2$ mag man sich als die vor einem ausgebreitete, sich ins Unendliche erstreckende Papierebene vorstellen und die Zahlen s, t als die Koordinaten eines Punkts in der Ebene (s. Abbildung I.1).

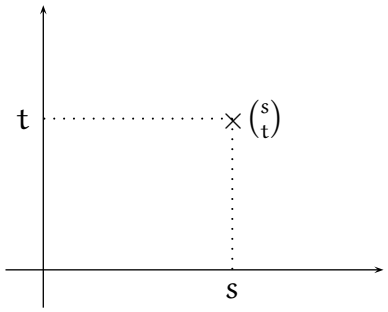


Abbildung I.1: Koordinaten im $\mathbb{R}^2$.

Wir betrachten das Gleichungssystem

$$\begin{array}{rcl} a \cdot x + b \cdot y & = & e \\ c \cdot x + d \cdot y & = & f. \end{array} \tag{I.1}$$

Dabei sind die Elemente a, b, c, d, e und f aus dem Körper K, also beispielsweise reelle Zahlen, und x und y sind Unbestimmte. Die *Lösungsmenge des Gleichungssystems* (I.1) ist

$$\mathbb{L} := \left\{ \begin{pmatrix} s \\ t \end{pmatrix} \in K^2 \,\middle|\, \begin{array}{rcl} a \cdot s + b \cdot t & = & e \\ c \cdot s + d \cdot t & = & f \end{array} \right\}.$$

Wir wollen nun allgemeine Aussagen über die *Struktur* von $\mathbb{L}$ gewinnen. Dazu unterscheiden wir verschiedene Fälle.

Fall 0: $a = b = c = d = 0$**. —** Dann gibt es offenbar zwei Möglichkeiten. Entweder gilt $e = f = 0$ und $\mathbb{L} = K^2$ oder $e \neq 0$ oder $f \neq 0$ und $\mathbb{L} = \varnothing$,[2] d.h. es gibt keine Lösung.

[2] Das Symbol „$\varnothing$“ steht für die leere Menge (s. Beispiel II.1.1, i).

Jetzt setzen wir voraus, dass mindestens eine der Zahlen a, b, c, d von Null verschieden ist. Nach etwaiger Vertauschung der beiden Zeilen des Gleichungssystems und/oder nach Umbenennung der Unbestimmten x in y und der Unbestimmten y in x können wir annehmen, dass $a \neq 0$ gilt.

Bemerkung I.1.1. Das Vertauschen der beiden Zeilen des Gleichungssystems ändert die Lösungsmenge offenbar nicht. Für die Lösungsmenge $\mathbb{L}'$ des Gleichungssystems, das durch Umbenennung der Variablen aus (I.1) entsteht, gilt

$$\mathbb{L}' = \left\{ \begin{pmatrix} s \\ t \end{pmatrix} \in K^2 \;\middle|\; \begin{pmatrix} t \\ s \end{pmatrix} \in \mathbb{L} \right\}.$$

Nun können wir in (I.1) das (c/a)-Fache der ersten Gleichung von der zweiten Gleichung abziehen und erhalten das Gleichungssystem

$$\begin{array}{rcl} a \cdot x + b \cdot y & = & e \\ \left(d - \dfrac{bc}{a}\right) \cdot y & = & \left(f - \dfrac{ec}{a}\right). \end{array} \tag{I.2}$$

Man überlegt sich leicht, dass die Lösungsmenge des Gleichungssystems (I.2) mit der Lösungsmenge des Gleichungssystems (I.1) übereinstimmt. Wegen $a \neq 0$ ist die zweite Gleichung in (I.2) äquivalent zu der Gleichung

$$(ad - bc) \cdot y = af - ec. \tag{I.3}$$

Definition I.1.2. Ein Zahlenschema

$$M := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

bezeichnen wir als (2×2)-*Matrix*. Die *Determinante von* M ist die Zahl

$$\det(M) := \delta(a, b, c, d) := ad - bc.$$

Damit lässt sich Gleichung (I.3) schreiben als

$$\delta(a, b, c, d) \cdot y = \delta(a, e, c, f). \tag{I.4}$$

Bemerkung I.1.3. Auch wenn $a = 0$ gilt, können wir c-mal die erste Gleichung in (I.1) von a-mal der zweiten Gleichung abziehen und die gültige Gleichung

$$\delta(a, b, c, d) \cdot y = \delta(a, e, c, f)$$

erhalten. Ebenso errechnet man

$$\delta(a, b, c, d) \cdot x = -\delta(b, e, d, f).$$

Wir unterscheiden nun zwei weitere Fälle.

Fall A: $\delta(a,b,c,d) \neq 0$. — In diesem Fall berechnet man

$$y = \frac{\delta(a,e,c,f)}{\delta(a,b,c,d)} \quad \text{und} \quad x = -\frac{\delta(b,e,d,f)}{\delta(a,b,c,d)}.$$

Das Gleichungssystem besitzt also genau eine Lösung,

$$\mathbb{L} = \left\{ \begin{pmatrix} s_0 \\ t_0 \end{pmatrix} \right\}.$$

Fall B: $\delta(a,b,c,d) = 0$. — Auf Grund von Bemerkung I.1.3 gilt $\mathbb{L} = \varnothing$, falls[3] $\delta(a,e,c,f) \neq 0$ oder $\delta(b,e,d,f) \neq 0$. Gilt jedoch $\delta(a,e,c,f) = 0$, dann ist (I.3) für alle $y = t \in K$ erfüllt, und aus der ersten Gleichung in (I.2) ergibt sich

$$x = \frac{e}{a} - \frac{b}{a} \cdot t,$$

also

$$\mathbb{L} = \left\{ \begin{pmatrix} \frac{e}{a} \\ 0 \end{pmatrix} + t \cdot \begin{pmatrix} -\frac{b}{a} \\ 1 \end{pmatrix} \,\middle|\, t \in K \right\}.$$

Für $K = \mathbb{R}$ ist die Lösungsmenge also eine Gerade in $\mathbb{R}^2$ (s. Abbildung 1.2).

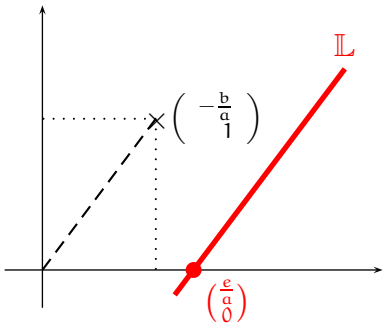


Abbildung I.2: Die Lösungsmenge im Fall B.

[3] Die Bedingungen $\delta(a,e,c,f) = 0$ und $\delta(b,e,d,f) = 0$ sind nicht unabhängig. Gilt $\delta(a,e,c,f) = af - ce = 0$, so folgt wegen $\delta(a,b,c,d) = ad - bc = 0$ auch

$$0 = b \cdot (af - ce) - e \cdot (ad - bc) = abf - ade = a \cdot (bf - de) = a \cdot \delta(b,e,d,f).$$

Mit $a \neq 0$ ergibt sich $\delta(b,e,d,f) = 0$. Ebenso schließt man von $\delta(a,b,c,d) = 0 = \delta(b,e,d,f)$ auf $\delta(a,e,c,f) = 0$.

Bemerkung I.1.4. Dem Gleichungssystem (I.1) hatten wir die Matrix

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

zugeordnet. Der Vertauschung der beiden Gleichungen entspricht auf Matrizenniveau die Umformung

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \rightsquigarrow \begin{pmatrix} c & d \\ a & b \end{pmatrix}.$$

Man beachte

$$\delta(a, b, c, d) = ad - bc = -(cb - da) = -\delta(c, d, a, b).$$

Der Umbenennung der Variablen entspricht die Operation

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \rightsquigarrow \begin{pmatrix} b & a \\ d & c \end{pmatrix}.$$

Dabei gilt

$$\delta(a, b, c, d) = ad - bc = -(bc - ad) = -\delta(b, a, d, c).$$

Die Bedingung „$\delta(a, b, c, d) = 0$“ ist also invariant unter Vertauschung der Gleichungen und Umbenennung der Unbestimmten. Dasselbe gilt für die Bedingung „$\delta(a, e, c, f) = 0 = \delta(b, e, d, f)$“.

Bemerkung I.1.5. Das Gleichungssystem (I.1) und seine Lösungen haben auch eine geometrische Bedeutung. Es seien a, b und e Elemente der Körpers K. Wir schauen uns die Gleichung

$$a \cdot x + b \cdot y = e$$

an. Die zugehörige Lösungsmenge ist

$$\mathbb{G}_1 := \left\{ \begin{pmatrix} s \\ t \end{pmatrix} \in K^2 \,\middle|\, a \cdot s + b \cdot t = e \right\}.$$

Gilt $a = b = e = 0$, so finden wir $\mathbb{G}_1 = K^2$. Falls $a = b = 0$ und $e \neq 0$, dann folgt $\mathbb{G}_1 = \varnothing$. In allen anderen Fällen ist $\mathbb{G}_1$ eine Gerade in K^2. Für das Folgende beschränken wir uns auf diesen Fall. Weiter seien $c, d, f \in K$. Wir nehmen $c \neq 0$ oder $d \neq 0$ an. Dann ist die Lösungsmenge

$$\mathbb{G}_2 := \left\{ \begin{pmatrix} s \\ t \end{pmatrix} \in K^2 \,\middle|\, c \cdot s + d \cdot t = f \right\}$$

der Gleichung

$$c \cdot x + d \cdot y = f$$

ebenfalls eine Gerade in K^2.

Für die Lösungsmenge von (I.1) gilt

$$\mathbb{L} = \mathbb{G}_1 \cap \mathbb{G}_2,$$

d.h. sie bildet den Durchschnitt der beiden Geraden. Zwei Geraden können gleich sein, verschieden und parallel, oder sie haben genau einen Schnittpunkt (s. Abbildung I.3). Wir überlassen es der Leserin bzw. dem Leser die verschiedenen Fälle den verschiedenen Fällen in der vorangegangen Diskussion zuzuordnen.

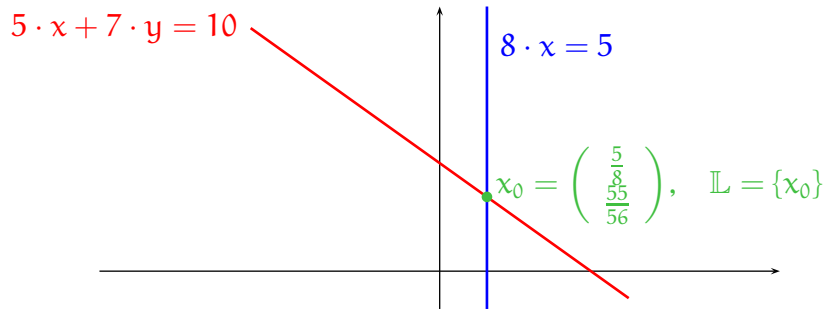


Abbildung I.3: Der Schnitt zweier nicht paralleler Geraden.

Zusammenfassung. — Aus unseren Überlegungen erhalten wir folgende Aussagen über die Struktur der Lösungsmenge eines (2×2)-Gleichungssystems.

Fall	Bedingung an die Koeffizienten	Gestalt von $\mathbb{L}$
A	$\delta(a,b,c,d) \neq 0$	Ein Punkt
B	$\delta(a,b,c,d) = 0$; $\delta(a,e,c,f) = 0$ und $\delta(b,e,d,f) = 0$; ein Koeffizient a, b, c, d ungleich null	Gerade in K^2
B	$\delta(a,b,c,d) = 0$; $\delta(a,e,c,f) \neq 0$ oder $\delta(b,e,d,f) \neq 0$	Leer
0	$a = b = c = d = e = f = 0$	K^2
0	$a = b = c = d = 0$; $e \neq 0$ oder $f \neq 0$	Leer

Die obige Tabelle stellt einen Zusammenhang zwischen den Ausgangsdaten des Gleichungssytems und der Gestalt der Lösungsmenge her. Gilt $e = f = 0$, dann sprechen wir von einem *homogenen Gleichungssystem*. Man beachte, dass ein homogenes Gleichungssystem immer die „triviale Lösung“ $\binom{0}{0}$ besitzt. Für homogene Gleichungssysteme ist der Zusammenhang zwischen den Koeffizienten a, b, c und d und der Gestalt der Lösungsmenge noch einprägsamer.

Fall	Bedingung an die Koeffizienten	Gestalt von $\mathbb{L}$
A	$\delta(a,b,c,d) \neq 0$	$\mathbb{L} = \left\{ \binom{0}{0} \right\}$
B	$\delta(a,b,c,d) = 0$; ein Koeff. a,b,c,d nicht null	Gerade in K^2 durch $\binom{0}{0}$
0	$a = b = c = d = 0$	K^2

Legt man den anschaulichen Dimensionsbegriff zu Grunde, nach dem K^2 zweidimensional, eine Gerade im K^2 eindimensional und ein Punkt nulldimensional ist, dann kann man diese Tabelle wie folgt interpretieren: Gilt $\delta(a,b,c,d) \neq 0$, so enthält das Gleichungssystem (I.1) zwei unabhängige Bedinungen (vgl. Aufgabe I.1.6), so dass die Dimension des Lösungsraumes $\mathbb{L}$ gegenüber der Dimension von K^2 um zwei vermindert ist. Im Fall B haben wir nur eine unabhängige Bedingung, so dass die Dimension von $\mathbb{L}$ um eins niedriger ist als die von K^2. Im Fall 0 haben wir natürlich gar keine Bedingung.

Schließlich wollen wir noch festhalten, dass das Studium linearer Gleichungssysteme zwei Aspekte hat: Zum Einen den „praktischen Aspekt“ der expliziten Bestimmung der Lösungsmenge, zum Anderen den „theoretischen Aspekt“ der Beschreibung der Struktur der Lösungsmenge in Abhängigkeit der Koeffizienten (z.B. Zusammenhang zwischen Anzahl der gestellten Bedingungen und der Dimension des Lösungsraumes). Im nächsten Abschnitt werden wir uns mit beliebigen Gleichungssystemen beschäftigen. Der praktische Aspekt ist dabei der Gauß-Algorithmus, während sich der theoretische Aspekt wesentlich umfangreicher gestaltet und die Einführung einer sehr formalen Sprache notwendig macht.

Aufgabe I.1.6.
Es seien K ein Körper und

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad a,b,c,d \in K,$$

eine (2×2)-Matrix. Zeigen Sie, dass die folgenden Bedingungen äquivalent sind.

1. $\delta(a,b,c,d) = 0$.
2. Es gibt eine Zahl $\lambda \in K$, so dass $c = \lambda \cdot a$ und $d = \lambda \cdot b$ oder $a = \lambda \cdot c$ und $b = \lambda \cdot d$.
3. Es gibt eine Zahl $\lambda \in K$, so dass $b = \lambda \cdot a$ und $d = \lambda \cdot c$ oder $a = \lambda \cdot b$ und $c = \lambda \cdot d$.

Was bedeutet die dritte Bedingung geometrisch für die 2-Tupel $\begin{pmatrix} a \\ c \end{pmatrix}$ und $\begin{pmatrix} b \\ d \end{pmatrix}$ in K^2?

Aufgabe I.1.7.
Gegeben seien Elemente $\lambda, e, f \in \mathbb{R}$. Wir betrachten das (2×2)-Gleichungssystem

$$\begin{array}{rcrcl} -x & + & (\lambda-2)\cdot y & = & e \\ (2-\lambda)\cdot x & + & y & = & f. \end{array}$$

Bestimmen Sie alle $\lambda \in K$, für die das Gleichungssystem **eindeutig** lösbar ist. Unter welchen Bedingungen an λ, e und f besitzt das Gleichungssystem keine Lösung bzw. eine Schar von Lösungen?

Aufgabe I.1.8.
Überprüfen Sie mit dem Determinantenkriterium die Lösbarkeit der folgenden (2×2)-Gleichungssysteme und geben Sie die Lösungsmenge an.

$$\text{a)} \begin{array}{rcrcl} 3\cdot x & + & y & = & 7 \\ -6\cdot x & + & 2\cdot y & = & -14 \end{array}, \quad \text{b)} \begin{array}{rcrcl} 3\cdot x & - & y & = & 7 \\ -6\cdot x & + & 2\cdot y & = & -14 \end{array},$$

$$\text{c)} \begin{array}{rcrcl} 3\cdot x & - & y & = & 7 \\ -6\cdot x & + & 2\cdot y & = & 14 \end{array}.$$

I.2 $(m \times n)$-Gleichungssysteme

Wir arbeiten wiederum über dem Körper K.

Definitionen I.2.1. i)

$$K^n := \left\{ \left(\begin{array}{c} s_1 \\ \vdots \\ s_n \end{array} \right) \;\middle|\; s_1, \ldots, s_n \in K \right\}.$$

Ein Element $\left(\begin{array}{c} s_1 \\ \vdots \\ s_n \end{array} \right)$ aus K^n heißt n-*Tupel* oder auch *Spaltenvektor*. Das *Nulltupel* oder der *Nullvektor* ist das Element

$$\mathbf{0} := \mathbf{0}_m := \left(\begin{array}{c} 0 \\ \vdots \\ 0 \end{array} \right).$$

ii) (Addition von n-Tupeln) Für zwei n-Tupel aus K^n ist die *Summe* erklärt durch

$$\left(\begin{array}{c} s_1 \\ \vdots \\ s_n \end{array} \right) + \left(\begin{array}{c} t_1 \\ \vdots \\ t_n \end{array} \right) := \left(\begin{array}{c} s_1 + t_1 \\ \vdots \\ s_n + t_n \end{array} \right).$$

iii) (Skalarmultiplikation) Für ein Element $a \in K$ und ein n-Tupel ist das *Produkt* gegeben als

$$a \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} := a \cdot \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} := \begin{pmatrix} a \cdot s_1 \\ \vdots \\ a \cdot s_n \end{pmatrix}.$$

iv) Ein *lineares* $(m \times n)$-*Gleichungssystem* besteht aus m Gleichungen

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

in den Unbestimmten $x_1,\ldots,x_n$. Die *Koeffizienten* a_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$, sowie $b_1,\ldots,b_m$ gehören dabei dem Körper K an. Die Koeffizienten tragen wir in eine sogenannte $(m \times n)$-*Matrix*

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

ein und die Elemente $b_1,\ldots,b_m$ in das m-Tupel

$$\mathbf{b} := \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Falls $\mathbf{b} = \mathbf{0}_m$ gilt, nennen wir das Gleichungssystem *homogen*, ansonsten *inhomogen*. Die *Lösungsmenge* ist

$$\mathbb{L}(A, \mathbf{b}) := \left\{ \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} \in K^n \;\middle|\; \begin{array}{ccccccccc} a_{11} \cdot s_1 & + & a_{12} \cdot s_2 & + & \cdots & + & a_{1n} \cdot s_n & = & b_1 \\ a_{21} \cdot s_1 & + & a_{22} \cdot s_2 & + & \cdots & + & a_{2n} \cdot s_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1} \cdot s_1 & + & a_{m2} \cdot s_2 & + & \cdots & + & a_{mn} \cdot s_n & = & b_m \end{array} \right\}.$$

Beispiel I.2.2. Die Skalarmultiplikation und die Addition von Vektoren kann man sich im $\mathbb{R}^2$ und $\mathbb{R}^3$ leicht veranschaulichen. Die Summe erhält man wie in Abbildung I.4. Die Skalarmultiplikation mit einem Element $a \in K$ ist geometrisch die zentrische Streckung (mit Zentrum $\mathbf{0}$ um den Faktor a (s. Abbildung I.5).

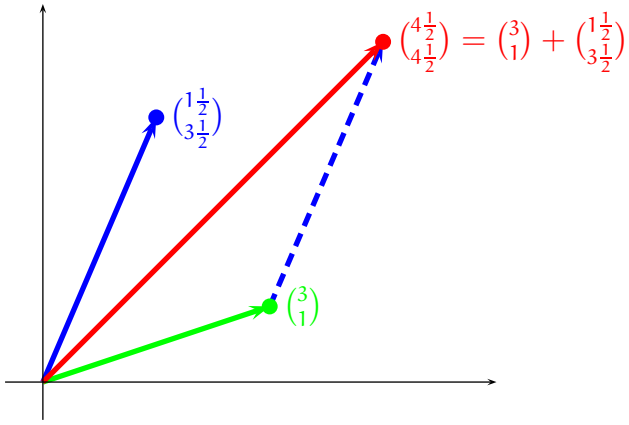


Abbildung I.4: Addition von Vektoren im $\mathbb{R}^2$.

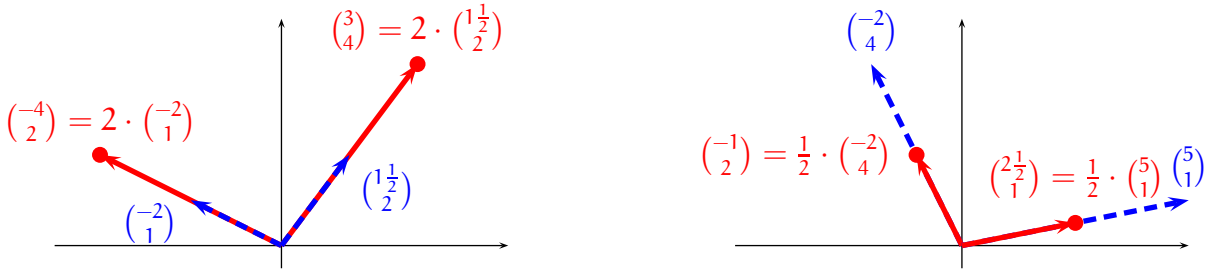


Abbildung I.5: Skalarmultiplikation im $\mathbb{R}^2$.

Bemerkungen I.2.3. i) Zwei n-Tupel $\begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix}, \begin{pmatrix} s'_1 \\ \vdots \\ s'_n \end{pmatrix} \in K^n$ sind genau dann gleich, wenn $s_i = s'_i$, $i = 1, ..., n$, gilt.

ii) Offenbar gilt immer

$$\mathbf{0}_n \in \mathbb{L}(A, \mathbf{0}_m).$$

iii) Zur Eingewöhnung werden wir Elemente aus K^n durch fett gedruckte Buchstaben kenntlich machen, z.B. $\mathbf{s} \in K^n$. Diese Gewohnheit werden wir aber mit der Zeit wieder ablegen.

iv) Der Raum K^n ist der Prototyp eines endlichdimensionalen K-Vektorraums. Die in K gültigen Rechenregeln implizieren dabei gewisse Rechenregeln, die in Aufgabe I.2.10 besprochen werden.

Die Lösungsmengen homogener Gleichungssysteme weisen eine zusätzliche Struktur auf.

Lemma I.2.4. *Für die Lösungsmenge $\mathbb{L}(A, \mathbf{0})$ eines homogenen linearen Gleichungssystems gilt:*

- $\mathbf{0} \in \mathbb{L}(A, \mathbf{0})$,
- $\mathbf{s}, \mathbf{t} \in \mathbb{L}(A, \mathbf{0}) \Longrightarrow \mathbf{s} + \mathbf{t} \in \mathbb{L}(A, \mathbf{0})$ *und*
- $\mathbf{s} \in \mathbb{L}(A, \mathbf{0})$, $a \in K \Longrightarrow a \cdot \mathbf{s} \in \mathbb{L}(A, \mathbf{0})$.

Beweis. Der einfache Beweis wird dem Leser als Übungsaufgabe überlassen. Die notwendigen Beweistechniken werden im Beweis von Lemma I.2.7 besprochen. □

Die Lösungsmenge eines homogenen Gleichungssystems erbt also vom K^n die zusätzlichen Strukturen „Addition“ und „Skalarmultiplikation“. Wir definieren allgemeiner.

Definition I.2.5. Eine Teilmenge $U \subseteq K^n$ wird *linearer Teilraum* oder *Unterraum* genannt, wenn die folgenden Bedingungen erfüllt sind:

- $\mathbf{0}_n \in U$,
- $\mathbf{s}, \mathbf{t} \in U \Longrightarrow \mathbf{s} + \mathbf{t} \in U$ und
- $\mathbf{s} \in U$, $a \in K \Longrightarrow a \cdot \mathbf{s} \in U$.

Beispiele I.2.6. i) Offensichtlich sind $\{0\}$ und K^n lineare Teilräume von K^n.

ii) Nach Lemma I.2.4 ist die Lösungsmenge eines homogenen linearen $(m \times n)$-Gleichungssystems ein linearer Unterraum von K^n.

iii) Für K^2 sind die linearen Unterräume $\{0\}$, K^2 sowie die Geraden durch $\binom{0}{0}$ (s. Aufgabe I.2.11). Im K^3 sind $\{0\}$, K^3, die Geraden durch $\mathbf{0}_3$ sowie die Ebenen, die $\mathbf{0}_3$ enthalten, die linearen Teilräume.

Das folgende Resultat erlaubt ein besseres Verständnis der Lösungsmengen inhomogener Gleichungssysteme.

Lemma I.2.7. *Es sei $\mathbf{s}^0 \in \mathbb{L}(A, \mathbf{b})$. Dann gilt*

$$\mathbb{L}(A, \mathbf{b}) = \mathbf{s}^0 + \mathbb{L}(A, \mathbf{0}_m) = \left\{ \mathbf{s}^0 + \mathbf{t} \in K^n \,|\, \mathbf{t} \in \mathbb{L}(A, \mathbf{0}_m) \right\}.$$

Man erhält also sämtliche Lösungen eines inhomogenen Gleichungssystems, indem man zu einer speziellen Lösung des inhomogenen Systems alle Lösungen des zugehörigen homogenen Gleichungssystems addiert.

Beweis. Es sei

$$\mathbf{v}_j := \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

die j-te Spalte der Matrix $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$, $j = 1, \dots, n$. Unter Benutzung der Addition und der Skalarmultiplikation erhalten wir folgende Charakterisierung für die Elemente der Lösungsmenge:

$$\left(\mathbf{s} = \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} \in \mathbb{L}(A, \mathbf{b})\right) \Longleftrightarrow \left(s_1 \cdot \mathbf{v}_1 + \cdots + s_n \cdot \mathbf{v}_n = \mathbf{b}\right). \tag{I.5}$$

Die gegebene spezielle Lösung sei $\mathbf{s}^0 = \begin{pmatrix} s_1^0 \\ \vdots \\ s_n^0 \end{pmatrix}$. Wir zeigen zunächst $\mathbb{L}(A, \mathbf{b}) \subseteq \mathbf{s}^0 + \mathbb{L}(A, \mathbf{0})$. Es sei $\mathbf{s} = \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} \in \mathbb{L}(A, \mathbf{b})$. Es gelten die Gleichungen

$$\begin{aligned} s_1 \cdot \mathbf{v}_1 + \cdots + s_n \cdot \mathbf{v}_n &= \mathbf{b} \\ s_1^0 \cdot \mathbf{v}_1 + \cdots + s_n^0 \cdot \mathbf{v}_n &= \mathbf{b}. \end{aligned}$$

Daraus ergibt sich

$$(s_1 - s_1^0) \cdot \mathbf{v}_1 + \cdots + (s_n - s_n^0) \cdot \mathbf{v}_n = \mathbf{0}.$$

In der Tat ist ja die i-te Komponente des Vektors auf der linken Seite

$$(s_1 - s_1^0)a_{i1} + \cdots + (s_n - s_n^0)a_{in} = (a_{i1}s_1 + \cdots + a_{in}s_n) - (a_{i1}s_1^0 + \cdots + a_{in}s_n^0) = b_i - b_i = 0,$$

$i = 1, \dots, m$. Somit gilt

$$\mathbf{t} := \begin{pmatrix} s_1 - s_1^0 \\ \vdots \\ s_n - s_n^0 \end{pmatrix} \in \mathbb{L}(A, \mathbf{0})$$

und

$$\mathbf{s} = \mathbf{s}^0 + \mathbf{t} \in \mathbf{s}^0 + \mathbb{L}(A, \mathbf{0}).$$

Es bleibt $\mathbf{s}^0 + \mathbb{L}(A, \mathbf{0}) \subseteq \mathbb{L}(A, \mathbf{b})$ zu zeigen. Dazu sei $\mathbf{t} = \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix} \in \mathbb{L}(A, \mathbf{0})$. Es sind also die beiden Gleichungen

$$\begin{aligned} s_1^0 \cdot \mathbf{v}_1 + \cdots + s_n^0 \cdot \mathbf{v}_n &= \mathbf{b} \\ t_1 \cdot \mathbf{v}_1 + \cdots + t_n \cdot \mathbf{v}_n &= \mathbf{0} \end{aligned}$$

erfüllt. Daher gilt auch

$$(s_1^0 + t_1) \cdot \mathbf{v}_1 + \cdots + (s_n^0 + t_n) \cdot \mathbf{v}_n = \mathbf{b}.$$

Wie bereits in (I.5) beobachtet wurde, ist dies gleichbedeutend mit der Tatsache, dass $\mathbf{s}^0 + \mathbf{t} \in \mathbb{L}(A, \mathbf{b})$. □

Beispiel I.2.8. Man beachte, dass für $\mathbf{b} \neq \mathbf{0}$ niemals $\mathbf{0} \in \mathbb{L}(A, \mathbf{b})$ gelten kann. Für eine $(m \times 3)$-Matrix A und einen Vektor $\mathbf{b} \neq \mathbf{0}_m$ ist die Lösungsmenge $\mathbb{L}(A, \mathbf{b})$ also entweder a) leer oder b) ein Punkt ungleich $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ oder c) eine Gerade bzw. d) Ebene, die $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ nicht enthält. Im Fall b) - d) wird $\mathbb{L}(A, \mathbf{b})$ durch Parallelverschiebung entlang einer speziellen Lösung $\mathbf{s}^0 \in \mathbb{L}(A, \mathbf{b})$ aus $\mathbb{L}(A, \mathbf{0}_m)$ erhalten.

Aufgabe I.2.9.
Es seien die beiden 2-Tupel

$$\begin{pmatrix} k \\ l \end{pmatrix}, \begin{pmatrix} m \\ n \end{pmatrix} \in K^2$$

gegeben, und es gelte $\delta(k, m, l, n) \neq 0$. Zeigen Sie, dass es für jedes 2-Tupel $\begin{pmatrix} v \\ w \end{pmatrix} \in K^2$ eindeutig bestimmte Zahlen $s_1, s_2 \in K$ mit

$$s_1 \cdot \begin{pmatrix} k \\ l \end{pmatrix} + s_2 \cdot \begin{pmatrix} m \\ n \end{pmatrix} = \begin{pmatrix} v \\ w \end{pmatrix} \quad \text{gibt.}$$

Aufgabe I.2.10.
Zeigen Sie, dass für die Addition und die Skalarmultiplikation auf K^n die folgenden Gesetze gelten. (Sie dürfen gerne $K = \mathbb{Q}$ oder $K = \mathbb{R}$ voraussetzen.)

1. $\mathbf{s} + \mathbf{s}' = \mathbf{s}' + \mathbf{s}$ für alle $\mathbf{s}, \mathbf{s}' \in K^n$ (Kommutativgesetz);
2. $\mathbf{0} + \mathbf{s} = \mathbf{s}$ für alle $\mathbf{s} \in K^n$ (Null ist Neutralelement);

3. $\mathbf{s} + (-\mathbf{s}) = \mathbf{0}$ für alle $\mathbf{s} \in K^n$, $-\mathbf{s} := (-1) \cdot \mathbf{s}$ ($-\mathbf{s}$ ist das inverse Element zu $\mathbf{s}$);
4. $\mathbf{s}+(\mathbf{s}'+\mathbf{s}'') = (\mathbf{s}+\mathbf{s}')+\mathbf{s}''$ für alle $\mathbf{s}, \mathbf{s}', \mathbf{s}'' \in K^n$ (Assoziativgesetz der Addition);
5. $1 \cdot \mathbf{s} = \mathbf{s}$ für alle $\mathbf{s} \in K^n$;
6. $(a + a') \cdot \mathbf{s} = a \cdot \mathbf{s} + a' \cdot \mathbf{s}$ für alle $a, a' \in K$, $\mathbf{s} \in K^n$ (Erstes Distributivgesetz);
7. $a \cdot (\mathbf{s} + \mathbf{s}') = a \cdot \mathbf{s} + a \cdot \mathbf{s}'$ für alle $a \in K$, $\mathbf{s}, \mathbf{s}' \in K^n$ (Zweites Distributivgesetz);
8. $(a \cdot a') \cdot \mathbf{s} = a \cdot (a' \cdot \mathbf{s})$ für alle $a, a' \in K$, $\mathbf{s} \in K^n$ (Assoziativgesetz der Skalarmultiplikation).

Aufgabe I.2.11.
a) Gegeben seien $\mathbf{v}_1, ..., \mathbf{v}_l \in K^n$. Die *lineare Hülle von* $\mathbf{v}_1, ..., \mathbf{v}_l$ ist gegeben als

$$\langle \mathbf{v}_1, ..., \mathbf{v}_l \rangle := \left\{ a_1 \cdot \mathbf{v}_1 + \cdots + a_l \cdot \mathbf{v}_l \,\middle|\, a_1, ..., a_l \in K \right\} \subseteq K^n.$$

Zeigen Sie, dass $\langle \mathbf{v}_1, ..., \mathbf{v}_l \rangle$ ein linearer Unterraum von K^n ist.
b) Bestimmen Sie die linearen Unterräume von K^2.

Aufgabe I.2.12.
Entwickeln Sie für Gleichungssysteme in drei Unbestimmten mit ein, zwei, oder drei Gleichungen eine geometrische Anschauung wie in Bemerkung I.1.5. Fertigen Sie Skizzen für die Lösungsmengen der Gleichungssysteme an, die aus einer, zwei oder allen drei der folgenden Gleichungen

$$\begin{array}{rcrcrcl} x & + & 2 \cdot y & - & z & = & 0 \\ 6 \cdot x & - & 3 \cdot y & - & z & = & 0 \\ 2 \cdot x & + & y & - & z & = & 6 \end{array}$$

bestehen, an.

I.3 Zeilenumformungen von Matrizen

Definition I.3.1. Gegeben sei eine $(p \times q)$-Matrix

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1q} \\ c_{21} & c_{22} & \cdots & c_{2q} \\ \vdots & \vdots & \cdots & \vdots \\ c_{p1} & c_{p2} & \cdots & c_{pq} \end{pmatrix}.$$

Wir betrachten die *Zeilenumformungen vom Typ*:

(ZI) Addition eines Vielfachen einer Zeile zu einer anderen Zeile;

(ZII) Vertauschung zweier Zeilen;

(ZIII) Multiplikation einer Zeile mit einem Element $a \in K^* := K \setminus \{0\}$.

Beispiele I.3.2. i) In der Matrix

$$C := \begin{pmatrix} 0 & -1 & 2 & -3 \\ 5 & 7 & -11 & 14 \\ 8 & 0 & -5 & 2 \\ -1 & -2 & 4 & -5 \end{pmatrix}$$

addieren wir fünfmal die vierte Zeile zur zweiten. Wir erhalten die Matrix

$$C' := \begin{pmatrix} 0 & -1 & 2 & -3 \\ 0 & -3 & 9 & -11 \\ 8 & 0 & -5 & 2 \\ -1 & -2 & 4 & -5 \end{pmatrix}.$$

ii) In der Matrix

$$D := \begin{pmatrix} 12 & -9 & 18 \\ 7 & 11 & -13 \\ 8 & 0 & 17 \\ -12 & -21 & 4 \end{pmatrix}$$

vertauschen wir die vierte und die zweite Zeile. Wir erhalten

$$D' := \begin{pmatrix} 12 & -9 & 18 \\ -12 & -21 & 4 \\ 8 & 0 & 17 \\ 7 & 11 & -13 \end{pmatrix}.$$

iii) In der Matrix

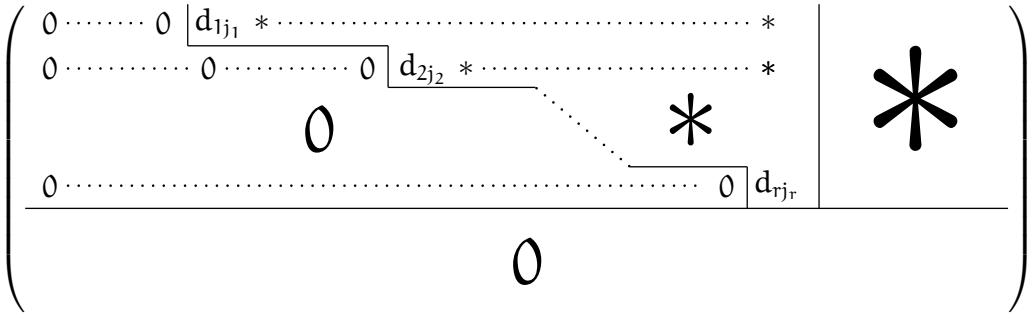
$$E := \begin{pmatrix} 12 & -21 & 42 & -9 & 18 \\ 8 & 0 & 17 & 11 & -13 \end{pmatrix}$$

multilplizieren wir die erste Zeile mit $1/3$. Es ergibt sich die Matrix

$$E' := \begin{pmatrix} 4 & -7 & 14 & -3 & 6 \\ 8 & 0 & 17 & 11 & -13 \end{pmatrix}.$$

Mit den oben angegebenen Zeilenumformungen lässt sich eine gegebene Matrix C vereinfachen.

Satz I.3.3. i) *Durch wiederholte Anwendung von Zeilenoperationen vom Typ* (ZI) *und* (ZII) *kann man eine gegebene Matrix* C *in eine Matrix* D *der Gestalt*



überführen. Genauer gesagt besitzt die Matrix D *die folgenden Eigenschaften: Es gibt eine ganze Zahl* r *mit* $0 \leq r \leq p, q$ *sowie Indizes* $1 \leq j_1 < j_2 < \cdots < j_r \leq q$*, so dass gilt*

a) $d_{ij_i} \neq 0$, $i = 1, ..., r$*;*

b) $d_{ij} = 0$ *für* $i = 1, ..., r$ *und* $1 \leq j < j_i$*;*

c) $d_{ij} = 0$ *für alle* $i > r$ *und* $j = 1, ..., q$.

ii) *Erlaubt man auch Zeilenoperationen vom Typ* (ZIII)*, dann kann man zusätzlich* $d_{ij_i} = 1$, $i = 1, ..., r$*, annehmen.*

Definition I.3.4. Wir sagen, dass sich eine Matrix D in *Zeilenstufenform* befindet, wenn sie die in Satz I.3.3, i), angegebenen Eigenschaften aufweist.

Bemerkung I.3.5. Addiert man in der Matrix D das $(-d_{1j_2}/d_{2j_2})$-Fache der zweiten Zeile zur ersten Zeile, das $(-d_{2j_3}/d_{3j_3})$-Fache der dritten zur zweiten, das $(-d_{1j_3}/d_{3j_3})$-Fache der dritten zur ersten usw., dann kann man zusätzlich noch $d_{i'j_i} = 0$ für $i = 1, ..., r$ und $1 \leq i' < i$ erreichen, d.h. über den Einträgen d_{ij_i}, $i = 1, ..., r$, an den Stufen stehen nur Nullen.

Beweis von Satz I.3.3. Zunächst bemerken wir, dass Teil ii) sofort aus Teil i) folgt. Man muss nur die ersten r-Zeilen der Matrix D noch mit dem entsprechenden Faktor multiplizieren.

Zu i). Falls alle Koeffizienten von C gleich null sind, dann ist $r = 0$ und C hat bereits die entsprechende Form. Ansonsten sei j_1 der kleinste unter den Indizes j, so dass die j-te Spalte ein von Null verschiedenes Element enthält. Es sei etwa $c_{ij_1} \neq 0$. Wir vertauschen die erste und i-te Zeile von C.[4] Die resultierende Matrix

[4]Falls $i = 1$, bedeutet dies, nichts an der Matrix zu ändern.

C' hat folgende Form:

$$C' = \left(\begin{array}{c|cccc} 0 \cdots\cdots 0 & \overset{\neq 0}{d_{1j_1}} & d_{1(j_1+1)} & \cdots\cdots & d_{1q} \\ \hline & c'_{2j_1} & c'_{2(j_1+1)} & \cdots\cdots & c'_{2q} \\ \mathbf{0} & \vdots & \vdots & & \vdots \\ & c'_{pj_1} & c'_{p(j_1+1)} & \cdots\cdots & c'_{pq} \end{array}\right).$$

Nun addiere man das $(-c'_{ij_1}/d_{1j_1})$-Fache der ersten Zeile zur i-ten Zeile, und zwar für $i = 2, ..., p$. Man erhält

$$C'' = \left(\begin{array}{c|cccc} 0 \cdots\cdots 0 & d_{1j_1} & d_{1(j_1+1)} & \cdots\cdots & d_{1q} \\ \hline & 0 & c''_{2(j_1+1)} & \cdots\cdots & c''_{2q} \\ \mathbf{0} & \vdots & \vdots & & \vdots \\ & 0 & c''_{p(j_1+1)} & \cdots\cdots & c''_{pq} \end{array}\right).$$

Es sei j_2 der kleinste der Indizes j, für die es einen Index $i \geq 2$ mit $c''_{ij} \neq 0$ gibt. Wie man der Form von C'' entnimmt, muss $j_2 > j_1$ gelten. Es gelte etwa $c''_{i'j_2} \neq 0$ mit $2 \leq i' \leq p$. Man vertausche nun in der Matrix C'' die zweite mit der i'-ten Zeile. Das Ergebnis ist die Matrix

$$C''' = \left(\begin{array}{c|ccccc} 0 \cdots\cdots 0 & d_{1j_1} & \cdots\cdots & \cdots\cdots & \cdots\cdots & d_{1q} \\ \hline 0 \cdots\cdots 0 \cdots\cdots 0 & d_{2j_2} & \cdots\cdots & \cdots\cdots & \cdots\cdots & d_{2q} \\ \hline & c'''_{3j_2} & c'''_{3(j_2+1)} & \cdots\cdots & c'''_{3q} \\ \mathbf{0} & \vdots & \vdots & & \vdots \\ & c'''_{pj_2} & c'''_{p(j_2+1)} & \cdots\cdots & c'''_{pq} \end{array}\right).$$

Jetzt können wir das $(-c'''_{ij_2}/d_{2j_2})$-Fache der ersten Zeile zu der i-ten Zeile addieren. Dabei durchlaufe i die Werte $3, ..., p$. Dies führt zu der Matrix

$$C'''' = \left(\begin{array}{c|ccccc} 0 \cdots\cdots 0 & d_{1j_1} & \cdots\cdots & \cdots\cdots & \cdots\cdots & d_{1q} \\ \hline 0 \cdots\cdots 0 \cdots\cdots 0 & d_{2j_2} & \cdots\cdots & \cdots\cdots & \cdots\cdots & d_{2q} \\ \hline & 0 & c''''_{3(j_2+1)} & \cdots\cdots & c''''_{3q} \\ \mathbf{0} & \vdots & \vdots & & \vdots \\ & 0 & c''''_{p(j_2+1)} & \cdots\cdots & c''''_{pq} \end{array}\right).$$

So fortfahrend erhält man schließlich die geforderte Matrix D. □

Sobald man einige Beispiele gerechnet hat, wird einem das obige Verfahren sofort verständlich.

Beispiel I.3.6.

$$\begin{pmatrix} 0 & -1 & 5 & 3 & 1 \\ 0 & 2 & -10 & -7 & -4 \\ 4 & 5 & -15 & -11 & -3 \\ 2 & 1 & 0 & -1 & 0 \end{pmatrix} \xrightarrow{\text{I}\longleftrightarrow\text{III}} \begin{pmatrix} 4 & 5 & -15 & -11 & -3 \\ 0 & 2 & -10 & -7 & -4 \\ 0 & -1 & 5 & 3 & 1 \\ 2 & 1 & 0 & -1 & 0 \end{pmatrix}$$

$$\xrightarrow{\text{IV}-\frac{1}{2}\cdot\text{I}} \begin{pmatrix} 4 & 5 & -15 & -11 & -3 \\ 0 & 2 & -10 & -7 & -4 \\ 0 & -1 & 5 & 3 & 1 \\ 0 & -1{,}5 & 7{,}5 & 4{,}5 & 1{,}5 \end{pmatrix}$$

$$\xrightarrow[\text{IV}+\frac{3}{4}\cdot\text{II}]{\text{III}+\frac{1}{2}\cdot\text{II}} \begin{pmatrix} 4 & 5 & -15 & -11 & -3 \\ 0 & 2 & -10 & -7 & -4 \\ 0 & 0 & 0 & -0{,}5 & -1 \\ 0 & 0 & 0 & -0{,}75 & -1{,}5 \end{pmatrix}$$

$$\xrightarrow{\text{IV}-1{,}5\cdot\text{III}} \begin{pmatrix} 4 & 5 & -15 & -11 & -3 \\ 0 & 2 & -10 & -7 & -4 \\ 0 & 0 & 0 & -0{,}5 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Diese Matrix kann noch auf die Form

$$\begin{pmatrix} 1 & 1{,}25 & -3{,}75 & -2{,}75 & -0{,}75 \\ 0 & 1 & -5 & -3{,}5 & -2 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

gebracht werden. Eine andere Möglichkeit ist

$$\begin{pmatrix} 0 & -1 & 5 & 3 & 1 \\ 0 & 2 & -10 & -7 & -4 \\ 4 & 5 & -15 & -11 & -3 \\ 2 & 1 & 0 & -1 & 0 \end{pmatrix} \xrightarrow{\text{I}\longleftrightarrow\text{IV}} \begin{pmatrix} 2 & 1 & 0 & -1 & 0 \\ 0 & 2 & -10 & -7 & -4 \\ 4 & 5 & -15 & -11 & -3 \\ 0 & -1 & 5 & 3 & 1 \end{pmatrix}$$

$$\xrightarrow{\text{III}-2\cdot\text{I}} \begin{pmatrix} 2 & 1 & 0 & -1 & 0 \\ 0 & 2 & -10 & -7 & -4 \\ 0 & 3 & -15 & -9 & -3 \\ 0 & -1 & 5 & 3 & 1 \end{pmatrix}$$

$$\xrightarrow[\text{IV}+\frac{1}{2}\cdot\text{II}]{\text{III}-\frac{3}{2}\cdot\text{II}} \begin{pmatrix} 2 & 1 & 0 & -1 & 0 \\ 0 & 2 & -10 & -7 & -4 \\ 0 & 0 & 0 & 1{,}5 & 3 \\ 0 & 0 & 0 & -0{,}5 & -1 \end{pmatrix}$$

$$\xrightarrow{\text{IV}+\frac{1}{3}\cdot\text{III}} \begin{pmatrix} 2 & 1 & 0 & -1 & 0 \\ 0 & 2 & -10 & -7 & -4 \\ 0 & 0 & 0 & 1{,}5 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Diese Matrix kann noch in die Matrix

$$\begin{pmatrix} 1 & 0,5 & 0 & -0,5 & 0 \\ 0 & 1 & -5 & -3,5 & -2 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

umgeformt werden. Die Zeilenstufenform einer Matrix ist also keineswegs eindeutig bestimmt. Man beachte auch, dass der zweite Weg eine augenscheinlich einfachere Matrix liefert, denn sie enthält mehr Nullen.

Bemerkung I.3.7. Man könnte die Mehrdeutigkeit im Algorithmus in Satz I.3.3, i), die sich durch die Wahl einer Zeile, in der ein Eintrag in der entsprechenden Spalte nicht null ist, ergibt, vermeiden, indem man immer die erste solche Zeile nimmt. Dann zeigt das obige Beispiel, dass sich die Zeilenstufenform ändert, wenn man in der ursprünglichen Matrix Zeilen vertauscht (oder gar andere Zeilenoperationen durchführt). Diese Modifikation des Algorithmus löst damit nicht die prinzipiellen Fragen, die in Problem I.4.6 aufgeworfen werden.

I.4 Das Gaußsche Lösungsverfahren

Gegeben seien eine $(m \times n)$-Matrix

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

sowie

$$\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in K^m.$$

Gesucht ist die Lösungsmenge $\mathbb{L}(A, \mathbf{b})$. Zu ihrer Bestimmung gehe man wie folgt vor. Zunächst bilde man die *erweiterte Koeffizientenmatrix*

$$M := \left(A \,\middle|\, \mathbf{b}\right) := \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right).$$

Die Matrix M ist also die $(m \times (n+1))$-Matrix, die aus A entsteht, indem man $\mathbf{b}$ als $(n+1)$-te Spalte hinzufügt.

Lemma I.4.1. *Es sei* $M = (A|\mathbf{b})$ *wie oben, und die Matrix* $M' = (A'|\mathbf{b}')$ *gehe aus* M *durch eine Zeilenoperation vom Typ* (ZI), (ZII) *oder* (ZIII) *hervor. Dann gilt*

$$\mathbb{L}(A, \mathbf{b}) = \mathbb{L}(A', \mathbf{b}').$$

Beweis. Für Zeilenoperationen vom Typ (ZII) und (ZIII) ist die Aussage klar. Im Fall einer Zeilenoperation vom Typ (ZI) sei $\mathbf{s} \in \mathbb{L}(A, \mathbf{b})$. Wir wollen zeigen, dass auch $\mathbf{s} \in \mathbb{L}(A', \mathbf{b}')$ gilt. Es gehe M' aus M durch Addition des a-Fachen der i-ten Zeile zur j-ten hervor, $a \in K^*$. Die durch M bzw. M' definierten $(m \times n)$-Gleichungssysteme unterscheiden sich nur in der j-ten Gleichung. Wir müssen also lediglich die Gültigkeit der Gleichung

$$a'_{j1} \cdot s_1 + \cdots + a'_{jn} \cdot s_n = b'_j$$

überprüfen. Nach Definition lautet diese Gleichnung

$$(a_{j1} + aa_{i1}) \cdot s_1 + \cdots + (a_{jn} + aa_{in}) \cdot s_n = b_j + ab_i.$$

Diese gilt offensichtlich, da $\mathbf{s} \in \mathbb{L}(A, \mathbf{b})$

$$\begin{aligned} aa_{i1} \cdot s_1 + \cdots + aa_{in} \cdot s_n &= ab_i, \\ a_{j1} \cdot s_1 + \cdots + a_{jn} \cdot s_n &= b_j \end{aligned}$$

beinhaltet. Damit ist $\mathbb{L}(A, \mathbf{b}) \subseteq \mathbb{L}(A', \mathbf{b}')$ nachgewiesen. Da man M aus M' ebenfalls durch eine Zeilenoperation vom Typ (ZI), nämlich die Addition des $(-a)$-Fachen der i-ten Zeile zur j-ten, erhalten kann, folgt auch $\mathbb{L}(A', \mathbf{b}') \subseteq \mathbb{L}(A, \mathbf{b})$. □

Man führe nun an der Matrix M sukzessiv Zeilenoperationen vom Typ (ZI), (ZII) und (ZIII) aus, bis man eine Matrix $M' = (A'|\mathbf{b}')$ erhält, in der sich A' in Zeilenstufenform befindet (s. Satz I.3.3), i.e.

$$M' = \left(\begin{array}{ccccc|c|c} 0 \cdots\cdots 0 & a'_{1j_1} & * \cdots\cdots & \cdots\cdots & * & & b'_1 \\ 0 \cdots\cdots 0 & \cdots\cdots 0 & a'_{2j_2} & * \cdots\cdots & * & & \vdots \\ & 0 & & \ddots & * & * & \vdots \\ 0 \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots 0 & a'_{rj_r} & & b'_r \\ \hline & & & 0 & & & b'_{r+1} \\ & & & & & & \vdots \\ & & & & & & b'_m \end{array}\right).$$

Es folgt.

Satz I.4.2. i) *Das durch* A *und* $\mathbf{b}$ *definierte Gleichungssystem ist genau dann lösbar, d.h.* $\mathbb{L}(A,\mathbf{b}) \neq \varnothing$*, wenn*

$$b'_{r+1} = \cdots = b'_m = 0 \tag{I.6}$$

gilt.

ii) *Wenn* (I.6) *erfüllt ist, dann erhält man jede Lösung*

$$\mathbf{s} = \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} \in \mathbb{L}(A,\mathbf{b}) = \mathbb{L}(A',\mathbf{b}')$$

auf folgende Weise: Für die Einträge s_j *mit* $j \in \{1,\ldots,n\}\backslash\{j_1,\ldots,j_r\}$ *(d.h.* $j \in \{1,\ldots,n\}$ *aber* $j \notin \{j_1,\ldots,j_r\}$*) wähle man beliebige Elemente aus* K*, und* s_{j_1}*,...,*s_{j_r} *bestimme man rekursiv:*

a)

$$s_{j_r} := \frac{1}{a'_{rj_r}} \cdot \Big(b'_r - \sum_{j=j_r+1}^{n} a'_{rj} \cdot s_j\Big).$$

b) *Sind die* s_{j_i} *für* $i > i_0$ *bestimmt, dann ist*

$$s_{j_{i_0}} := \frac{1}{a'_{i_0 j_{i_0}}} \cdot \Big(b'_{i_0} - \sum_{j=j_{i_0}+1}^{n} a'_{i_0 j} \cdot s_j\Big).$$

Beweis. Wir bemerken zunächst, dass die Gleichheit $\mathbb{L}(A,\mathbf{b}) = \mathbb{L}(A',\mathbf{b}')$ eine Konsequenz von Lemma I.4.1 ist. Damit ist die Notwendigkeit der Bedingung (I.6) evident. Ist umgekehrt (I.6) erfüllt, dann liefert das in ii) beschriebene Verfahren tatsächlich Lösungen. Somit ist i) bewiesen. Da schließlich jede Lösung $\mathbf{s} \in \mathbb{L}(A',\mathbf{b}')$ offenbar der angegebenen Rekursionsvorschrift genügen muss, ist auch ii) nachgewiesen. □

Auch hier sagt ein Beispiel wieder mehr als tausend Worte.

Beispiel I.4.3. Es seien

$$A := \begin{pmatrix} 0 & -1 & 5 & 3 & 1 \\ 0 & 2 & -10 & -7 & -4 \\ 4 & 5 & -15 & -11 & -3 \\ 2 & 1 & 0 & -1 & 0 \end{pmatrix}, \quad \mathbf{b} := \begin{pmatrix} 2 \\ -14 \\ -6 \\ 0 \end{pmatrix}, \quad \mathbf{c} := \begin{pmatrix} 2 \\ -14 \\ -5 \\ 0 \end{pmatrix}.$$

Wir können die durch A und $\mathbf{b}$ bzw. A und $\mathbf{c}$ definierten Gleichungssysteme gleichzeitig behandeln, indem wir die Matrix $M := (A|\mathbf{b}|\mathbf{c})$ umformen:

$$\left(\begin{array}{ccccc|c|c} 0 & -1 & 5 & 3 & 1 & 2 & 2 \\ 0 & 2 & -10 & -7 & -4 & -14 & -14 \\ 4 & 5 & -15 & -11 & -3 & -6 & -5 \\ 2 & 1 & 0 & -1 & 0 & 0 & 0 \end{array}\right)$$

$$\xrightarrow{\mathrm{I}\longleftrightarrow\mathrm{IV}} \left(\begin{array}{ccccc|c|c} 2 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 2 & -10 & -7 & -4 & -14 & -14 \\ 4 & 5 & -15 & -11 & -3 & -6 & -5 \\ 0 & -1 & 5 & 3 & 1 & 2 & 2 \end{array}\right)$$

$$\xrightarrow{\mathrm{III}-2\cdot\mathrm{I}} \left(\begin{array}{ccccc|c|c} 2 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 2 & -10 & -7 & -4 & -14 & -14 \\ 0 & 3 & -15 & -9 & -3 & -6 & -5 \\ 0 & -1 & 5 & 3 & 1 & 2 & 2 \end{array}\right)$$

$$\xrightarrow{\mathrm{III}-\frac{3}{2}\cdot\mathrm{II}\ \&\ \mathrm{IV}+\frac{1}{2}\cdot\mathrm{II}} \left(\begin{array}{ccccc|c|c} 2 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 2 & -10 & -7 & -4 & -14 & -14 \\ 0 & 0 & 0 & 1,5 & 3 & 15 & 16 \\ 0 & 0 & 0 & -0,5 & -1 & -5 & -5 \end{array}\right)$$

$$\xrightarrow{\mathrm{IV}+\frac{1}{3}\cdot\mathrm{III}} \left(\begin{array}{ccccc|c|c} 2 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 2 & -10 & -7 & -4 & -14 & -14 \\ 0 & 0 & 0 & 1,5 & 3 & 15 & 16 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} \end{array}\right).$$

Wir erkennen sofort, dass $\mathbb{L}(A, \mathbf{c}) = \varnothing$ gilt. Jetzt bestimmen wir $\mathbb{L}(A, \mathbf{0})$. Nach Satz I.4.2 erhalten wir eine Lösung $\begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{pmatrix}$, indem wir zunächst s_3 und s_5 frei wählen, z.B. $s_3 := t \in \mathbb{R}$ und $s_5 := u \in \mathbb{R}$. Aus der vierten Gleichnung ergibt sich $1,5s_4 + 3u = 0$, d.h. $s_4 = -2u$. Die zweite Gleichung wird zu $2s_2 - 10t + 14u - 4u = 0$, und wir erhalten $s_2 = 5t - 5u$. Schließlich ergibt die erste Gleichung $2s_1 + 5t - 5u + 2u = 0$ und somit $s_1 = -2,5t + 1,5u$. Also

$$\begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{pmatrix} = \begin{pmatrix} -2,5t & + & 1,5u \\ 5t & - & 5u \\ t & & \\ & - & 2u \\ & & u \end{pmatrix} = t \cdot \begin{pmatrix} -\frac{5}{2} \\ 5 \\ 1 \\ 0 \\ 0 \end{pmatrix} + u \cdot \begin{pmatrix} \frac{3}{2} \\ -5 \\ 0 \\ -2 \\ 1 \end{pmatrix},$$

i.e.

$$\mathbb{L}(A,\mathbf{0}) = \left\{ t\cdot\begin{pmatrix} -\frac{5}{2}\\ 5\\ 1\\ 0\\ 0\end{pmatrix} + u\cdot\begin{pmatrix} \frac{3}{2}\\ -5\\ 0\\ -2\\ 1\end{pmatrix} \;\middle|\; t,u\in\mathbb{R}\right\}.$$

Um $\mathbb{L}(A,\mathbf{b})$ zu ermitteln, müssen wir nach Lemma I.2.7 noch eine spezielle Lösung $s^0 \in \mathbb{L}(A,\mathbf{b})$ suchen. Dazu setzen wir $s_3^0 = s_5^0 = 0$. Der Rekursionsvorschrift entnehmen wir $s_4^0 = 10$, $2s_2^0 - 70 = -14$, i.e. $s_2^0 = 28$, $2s_1^0 + 28 - 10 = 0$, d.h. $s_1^0 = -9$. Folglich gilt

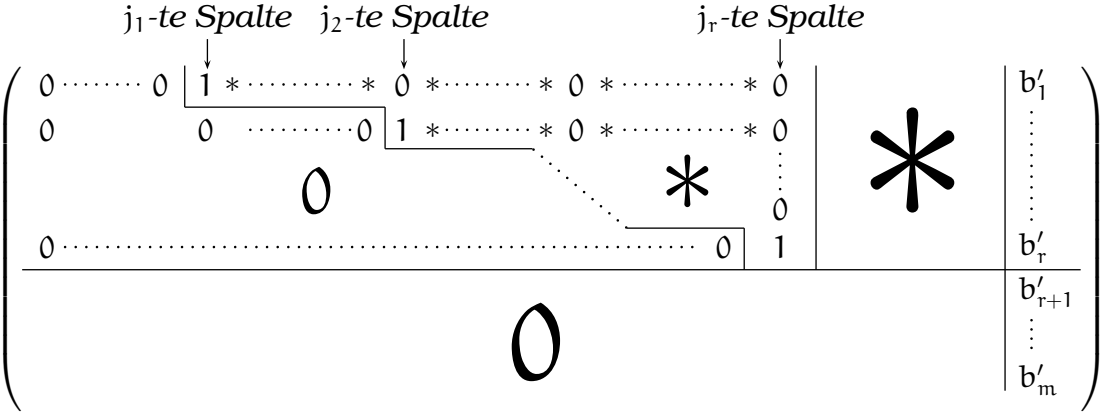
$$\mathbb{L}(A,\mathbf{b}) = \begin{pmatrix} -9\\ 28\\ 0\\ 10\\ 0\end{pmatrix} + \mathbb{L}(A,\mathbf{0}) = \left\{ \begin{pmatrix} -9\\ 28\\ 0\\ 10\\ 0\end{pmatrix} + t\cdot\begin{pmatrix} -\frac{5}{2}\\ 5\\ 1\\ 0\\ 0\end{pmatrix} + u\cdot\begin{pmatrix} \frac{3}{2}\\ -5\\ 0\\ -2\\ 1\end{pmatrix} \;\middle|\; t,u\in\mathbb{R}\right\}.$$

Wir geben jetzt eine weitere Version von Satz I.4.2 an, die für die Entwicklung der Theorie vorteilhaft ist.

Satz I.4.4. *Es seien* $A := (a_{ij})_{\substack{i=1,\dots,m\\ j=1,\dots,n}}$ *eine* $(m\times n)$*-Matrix und*

$$\mathbf{b} := \begin{pmatrix} b_1\\ \vdots\\ b_m\end{pmatrix} \in K^m.$$

Gemäß Satz I.3.3 *und Bemerkung* I.3.5 *überführe man die erweiterte Koeffizientenmatrix* $M = (A|\mathbf{b})$ *durch Zeilenoperationen vom Typ* (ZI), (ZII) *und* (ZIII) *in eine Matrix* $M' = (A'|\mathbf{b}')$ *der Form*



Für $j\in\{1,\dots,n\}\setminus\{j_1,\dots,j_r\}$ *setze man*

$$i(j) := \max\Big\{ i = 1,\dots,r \,|\, j_i < j \Big\}$$

$$\mathbf{v}^j = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -a'_{1j} \\ 0 \\ \vdots \\ 0 \\ -a'_{2j} \\ \vdots \\ -a'_{i(j)j} \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{array}{l} \\ \\ \\ \longleftarrow j_1\text{-ter Eintrag} \\ \\ \\ \\ \longleftarrow j_2\text{-ter Eintrag} \\ \\ \longleftarrow j_{i(j)}\text{-ter Eintrag} \\ \\ \\ \\ \longleftarrow j\text{-ter Eintrag} \\ \\ \\ \end{array}$$

Abbildung I.6: Der Vektor $\mathbf{v}^j$.

und definiere

$$\mathbf{v}^j = \begin{pmatrix} v_1^j \\ \vdots \\ v_n^j \end{pmatrix} \in K^n$$

durch

$$v_k^j := \begin{cases} -a'_{ij}, & k = j_i,\ i = 1, ..., i(j) \\ 1, & k = j \\ 0, & \textit{sonst} \end{cases} \qquad \textit{(s. Abbildung 1.5)}\,.$$

i) *Es sei* $\mathbf{b}' = \mathbf{b} = \mathbf{0}$. *Zu jeder Lösung* $\mathbf{s} \in \mathbb{L}(A, \mathbf{0})$ *gibt es eindeutig bestimmte Zahlen* t_j, $j \in \{1, ..., n\} \setminus \{j_1, ..., j_r\}$, *mit*

$$\mathbf{s} = \sum_{j \in \{1,...,n\} \setminus \{j_1,...,j_r\}} t_j \cdot \mathbf{v}^j. \tag{I.7}$$

Sind umgekehrt Zahlen t_j, $j \in \{1, ..., n\} \setminus \{j_1, ..., j_r\}$, *gegeben und ist* $\mathbf{s}$ *durch* (I.7) *definiert, dann gilt* $\mathbf{s} \in \mathbb{L}(A, \mathbf{0})$.

ii) *Es gelte* $b'_{r+1} = \cdots = b'_m = 0$. *Man definiere*

$$\mathbf{s}^0 = \begin{pmatrix} s_1^0 \\ \vdots \\ s_n^0 \end{pmatrix} \in K^n$$

durch

$$s_j^0 := \begin{cases} b'_i, & j = j_i,\ i = 1, ..., r \\ 0, & \textit{sonst} \end{cases}.$$

Dann gibt es zu jeder Lösung $\mathbf{s} \in \mathbb{L}(A, \mathbf{b})$ *eindeutig bestimmte Zahlen* t_j, $j \in \{1, ..., n\} \setminus \{j_1, ..., j_r\}$, *mit*

$$\mathbf{s} = \mathbf{s}^0 + \sum_{j \in \{1,...,n\} \setminus \{j_1,...,j_r\}} t_j \cdot \mathbf{v}^j. \tag{I.8}$$

Falls umgekehrt Zahlen t_j, $j \in \{1, ..., n\} \setminus \{j_1, ..., j_r\}$, *gegeben sind und* $\mathbf{s}$ *durch* (I.8) *definiert wird, so folgt* $\mathbf{s} \in \mathbb{L}(A, \mathbf{b})$.

Beweis. Wir erinnern daran, dass nach Lemma I.4.1 $\mathbb{L}(A, \mathbf{0}) = \mathbb{L}(A', \mathbf{0})$ und $\mathbb{L}(A, \mathbf{b}) = \mathbb{L}(A', \mathbf{b}')$ gilt.

Zu i). Es gilt $\mathbf{v}^j \in \mathbb{L}(A', \mathbf{0}) = \mathbb{L}(A, \mathbf{0})$, $j \in \{1, ..., n\} \setminus \{j_1, ..., j_r\}$. Da $\mathbb{L}(A, \mathbf{0})$ ein linearer Teilraum von K^n ist (Lemma I.2.4), ist jedes Element, das durch (I.7) definiert wird, eine Lösung des durch A definierten homogenen Gleichungssystems. Für eine gegebene Lösung $\mathbf{s} = \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} \in \mathbb{L}(A, \mathbf{0})$ definieren wir

$$\mathbf{s}' = \sum_{j \in \{1,...,n\} \setminus \{j_1,...,j_r\}} s_j \cdot \mathbf{v}^j.$$

Dann gilt $\mathbf{s}, \mathbf{s}' \in \mathbb{L}(A, \mathbf{0})$ und $s_j = s'_j$ für $j \in \{1, ..., n\} \setminus \{j_1, ..., j_r\}$. Die Rekursionsformel aus Satz I.4.2 impliziert daher, dass $s_j = s'_j$ für alle j gelten muss. Damit ist jede Lösung von der behaupteten Gestalt. Besteht schließlich eine Darstellung (I.7), dann muss $t_j = s_j$, $j \in \{1, ..., n\} \setminus \{j_1, ..., j_r\}$, gelten, so dass die Eindeutigkeit einer solchen Darstellung folgt.

Zu ii). Nach Konstruktion gilt $\mathbf{s}^0 \in \mathbb{L}(A', \mathbf{b}') = \mathbb{L}(A, \mathbf{b})$. Die Behauptung folgt sofort aus der Tatsache $\mathbb{L}(A, \mathbf{b}) = \mathbf{s}^0 + \mathbb{L}(A, \mathbf{0})$ (s. Lemma I.2.7). □

Folgerung I.4.5. i) *Es sei* $n > m$, *d.h. es gebe mehr Unbekannte als Gleichungen,*[5] *dann besitzt das durch die Matrix* A *definierte homogene lineare Gleichungssystem eine nichttriviale Lösung* $\mathbf{s}$ *(i.e.* $\mathbf{s} \neq \mathbf{0}$*).*

ii) *Es sei* $m = n$. *Falls das durch die Matrix* A *definierte homogene lineare Gleichungssystem nur die triviale Lösung* $\mathbf{s} = \mathbf{0}$ *besitzt, dann ist für* ***jedes*** $\mathbf{b} \in K^m$ *das durch* A *und* $\mathbf{b}$ *definierte lineare Gleichungssystem eindeutig lösbar, d.h.* $\mathbb{L}(A, \mathbf{b}) = \{\mathbf{s}^0\}$.

[5] Man sagt, das Gleichungssytem ist *unterbestimmt*

Beweis. Zu i). Es gilt $r \leq m < n$. Daher exisitiert ein Index $j \in \{1, ..., n\} \backslash \{j_1, ..., j_r\}$, und $\mathbf{v}^j$ ist eine nichttriviale Lösung.

Zu ii). Aus der Annahme $\mathbb{L}(A, \mathbf{0}) = \{\mathbf{0}\}$ und i) folgt $r = m = n$, also

$$M' = \left(\begin{array}{ccc|c} 1 & & & b'_1 \\ & 1 & 0 & b'_2 \\ 0 & & \ddots & \vdots \\ & & 1 & b'_m \end{array}\right).$$

Dies beinhaltet offensichtlich

$$\mathbb{L}(A, \mathbf{b}) = \mathbb{L}(A', \mathbf{b}') = \left\{ \begin{pmatrix} b'_1 \\ \vdots \\ b'_m \end{pmatrix} \right\}$$

und damit die Behauptung. □

Satz I.4.4 suggeriert, dass die Dimension von $\mathbb{L}(A, \mathbf{0})$ genau $n - r$ beträgt. Wir formulieren daher folgende Fragestellungen.

Problem I.4.6. i) Ist es möglich, jeder $(m \times n)$-Matrix A eine Zahl r zuzuordnen, die invariant unter Zeilenoperationen vom Typ (ZI), (ZII) und (ZIII) ist, und jedem linearen Unterraum U von K^n eine Zahl d, so dass

- für eine Matrix in Zeilenstufenform r die Anzahl der Stufen ist und
- $d(\mathbb{L}(A, \mathbf{0})) = n - r$ gilt?

Dies würde wie im Fall der (2×2)-Gleichungssysteme den erwarteten Zusammenhang zwischen der Anzahl r der durch A gestellten Bedingungen und der Dimension des Lösungsraumes liefern.

ii) Es sei $U \subseteq K^n$ ein linearer Teilraum. Gibt es eine geeignete $(m \times n)$-Matrix A, für die $U = \mathbb{L}(A, \mathbf{0})$ gilt?

Aufgabe I.4.7.
Es seien folgende Vektoren in $\mathbb{Q}^4$ gegeben:

$$\mathbf{v}_1 := \begin{pmatrix} 1 \\ 1 \\ 2 \\ -1 \end{pmatrix}, \mathbf{v}_2 := \begin{pmatrix} -1 \\ 1 \\ -1 \\ 2 \end{pmatrix}, \mathbf{v}_3 := \begin{pmatrix} 0 \\ -3 \\ 1 \\ -2 \end{pmatrix}, \mathbf{v}_4 := \begin{pmatrix} 1 \\ 0 \\ -1 \\ -1 \end{pmatrix} \text{ und } \mathbf{b} := \begin{pmatrix} 3 \\ 6 \\ 5 \\ -1 \end{pmatrix}.$$

Bestimmen Sie alle Elemente

$$\mathbf{s} = \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{pmatrix} \in \mathbb{Q}^4,$$

für die

$$s_1 \cdot \mathbf{v}_1 + s_2 \cdot \mathbf{v}_2 + s_3 \cdot \mathbf{v}_3 + s_4 \cdot \mathbf{v}_4 = \mathbf{b}$$

gilt.

Aufgabe I.4.8.

a) Bestimmen Sie die Lösungsmengen der folgenden Gleichungssysteme über $\mathbb{Q}$.

$$\text{(i)} \quad \begin{array}{rcrcrcl} x_1 & - & 2x_2 & - & x_3 & = & 1 \\ -2x_2 & + & 4x_2 & + & 2x_3 & = & 2 \end{array}, \qquad \text{(ii)} \quad \begin{array}{rcrcl} x_2 & - & x_3 & = & 2 \\ x_1 & + & x_2 & = & 5 \\ x_1 & + & x_3 & = & 0 \end{array},$$

$$\text{(iii)} \quad \begin{array}{rcrcrcl} x_1 & + & 2x_2 & - & 3x_3 & = & 2 \\ x_1 & + & 4x_2 & + & x_3 & = & 4 \\ x_1 & + & \frac{5}{2}x_2 & - & 2x_3 & = & \frac{5}{2} \\ x_1 & + & 3x_2 & - & x_3 & = & 3 \end{array}.$$

b) Es seien

$$\text{(i)} \quad A := \begin{pmatrix} 1 & -2 & -1 & 1 \\ -2 & 4 & 2 & 1 \end{pmatrix}, \quad \mathbf{b} := \begin{pmatrix} 1 \\ 2 \end{pmatrix};$$

$$\text{(ii)} \quad A := \begin{pmatrix} 2 & 4 & 0 & 2 & 4 \\ 1 & 1 & -1 & 3 & 2 \\ 0 & 3 & 3 & -6 & 0 \\ 1 & 2 & 0 & 1 & 2 \\ 2 & 6 & 2 & -2 & 4 \end{pmatrix}, \quad \mathbf{b} := \begin{pmatrix} 10 \\ 4 \\ 3 \\ 5 \\ 12 \end{pmatrix}.$$

Berechnen Sie $\mathbb{L}(A, \mathbf{0})$ und $\mathbb{L}(A, \mathbf{b})$ in den angegebenen Fällen und schreiben Sie $\mathbb{L}(A, \mathbf{b}) = \mathbf{s}^0 + \mathbb{L}(A, \mathbf{0})$ für geeignetes $\mathbf{s}^0 \in \mathbb{L}(A, \mathbf{b})$.

c) Gegeben seien folgende 3-Tupel im $\mathbb{R}^3$:

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

Bestimmen Sie reelle Zahlen s_1, s_2 und s_3 mit

$$s_1 \cdot \mathbf{v}_1 + s_2 \cdot \mathbf{v}_2 + s_3 \cdot \mathbf{v}_3 = \mathbf{b}.$$

Aufgabe I.4.9.

Es seien $K = \mathbb{Q}$ und

$$A := \begin{pmatrix} 1 & 2 & 0 & 3 \\ 2 & 3 & 1 & 6 \\ -1 & 2 & 1 & 2 \\ 4 & -2 & 1 & 3 \end{pmatrix}, \quad \mathbf{b}_{u,v} := \begin{pmatrix} 0 \\ 0 \\ u \\ v \end{pmatrix}.$$

Welche Bedingungen müssen u und v erfüllen, damit $\mathbb{L}(A, \mathbf{b}_{u,v}) \neq \varnothing$ gilt? Berechnen Sie $\mathbb{L}(A, \mathbf{b}_{5,-9})$.

Aufgabe I.4.10.
In der Vorlesung wurden die Zeilenoperationen von Typ (ZI), (ZII) und (ZIII) angegeben. In Analogie definiert man die *Spaltenoperationen vom Typ*

(SI): Addition eines Vielfachen einer Spalte zu einer anderen Spalte.
(SII): Vertauschung zweier Spalten.
(SIII): Multiplikation einer Spalte mit einem Element $a \in K^*$.

a) Illustrieren Sie die Spaltenoperationen vom Typ (SI), (SII) und (SIII) an jeweils einem Beispiel.
b) Zeigen Sie, dass man eine $(m \times n)$-Matrix $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$ durch Zeilenoperationen vom Typ (ZI), (ZII), (ZIII) und Spaltenoperationen vom Typ (SII) in eine Matrix der Gestalt

$$A' = \left(\begin{array}{ccccc|c} 1 & 0 & \cdots & \cdots & 0 & \\ 0 & 1 & 0 & \cdots & 0 & \\ \vdots & 0 & \ddots & \ddots & \vdots & * \\ \vdots & \vdots & \ddots & \ddots & 0 & \\ 0 & 0 & \cdots & 0 & 1 & \\ \hline 0 & \cdots & \cdots & \cdots & 0 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & \end{array}\right)$$

überführen kann. Entwickeln Sie aus dieser Beobachtung ein Lösungsverfahren für lineare Gleichungssysteme und testen Sie es für

$$A := \begin{pmatrix} 0 & -1 & 5 & 3 & 1 \\ 0 & 2 & -10 & -7 & -4 \\ 4 & 5 & -15 & -11 & -3 \\ 2 & 1 & 0 & -1 & 0 \end{pmatrix} \quad \text{und} \quad \mathbf{b} := \begin{pmatrix} 2 \\ -14 \\ -6 \\ 0 \end{pmatrix}.$$

(Spaltenvertauschungen können die Lösungsmenge ändern! Wie?)
c) Zeigen Sie: Lässt man zusätzlich auch Spaltenoperationen vom Typ (SI) zu, dann kann man

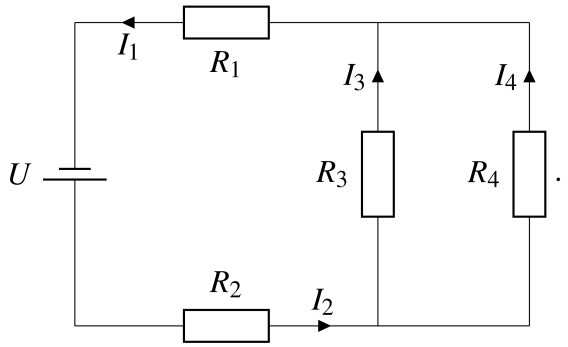
$$A' = \left(\begin{array}{ccccc|c} 1 & 0 & \cdots & \cdots & 0 & \\ 0 & 1 & 0 & \cdots & 0 & \\ \vdots & 0 & \ddots & \ddots & \vdots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 & \\ 0 & 0 & \cdots & 0 & 1 & \\ \hline 0 & \cdots & \cdots & \cdots & 0 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & \end{array}\right)$$

erreichen.
d) Zeigen Sie, dass man eine Zeilenoperation vom Typ (ZII) durch eine gewisse Folge von Zeilenoperationen vom Typ (ZI) und (ZIII) erhalten kann.

Aufgabe I.4.11.
Die **Kirchhoffschen**[6] **Regeln** für ein elektrisches Gleichstromnetzwerk lauten:

- **1. Kirchhoffsche Regel (Knotenregel):** In einem Knotenpunkt des Netzwerks ist die Summe der zufließenden Ströme gleich der Summe der abfließenden Ströme.
- **2. Kirchhoffsche Regel (Maschenregel):** Die Summe aller Teilspannungen in einer Masche des Netzwerks ist null.[7]

Man betrachte das Netzwerk:



Berechnem Sie die Stromstärken in diesem Netzwerk im Fall, dass die Spannungsquelle 36V hat und die Widerstände die Werte $R_1 = 200\Omega$, $R_2 = 400\Omega$, $R_3 = 300\Omega$ und $R_4 = 200\Omega$.

[6] Gustav Robert Kirchhoff (1824 - 1887), deutscher Physiker.
[7] Man beachte, dass Spannung, Widerstand und Stromstärke über die **Ohmsche Gleichung** $U = R \cdot I$ verbunden sind. (Georg Simon Ohm (1789 - 1854), deutscher Physiker.)

II

Grundbegriffe

In diesem Kapitel entwickeln wir die abstrakte Sprache, in der die Lineare Algebra und weite Teile der Mathematik formuliert werden. Der Leser sollte sich möglichst früh an diese Sprache gewöhnen.

II.1 Mengen und Abbildungen

Die Theorie der Mengen und Abbildungen bildet das Fundament der Mathematik. Alle Objekte, die wir betrachten werden, sind zunächst einmal Mengen. Auf diesen Mengen werden mit Hilfe von Abbildungen zwischen Mengen zusätzliche Strukturen wie die einer Gruppe, eines Körpers oder eines Vektorraums erklärt. Es mag daher besonders störend erscheinen, dass gerade die präzise Formulierung der Mengenlehre große Probleme bereitet.

Im Jahre 1895 schlug Cantor[1] folgende „Definition“ für den Mengenbegriff vor: *„Unter einer „Menge“ verstehen wir jede Zusammenfassung* M *von bestimmten wohlunterschiedenen Objekten* m *unserer Anschauung oder unseres Denkens (welche Elemente von* M *genannt werden) zu einem Ganzen.“*

Natürlich ist dies keine mathematische Definition. Von einer mathematischen Definition verlangen wir, dass sie neue Begriffe unter Verwendung uns bereits bekannter mathematischer Konzepte einführt. „Anschauung“, „Denken“ oder auch „Zusammenfassung“ sind sicherlich keine solchen Konzepte. Es ist aber auch klar, dass wir mit dieser Forderung an eine Definition die Grundbausteine der Mathematik, also z.B. Mengen, gar nicht definieren können. (Irgendwo muss man schließlich einmal anfangen!) Man kann sich bloß, wie mit

[1] Georg Ferdinand Ludwig Philipp Cantor (1845 - 1918), deutscher Mathematiker.

Cantors Definition, einige Charakteristika der Mengen klarmachen. Der Ausweg aus dieser verfahrenen Situation ist die axiomatische Mengenlehre. Hier versucht man nicht, den Begriff „Menge“ zu definieren, sondern legt lediglich die Spielregeln (d.h. die Axiome) für Mengen fest. (Eines dieser Axiome besagt z.B., dass überhaupt eine Menge existiert.) Die in den Axiomen formulierten Eigenschaften nimmt man dann als wahre Tatsachen an und baut die Mathematik darauf auf. In der Vorlesung gehen wir einen ähnlichen Weg, allerdings in wesentlich naiverer Form. Wir gehen von einigen uns bekannten Mengen wie den Mengen der natürlichen, ganzen oder rationalen Zahlen aus und geben einige Vorschriften an, mit denen man aus gegebenen Mengen neue konstruieren kann. Dem Leser, der sich für den axiomatischen Aufbau der Mengenlehre interessiert, sei empfohlen, sich zunächst, etwa durch die Vorlesung „Lineare Algebra“, an die Sprache der Mathematik zu gewöhnen und sich in einem späteren Semester an die Mengenlehre zu wagen. Einen guten ersten Eindruck vermittelt dann das Buch [4]. Einige weitere Angaben sind auch in [8], Abschnitt 1.2, enthalten.

Mengen setzen sich aus ihren Elementen zusammen, und wir schreiben kurz „$x \in M$“ für „x ist Element der Menge M“. Mengen sind in folgendem Sinne durch ihre Elemente bestimmt:

Extensionalitätsaxiom. *Zwei Mengen A und B sind genau dann gleich, wenn sie dieselben Elemente besitzen, d.h.*

$$A = B \iff \Big(\forall x : x \in A \iff x \in B\Big).$$

Die Formel auf der rechten Seite wird dabei wie folgt gelesen: „Für alle x ($\forall x$) gilt: x ist Element von A genau dann, wenn ($\iff$) x Element von B ist.“ Ohne dieses Axiom könnten wir mit Mengen überhaupt nicht arbeiten. Da wir aber Mengen nicht definieren können, haben wir auch keine Möglichkeit, die genannte Eigenschaft zu beweisen. Wir müssen sie daher als eine unserer Spielregeln akzeptieren. Gängige Beispiele für Mengen sind:

Beispiele II.1.1. i) Die *leere Menge* $\varnothing$ ist die Menge, die kein einziges Element enthält ($\forall x : x \notin \varnothing$). Nach dem Extensionalitätsaxiom ist sie die einzige Menge mit dieser Eigenschaft.

ii) Die *Menge der natürlichen Zahlen* ist die Menge

$$\mathbb{N} = \Big\{\, 0, 1, 2, 3, \ldots \Big\}.$$

iii) Weitere bekannte Mengen sind die Mengen $\mathbb{Z}$, $\mathbb{Q}$ und $\mathbb{R}$ der *ganzen*, *rationalen* bzw. *reellen Zahlen*.

Definition II.1.2. Es seien A und B Mengen. Wir sagen, A *ist Teilmenge* (oder auch *Untermenge*) *von* B, wenn jedes Element von A auch Element von B ist.

Formal:

$$\forall x : x \in A \Longrightarrow x \in B.$$

Schreibweise. $A \subseteq B$. Falls klar ist, dass $A \neq B$, schreiben wir auch $A \subset B$ oder $A \subsetneq B$.

Bemerkung II.1.3. Zum Beweis der Gleichheit zweier Mengen benutzt man oft die folgende Eigenschaft:

$$A = B \Longleftrightarrow \big(A \subseteq B \wedge B \subseteq A\big).$$

Dabei steht „$\wedge$“ für „und“.

Beispiel II.1.4. Es gilt z.B. $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.

Konstruktion neuer Mengen. — Um mit Hilfe des obigen Beispielvorrats weitere Mengen konstruieren zu können, geben wir in den folgenden Definitionen einige Vorschriften an, die dies ermöglichen. Man beachte, dass wir dabei stillschweigend annehmen, dass diese Bildungsvorschriften in der Tat wieder Mengen liefern.

Definitionen II.1.5. i) (*Aufzählende Form*). Endliche Mengen kann man durch die Aufzählung ihrer Elemente angeben, z.B.

$$A := \Big\{ 1, 2, 3, 5, 223 \Big\} \quad \text{und} \quad B := \Big\{ 13, \varnothing, \mathbb{N} \Big\}.$$

ii) (*Komprehension*). Es seien A eine Menge und $\mathcal{E}$ eine Eigenschaft,[2] die Elemente der Menge A haben können. Dann kann man die Menge

$$\Big\{ x \,\Big|\, x \in A \text{ und } x \text{ hat } \mathcal{E} \Big\} = \Big\{ x \in A \,\Big|\, x \text{ hat } \mathcal{E} \Big\}$$

bilden.

Bemerkungen und Beispiele II.1.6. i) Bei der aufzählenden Form kommt es nicht auf die Reihenfolge an. Außerdem sind Wiederholungen erlaubt:

$$\Big\{ 1, 3, 5 \Big\} = \Big\{ 5, 1, 3 \Big\} = \Big\{ 5, 3, 1 \Big\},$$

[2] Um ganz präzise zu sein, müssten wir auch erklären, was „Eigenschaft“ bedeutet. Prinzipiell müssten solche Eigenschaften in der Sprache der Logik formuliert werden. Dort verwendet man Formeln, die aus den logischen Symbolen $\wedge$, $\vee$, $\neg$, ..., den Quantoren $\forall$, $\exists$, Buchstaben und, im Fall der Mengenlehre, dem Symbol $\in$ nach gewissen Regeln aufgebaut sind (s. [12], Kapitel 1). Die Eigenschaften sind dann von der Form, dass eine gewisse Formel wahr ist. Wir werden im Folgenden Eigenschaften weniger formal beschreiben. Im Laufe des Kurses sollte aber deutlich werden, dass sich die Eigenschaften, die wir benutzen, präzise „formalisieren“ lassen.

$$\{7,7,8,8,8\} = \{7,8\} = \{7,8,7,7,8,8\}.$$

Diese Beobachtung ist eine Folge des Extensionalitätsaxioms.

ii) Mit Hilfe der Komprehension lassen sich z.B. folgende Mengen bilden:

a) $\{x \in \mathbb{N} \mid x \text{ ist gerade}\} = \{0,2,4,6,\ldots\}$.

b) $\{x \in \mathbb{N} \mid x \text{ ist Primzahl}\} = \{2,3,5,7,11,13,\ldots\}$.

c) $\mathbb{R}_{>0} := \{x \in \mathbb{R} \mid x > 0\}$.

d) Für jede Menge A gilt: $\varnothing = \{x \in A \mid x \neq x\}$.

Definitionen II.1.7. Es seien A und B Mengen.

i) Die *Vereinigung der Mengen* A *und* B ist die Menge

$$A \cup B := \{x \mid x \in A \vee x \in B\}.$$

(Dabei wird „$\vee$“ als „oder“ gelesen.)

ii) Der *Durchschnitt der Mengen* A *und* B ist die Menge

$$A \cap B := \{x \mid x \in A \wedge x \in B\}.$$

iii) Die *Differenz von* A *und* B ist die Menge

$$A \setminus B := \{x \mid x \in A \wedge x \notin B\}.$$

Beispiel II.1.8. Es gilt z.B. $\mathbb{Z} \setminus \mathbb{N} = \{x \in \mathbb{Z} \mid x < 0\}$.

Lemma II.1.9. *Für die Operationen „$\cup$“ und „$\cap$“ gelten die* Distributivgesetze*, d.h. für drei Mengen* A*,* B *und* C *hat man:*

a) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

b) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

Beweis. Die fundamentalen Eigenschaften der Mengenoperationen sind oftmals nur Gegenstücke einfacher Gesetzmäßigkeiten der Logik. Die Logik ist normalerweiser fest in uns „eingebaut“ und muss lediglich ein wenig trainiert werden. Als Beispiel geben wir den Beweis von a) an. Es gilt:

$$\begin{aligned} x \in A \cup (B \cap C) &\iff x \in A \vee (x \in B \wedge x \in C) \\ &\iff (x \in A \vee x \in B) \wedge (x \in A \vee x \in C) \\ &\iff (x \in A \cup B) \wedge (x \in A \cup C) \\ &\iff x \in (A \cup B) \cap (A \cup C). \end{aligned}$$

Nach dem Extensionalitätsaxiom II.1 sind $A \cup (B \cap C)$ und $(A \cup B) \cap (A \cup C)$ gleich. □

Definition II.1.10. Es seien A und B Mengen. Das *kartesische*[3] *Produkt von* A *und* B ist die Menge

$$A \times B := \Big\{ (x, y) \,\Big|\, x \in A \wedge y \in B \Big\}.$$

Ein Element $(x, y) \in A \times B$ heißt ein *geordnetes Paar*.

Bemerkung II.1.11. Die charakteristische Eigenschaft der geordneten Paare ist:

$$(x_1, y_1) = (x_2, y_2) \Longleftrightarrow x_1 = x_2 \wedge y_1 = y_2.$$

Definition II.1.12. Es sei A eine Menge. Die *Potenzmenge von* A ist die Menge

$$\mathcal{P}(A) := \Big\{ B \,\Big|\, B \subseteq A \Big\}.$$

Beispiele II.1.13. i) Für jede Menge A gilt $\varnothing \in \mathcal{P}(A)$ und $A \in \mathcal{P}(A)$.
ii) $\mathcal{P}\big(\{0, 1\}\big) = \Big\{ \varnothing, \{0\}, \{1\}, \{0, 1\} \Big\}$.

Aufgabe II.1.14.
a) Es seien A, B und C Mengen. Beweisen Sie die *Distributivgesetze für den Durchschnitt und die Vereinigung*:

$$\begin{aligned} A \cap (B \cup C) &= (A \cap B) \cup (A \cap C), \\ A \cup (B \cap C) &= (A \cup B) \cap (A \cup C). \end{aligned}$$

b) A und B seien Mengen. Zeigen Sie, dass folgende Aussagen äquivalent sind:

i) $A \subseteq B$.

ii) $A \cup B = B$.

iii) $A \cap B = A$.

c) Es seien A und B Mengen. Die Menge

$$A \triangle B := (A \setminus B) \cup (B \setminus A)$$

heißt *die symmetrische Mengendifferenz von* A *und* B. Beweisen Sie:

$$A \triangle B = (A \cup B) \setminus (A \cap B).$$

[3] Nach René Descartes (1596 - 1650), französischer Philosoph, Mathematiker und Naturwissenschaftler.

d) Es sei Ω eine feste Menge. Für eine Teilmenge $A \subseteq \Omega$ ist das *Komplement von* A *(in* Ω*)* die Differenz $\overline{A} := \Omega \setminus A$. Beweisen Sie für Teilmengen $A, B \subset \Omega$ die *De Morganschen*[4] *Gesetze*

$$\overline{A \cup B} = \overline{A} \cap \overline{B}, \quad \overline{A \cap B} = \overline{A} \cup \overline{B}$$

sowie

$$\overline{\overline{A}} = A.$$

Aufgabe II.1.15.
Es sei $A = \{x_1, ..., x_n\}$ eine endliche Menge mit genau n Elementen. Man schreibt auch $\#A = n$. Beweisen Sie für die Potenzmenge $P(A)$ von A

$$\#P(A) = 2^n.$$

(**Hinweis.** Verwenden Sie vollständige Induktion.)
Zum Prinzip der vollständigen Induktion: Eines der **Peano-Axiome** (s. [8], Definition 1.3.1, i) für die natürlichen Zahlen sagt aus, dass eine Teilmenge $A \subseteq \mathbb{N}$ der natürlichen Zahlen, für die

$$0 \in A \quad \text{und} \quad \big(n \in A \Longrightarrow n + 1 \in A\big)$$

gilt, bereits die Menge der natürlichen Zahlen selbst ist, i.e. $A = \mathbb{N}$. Darauf beruht das **Prinzip der vollständigen Induktion**: Es sei $\mathcal{A}(n)$, $n \in \mathbb{N}$, eine Familie von Aussagen, hier $\mathcal{A}(n)$=„Die Potenzmenge einer Menge mit genau n Elementen hat genau 2^n Elemente“. Nach dem genannten Peano-Axiom ist $\mathcal{A}(n)$ genau dann für alle natürlichen Zahlen n wahr, wenn gilt 1. $\mathcal{A}(0)$ ist wahr (*Induktionsanfang*) und 2. Ist $\mathcal{A}(n)$ wahr, dann ist auch $\mathcal{A}(n+1)$ wahr (*Induktionsschritt*). (Dazu betrachte man die Menge $A := \{n \in \mathbb{N} \mid \mathcal{A}(n) \text{ ist wahr}\}$.)

Abbildungen. — Um im Weiteren komplexere Objekte[5] als Mengen einführen zu können, benötigen wir Abbildungen zwischen Mengen.

Definitionen II.1.16. Es seien A und B Mengen.

i) Eine *Abbildung von* A *nach* B ist eine Vorschrift f, die jedem Element $x \in A$ (genau) ein Element $y \in B$ zuordnet. Dabei nennt man y das *Bild von* x (*unter der Abbildung* f).

Bezeichnung. a) $y = f(x)$.

b) $f\colon A \longrightarrow B$
$x \longmapsto f(x)$.

[4] Augustus De Morgan (1806 - 1871), englischer Mathematiker.

[5] Die Objekte sind streng genommen nur scheinbar komplexer als Mengen, weil alle Objekte, die wir in der Mathematik betrachten, letztendlich Menge sein müssen. Wir stellen uns aber z.B. eine Gruppe als Menge mit Zusatzstruktur vor. Die Zusatzstruktur ist die Multiplikation.

ii) Es seien A und B Mengen und $f\colon A \longrightarrow B$ eine Abbildung von A nach B. Für eine Teilmenge $C \subseteq B$ nennt man die Menge

$$f^{-1}(C) := \left\{ x \in A \,\middle|\, f(x) \in C \right\}$$

das *Urbild von* C (*unter* f). Ein Element $x \in f^{-1}(\{y\})$, $y \in B$, nennt man auch ***ein*** *Urbild von* y.
Schreibweise. $f^{-1}(y) := f^{-1}\big(\{y\}\big)$, $y \in B$.

iii) Für eine Teilmenge $D \subseteq A$ heißt die Menge

$$f(D) := \left\{ y \in B \,\middle|\, \exists x \in D : y = f(x) \right\} = \left\{ f(x) \,\middle|\, x \in D \right\}$$

das *Bild von* D (*unter* f). (Das Symbol „$\exists$“ wird „es existiert“ oder „es gibt“ gelesen.)
Speziell. Die Menge $\mathrm{Im}(f) := \mathrm{Bild}(f) := f(A)$ bezeichnet man als das *Bild von* f.

Bemerkungen II.1.17. i) Man beachte, dass zwei Abbildungen $f, g\colon A \longrightarrow B$ genau dann gleich sind, wenn $f(x) = g(x)$ für alle $x \in A$ gilt.

ii) Man mag einwenden, dass die Definition von Abbildung ebenfalls keine mathematisch korrekte Definition ist, da wir nicht wissen, was eine „Vorschrift“ sein soll. Dies lässt sich aber mit den uns bekannten Begriffen leicht beheben. Für eine Abbildung $f\colon A \longrightarrow B$ können wir nämlich den *Graphen*

$$\Gamma_f := \left\{ (x, y) \in A \times B \,\middle|\, y = f(x) \right\}$$

betrachten. Diese Teilmenge erfüllt die folgenden zwei Bedingungen:
a) $\forall x \in A \exists y \in B : (x, y) \in \Gamma_f$.
b) $\forall (x_1, y_1), (x_2, y_2) \in \Gamma_f : x_1 = x_2 \Longrightarrow y_1 = y_2$.
Benutzen wir das Symbol „$\exists!$“ („Es existiert/gibt **genau** ein“), dann können wir a) und b) zusammenfassen in

$$\forall x \in A \exists! y \in B : (x, y) \in \Gamma_f. \tag{II.1}$$

Wenn wir nun umgekehrt eine Teilmenge $\Gamma \subseteq A \times B$ vorgeben, für die Bedingung (II.1) erfüllt ist, dann definieren wir

$$\begin{array}{rcl} f\colon A & \longrightarrow & B \\ x & \longmapsto & y, \text{ so dass } (x, y) \in \Gamma. \end{array}$$

Damit kann man eine Abbildung formal als eine Teilmenge Γ von $A \times B$, für die (II.1) gilt, einführen. Die in i) vermerkte Eigenschaft wird somit zum Spezialfall des Extensionalitätsaxioms.

Definition II.1.18. Es seien A und B Mengen. Wir setzen

$$\mathrm{Abb}(A,B) := \Big\{ f\colon A \longrightarrow B \,|\, f \text{ ist Abbildung} \Big\}.$$

In der Sichtweise von Bemerkung II.1.17, ii), können wir die Menge $\mathrm{Abb}(A, B)$ mit Hilfe der Komprehension (Definition II.1.5, ii) als Teilmenge der Potenzmenge $\mathcal{P}(A \times B)$ (Definition II.1.12) von $A \times B$ definieren.

Beispiele II.1.19. i) Es sei A eine Menge. Die *identische Abbildung* (oder *Identität*) *auf* A ist die Abbildung

$$\begin{array}{rcl} \mathrm{Id}_A\colon A & \longrightarrow & A \\ x & \longmapsto & x. \end{array}$$

Dabei ist

$$\Gamma_{\mathrm{Id}_A} = \Big\{ (x,x) \in A \times A \,\Big|\, x \in A \Big\}$$

die *Diagonale in* $A \times A$.

ii) Es seien A und B Mengen und $y_0 \in B$. Damit definiert man die konstante Abbildung

$$\begin{array}{rcl} k_{y_0}\colon A & \longrightarrow & B \\ x & \longmapsto & y_0. \end{array}$$

Offenbar gilt $\mathrm{Bild}(f) = \{y_0\}$, und für eine Teilmenge $C \subseteq B$ hat man

$$f^{-1}(C) = \begin{cases} A, & y_0 \in C \\ \varnothing, & y_0 \notin C \end{cases}.$$

iii) Für eine Zahl $m \in \mathbb{N}$ ist die *Multiplikation mit* m gegeben durch:

$$\begin{array}{rcl} \mu_m\colon \mathbb{N} & \longrightarrow & \mathbb{N} \\ x & \longmapsto & m \cdot x. \end{array}$$

Für $m \neq 0$ ist $\mathrm{Bild}(\mu_m)$ die Menge der durch m teilbaren Zahlen.

Definition II.1.20. Es seien A eine Menge und I eine Menge, die wir als Indexmenge auffassen wollen. Eine (*durch* I *indizierte*) *Familie von Elementen aus* A ist eine Abbildung $f\colon I \longrightarrow A$, $i \longmapsto a_i$, die man gewöhnlicherweise in der Form $(a_i)_{i\in I}$, notiert.[6] Anstatt A kann man auch die Potenzmenge $\mathcal{P}(A)$ wählen. Eine Familie A_i, $i \in I$, mit $A_i \in \mathcal{P}(A)$, i.e. $A_i \subseteq A$, ist dann eine (*durch* I *indizierte*) *Familie von Teilmengen von* A.

Schreibweise. Eine Familie von Teilmengen schreibt man auch in der Form $(A_i)_{i\in I}$.

[6] Diese Notation ist insbesondere für Folgen in der Analysis gebräuchlich (s. [8], Definition 2.1.1).

Eigenschaften von Abbildungen. — In der folgenden Definition formulieren wir einige wichtige Eigenschaften, die Abbildungen haben können und die in der Vorlesung wiederholt von Bedeutung sein werden.

Definitionen II.1.21. Es seien A und B Mengen, und $f\colon A \longrightarrow B$ sei eine Abbildung von A nach B.

i) Die Abbildung f heißt *surjektiv*, wenn $f(A) = B$ gilt, d.h. wenn für jedes Element $y \in B$ ein Element $x \in A$ existiert mit $y = f(x)$.

ii) Man nennt f *injektiv*, wenn für alle $x_1, x_2 \in A$ mit $f(x_1) = f(x_2)$ bereits $x_1 = x_2$ folgt.

iii) Die Abbildung f ist *bijektiv*, wenn sie sowohl injektiv als auch surjektiv ist.

Beispiele II.1.22. i) Für endliche Mengen $A = \{x_1, ..., x_m\}$ und $B = \{y_1, ..., y_n\}$ kann man sich eine Abbildung $f\colon A \longrightarrow B$ als eine Ansammlung von Pfeilen vorstellen, die von Elementen von A ausgehen und in Elementen von B enden. Die Definition von Abbildung verlangt, dass von jedem Element von A genau ein Pfeil ausgeht. Die Bedingung „surjektiv“ bedeutet, dass in jedem Element von B mindestens ein Pfeil ankommt, die Bedingung „injektiv“, dass höchstens ein Pfeil ankommt, und „bijektiv“, dass genau ein Pfeil ankommt (vgl. Abbildung II.1).

ii) Die Abbildung $\mathrm{Id}_A\colon A \longrightarrow A$ ist für jede Menge A bijektiv.

iii) Die Abbildung $\mu_1\colon \mathbb{N} \longrightarrow \mathbb{N}$ ist die identische Abbildung. Für $m > 1$ ist μ_m injektiv, denn für natürliche Zahlen n_1, n_2 ist $\mu_m(n_1) = m \cdot n_1 = m \cdot n_2 = \mu_m(n_2)$ gleichbedeutend mit der Tatsache, dass $m \cdot (n_1 - n_2) = m \cdot n_1 - m \cdot n_2 = 0$. Da $m \neq 0$, muss $n_1 - n_2 = 0$ und damit $n_1 = n_2$ gelten. Die Abbildung μ_m ist für $m > 1$ nicht surjektiv, denn $1 \notin \mu_m(\mathbb{N})$.

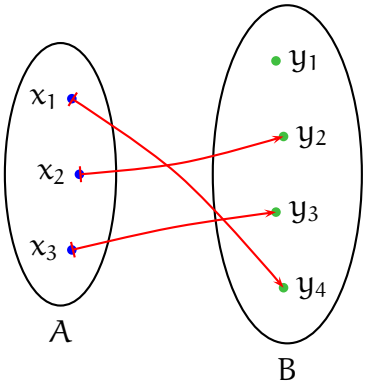
iv) Die Abbildung $f\colon \mathbb{R} \longrightarrow \mathbb{R}$, $x \longmapsto x^3 - 4x^2 + 5x - 2 = (x-1)^2(x-2)$, ist nach dem Zwischenwertsatz der Analysis ([8], Folgerung 3.5.2) surjektiv. Sie ist aber nicht injektiv, da z.B. $f(1) = 0 = f(2)$ gilt (s. Abbildung II.2).

Definition II.1.23. Es seien A, B und C Mengen und $f\colon A \longrightarrow B$ und $g\colon B \longrightarrow C$ Abbildungen. Die *Verknüpfung* oder *Komposition von* f *mit* g ist die Abbildung

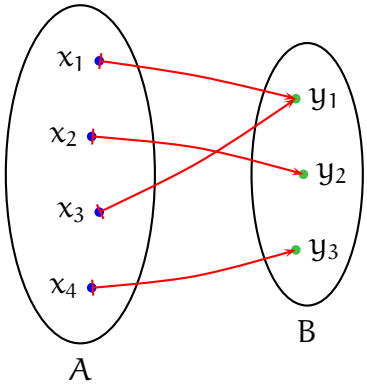
$$\begin{array}{rcl} g \circ f\colon A & \longrightarrow & C \\ x & \longmapsto & g\Big(f(x)\Big). \end{array}$$

Vorsicht. Bei der Verknüpfung von Abbildungen müssen wir ausnahmsweise von rechts nach links lesen: In $g \circ f$ wird zuerst f und dann g ausgeführt.

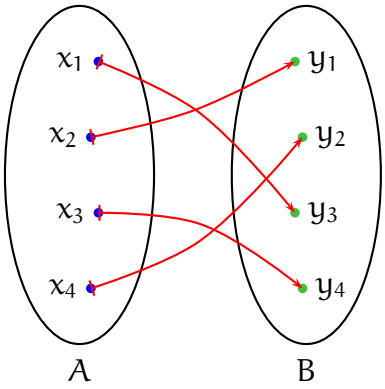
Beispiele II.1.24. i) Es seien A eine Menge und $f\colon A \longrightarrow A$ eine Abbildung. Dann gilt $f \circ \mathrm{Id}_A = f = \mathrm{Id}_A \circ f$.



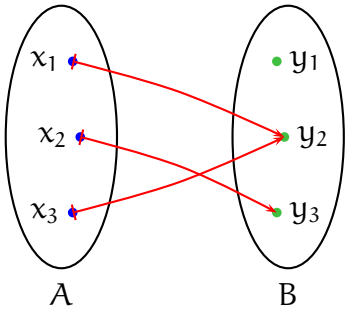
Injektiv aber nicht surjektiv.



Surjektiv aber nicht injektiv.



Bijektiv.



Weder surjektiv noch injektiv.

Abbildung II.1: Die Eigenschaften von Abbildungen.

ii) Es sei $A = \{1, 2, 3\}$. Wir definieren die (bijektiven) Abbildungen $f\colon A \longrightarrow A$, $1 \longmapsto 1$, $2 \longmapsto 3$, $3 \longmapsto 2$, und $g\colon A \longrightarrow A$, $1 \longmapsto 2$, $2 \longmapsto 1$, $3 \longmapsto 3$. Dann erhalten wir

$$\begin{aligned} g \circ f\colon A &\longrightarrow A, \quad 1 \longmapsto 2, 2 \longmapsto 3, 3 \longmapsto 1, \\ f \circ g\colon A &\longrightarrow A, \quad 1 \longmapsto 3, 2 \longmapsto 1, 3 \longmapsto 2. \end{aligned}$$

Insbesondere erkennen wir $f \circ g \neq g \circ f$.

Satz II.1.25. i) *Es sei* $f\colon A \longrightarrow B$ *eine bijektive Abbildung zwischen den Mengen* A *und* B. *Dann gibt es genau eine Abbildung* $g\colon B \longrightarrow A$ *mit*

$$g \circ f = \mathrm{Id}_A \quad \textit{und} \quad f \circ g = \mathrm{Id}_B. \tag{II.2}$$

ii) *Für Mengen* A, B, C *und* D *und Abbildungen* $f\colon A \longrightarrow B$, $g\colon B \longrightarrow C$ *und* $h\colon C \longrightarrow D$ *gilt:*

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

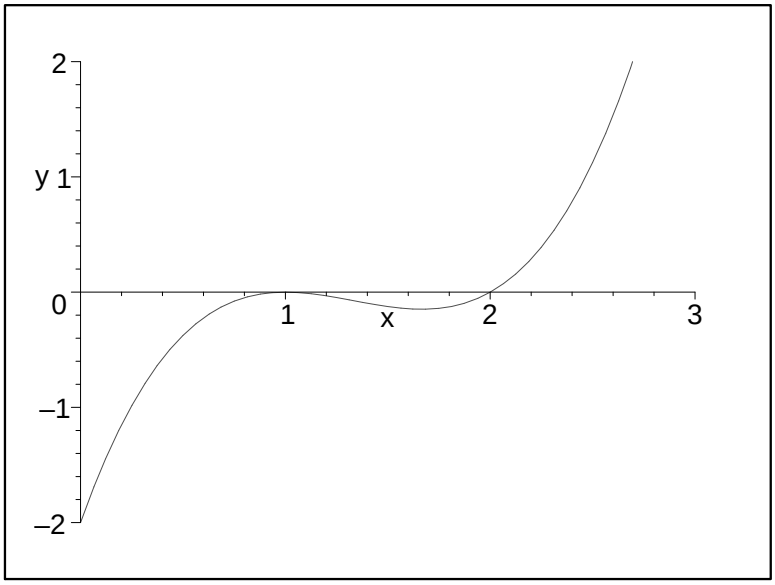


Abbildung II.2: Der Graph der Funktion $x \longmapsto x^3 - 4x^2 + 5x - 2$.

Definition II.1.26. In Teil i) des Satzes setzen wir $f^{-1} := g$. Man nennt f^{-1} die *Umkehrabbildung* oder *inverse Abbildung von* f. Auf Grund von Teil ii) schreiben wir $h \circ g \circ f := h \circ (g \circ f)$ (s. Abbildung II.3).

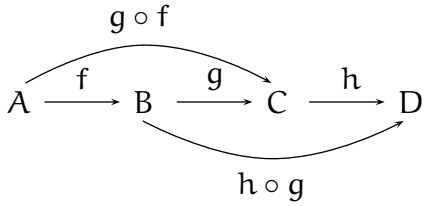


Abbildung II.3: Das Assoziativgesetz für Abbildungen.

Beweis von Satz II.1.25. **Zu i).** Es sei $\Gamma_f \subseteq A \times B$ der Graph von f (s. Bemerkung II.1.17, ii). Man setze

$$\Gamma := \left\{ (y, x) \in B \times A \,\middle|\, (x, y) \in \Gamma_f \right\}.$$

a) Es sei $y \in B$. Da f surjektiv ist, existiert ein $x \in A$ mit $y = f(x)$ oder, anders gesagt, $(x, y) \in \Gamma_f$. Für jedes $y \in B$ existiert also ein $x \in A$, so dass $(y, x) \in \Gamma$.

b) Es seien $y \in B$ und $x_1, x_2 \in A$ gegeben, so dass $(y, x_1), (y, x_2) \in \Gamma$. Nach Definition gilt $(x_1, y), (x_2, y) \in \Gamma_f$. Insbesondere gilt $f(x_1) = y = f(x_2)$. Weil f injektiv ist, folgt $x_1 = x_2$.

Damit definiert Γ gemäß Bemerkung II.1.17, ii), eine Abbildung $g\colon B \longrightarrow A$ mit $\Gamma_g = \Gamma$. Eigenschaft (II.2) ist dabei sofort klar.

Die Eindeutigkeit sieht man folgendermaßen ein: Es seien $g\colon B \longrightarrow A$ und $g'\colon A \longrightarrow B$ zwei Abbildungen mit $f \circ g = \mathrm{Id}_B = f \circ g'$. Wir zeigen die Gleichheit mit Bemerkung II.1.17, i). Es sei $y \in B$. Aus

$$f\big(g(y)\big) = (f \circ g)(y) = \mathrm{Id}_B(y) = y = \mathrm{Id}_B(y) = (f \circ g')(y) = f\big(g'(y)\big)$$

folgt $g(y) = g'(y)$, denn f ist injektiv.

Zu ii). Für ein Element $x \in A$ erhält man:

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

Nach Bemerkung II.1.17, i), gilt daher $(h \circ g) \circ f = h \circ (g \circ f)$. □

Aufgabe II.1.27.
a) Es seien A und B zwei endliche Mengen und $f\colon A \longrightarrow B$ eine Abbildung. Überprüfen Sie:

i) Ist f injektiv, dann gilt $\#A \leq \#B$.

ii) Wenn f surjektiv ist, dann gilt $\#A \geq \#B$.

iii) Gilt $\#A = \#B$ und ist f surjektiv (injektiv), dann ist f auch injektiv (surjektiv).

b) Es seien A, B und C Mengen und $f\colon A \longrightarrow B$ und $g\colon B \longrightarrow C$ Abbildungen. Zeigen Sie:

i) Wenn $g \circ f$ surjektiv ist, dann ist g surjektiv.

ii) Wenn $g \circ f$ injektiv ist, dann ist f injektiv.

Schließen Sie, dass eine Abbildung $f\colon A \longrightarrow B$ genau dann bijektiv ist, wenn es eine Abbildung $f^{-1}\colon B \longrightarrow A$ mit $f^{-1} \circ f = \mathrm{Id}_A$ und $f \circ f^{-1} = \mathrm{Id}_B$ gibt.

Aufgabe II.1.28.
a) Geben Sie eine injektive Abbildung $f\colon \mathbb{N} \longrightarrow \mathbb{N}$ an, für die $\mathbb{N} \setminus f(\mathbb{N})$ unendlich viele Elemente hat.
b) Geben Sie eine surjektive Abbildung $f\colon \mathbb{N} \longrightarrow \mathbb{N}$ an, so dass $\#f^{-1}(\{n\}) = 1$ für alle $n \geq 1$ und $f^{-1}(\{0\})$ unendlich viele Elemente enthält.
c) Geben Sie ein Beispiel für Mengen A, B, C und Abbildungen $f\colon A \longrightarrow B$, $g\colon B \longrightarrow C$ an, in dem $g \circ f$ bijektiv ist, aber f nicht surjektiv und g nicht injektiv.

II.2 Gruppen

Bemerkung II.2.1. Es sei A eine Menge. Wir betrachten[7]

$$\mathcal{S}(A) := \Big\{ f\colon A \longrightarrow A \,\Big|\, f \text{ ist bijektiv} \Big\}.$$

Die Abbildung

$$\begin{aligned} \circ\colon \mathcal{S}(A) \times \mathcal{S}(A) &\longrightarrow \mathcal{S}(A) \\ (f, g) &\longmapsto f \circ g \end{aligned}$$

genügt den folgenden Gesetzmäßigkeiten:

i) Für alle $f, g, h \in \mathcal{S}(A)$ gilt nach Satz II.1.25, ii), $(f \circ g) \circ h = f \circ (g \circ h)$.

ii) Für das Element $\mathrm{Id}_A \in \mathcal{S}(A)$ und alle $f \in \mathcal{S}(A)$ gilt $\mathrm{Id}_A \circ f = f$.

iii) Für alle $f \in \mathcal{S}(A)$ existiert nach Satz II.1.25, i), die Abbildung $f^{-1} \in \mathcal{S}(A)$ mit $f^{-1} \circ f = \mathrm{Id}_A$.

iv) Hat die Menge A mindestens drei Elemente, so gibt es Elemente $f, g \in \mathcal{S}(A)$ mit $f \circ g \neq g \circ f$ (vgl. Beispiel II.1.24, ii).

Man vergleiche diese Eigenschaften mit den gängigen Rechenregeln für $\mathbb{Z}$ und die Abbildung $+\colon \mathbb{Z} \times \mathbb{Z} \longrightarrow \mathbb{Z}$, $(x, y) \longmapsto x + y$:

i) Für alle $x, y, z \in \mathbb{Z}$ gilt $(x + y) + z = x + (y + z)$.

ii) Für die Zahl $0 \in \mathbb{Z}$ und alle $x \in \mathbb{Z}$ gilt $0 + x = x$.

iii) Für alle $x \in \mathbb{Z}$ existiert die Zahl $-x \in \mathbb{Z}$ mit $-x + x = 0$.

iv) Für je zwei ganze Zahlen x, y gilt allerdings $x + y = y + x$.

Die anscheinend so verschiedenen Quadrupel $(\mathcal{S}(A), \circ, \mathrm{Id}_A, {}^{-1})$ und $(\mathbb{Z}, +, 0, -)$ genügen beide den Rechenregeln i) - iii), unterscheiden sich jedoch deutlich in Punkt iv). Es ist eine wichtige Erkenntnis, dass die Eigenschaften i) - iii) eine fundamentale Struktur in der Mathematik definieren, nämlich die der Gruppe.

Definition II.2.2. Eine *Gruppe* ist ein Paar $(G, \star)$, das aus einer Menge G und einer Abbildung $\star\colon G \times G \longrightarrow G$ besteht, so dass folgende Bedingungen erfüllt sind:

i) (*Assoziativgesetz*). Für je drei Elemente $x, y, z \in G$ gilt $(x \star y) \star z = x \star (y \star z)$.

[7] Die Bildung der Menge $\mathcal{S}(A)$ erfolgt mit Hilfe der Komprehension (Definition II.1.5, ii) als Teilmenge von $\mathrm{Abb}(A, A)$ (s. Definition II.1.18).

ii) (*Existenz eines Neutralelements*). Es gibt ein Element $e \in G$, so dass $e \star x = x$ für alle $x \in G$ gilt.

iii) (*Existenz eines inversen Elements*). Zu jedem Element $x \in G$ existiert ein Element $x' \in G$ mit $x' \star x = e$.

Gilt zusätzlich noch

iv) (*Kommutativgesetz).* $x \star y = y \star x$ *für alle* $x, y \in G$,

dann nennt man die Gruppe G *abelsch* oder *kommutativ*.
Schreibweise. I.A. schreibt man nur „die Gruppe G“ anstatt „die Gruppe $(G, \star)$“. Üblicherweise schreibt man das Gruppengesetz „$\star$“ als Multiplikation, d.h. „xy“ oder „$x \cdot y$“ anstatt „$x \star y$“. In abelschen Gruppen ist auch die additive Notation „$x + y$“ gebräuchlich.

Beispiele II.2.3. i) $(\mathbb{Z}, +)$ ist eine abelsche Gruppe.
ii) $(\mathbb{Q}^* := \mathbb{Q} \setminus \{0\}, \cdot)$ und $(\mathbb{R}^* := \mathbb{R} \setminus \{0\}, \cdot)$ sind abelsche Gruppen.
iii) $(\mathbb{R}_{>0}, \cdot)$ ist eine abelsche Gruppe.
iv) Für jede Menge A ist $(\mathcal{S}(A), \circ)$ eine Gruppe, die allerdings nicht abelsch ist, wenn A mindestens drei verschiedene Elemente enthält.
v) Es seien K ein Körper und $n \geq 1$ eine natürliche Zahl. Dann ist die Menge K^n zusammen mit der Addition aus Definition I.2.1, i) und ii), eine abelsche Gruppe. Das wurde in Aufgabe I.2.10 gezeigt.
vi) $(\mathbb{Z}, \cdot)$ und $(\mathbb{N}, +)$ erfüllen die Gesetze i), ii) und iv), jedoch nicht Gesetz iii).

Der Vorteil des abstrakten Gruppenbegriffs ist die Möglichkeit, allgemeine Sätze über Gruppen zu beweisen, die in jeder Gruppe gelten, also insbesondere in $(\mathcal{S}(A), \circ)$ und $(\mathbb{Z}, +)$, anstatt jedesmal, wenn man eine Struktur $(G, \star)$, die die Axiome i) - iii) erfüllt, antrifft, dieselben Eigenschaften aufs Neue zu verifizieren. Mit dem Begriff der Gruppe hat man die wesentlichen Merkmale herauskristallisiert und abstrahiert.

Satz II.2.4. *Es sei* $(G, \star)$ *eine Gruppe. Dann sind auch folgende Eigenschaften erfüllt:*
i) *Für jedes* $x \in G$ *gilt* $x \star e = x$.
ii) *Es sei* $e' \in G$, *so dass* $e' \star x = x$ *für alle* $x \in G$ *gilt. Dann folgt* $e' = e$.
iii) *Für jedes* $x \in G$ *und jedes* $x' \in G$ *mit* $x' \star x = e$ *gilt auch* $x \star x' = e$.
iv) *Es seien* x, x' *und* $x'' \in G$, *so dass* $x' \star x = e = x'' \star x$. *Dann folgt* $x' = x''$.

Definitionen II.2.5. i) Das Element e ist ***das*** *Neutralelement von* G.
ii) Es sei $x \in G$. Das Element $x' \in G$ mit $x' \star x = e$ ist ***das*** *zu* x *inverse Element*.
Schreibweise. $x^{-1} := x'$.

Beweis von Satz II.2.4. Wir beginnen mit iii). Wir wissen, dass es ein Element $x'' \in G$ mit $x'' \star x' = e$ gibt. Damit ergibt sich:

$$\begin{aligned} x \star x' &= e \star (x \star x') = (x'' \star x') \star (x \star x') = x'' \star \Big(x' \star (x \star x')\Big) = \\ &= x'' \star \Big((x' \star x) \star x'\Big) = x'' \star \Big(e \star x'\Big) = x'' \star x' = e. \end{aligned}$$

Teil i) sieht man jetzt wie folgt:

$$x \star e = x \star (x' \star x) = (x \star x') \star x = e \star x = x.$$

Aus i) folgt

$$e' = e' \star e = e.$$

Schließlich erhält man iv) auf folgende Weise:

$$x' = x' \star e = x' \star (x \star x'') = (x' \star x) \star x'' = e \star x'' = x''.$$

Dabei wurde für die zweite Gleichung die bereits gewonnene Erkenntnis iii) sinngemäß auf x'' angewandt. □

Bemerkungen und Beispiele II.2.6. i) Man kann noch auf weitere Eigenschaften schließen. Z.B. folgt aus der Tatsache iii) $x \star x^{-1} = e$ zusammen mit iv) $(x^{-1})^{-1} = x$.

ii) Ohne irgendwelche Extraüberlegungen anstellen zu müssen, ergibt sich aus rein formalen Gründen, dass es zu einer bijektiven Abbildung $f\colon A \longrightarrow A$ **genau eine** inverse Abbildung $f^{-1}\colon A \longrightarrow A$ gibt. Auch sieht man, dass die Identität durch die Eigenschaft $\mathrm{Id}_A \circ f = f = f \circ \mathrm{Id}_A$ für alle $f \in \mathcal{S}(A)$ ausgezeichnet ist.

iii) Für eine endliche Gruppe $G = \{x_1, ..., x_m\}$ kann man das Gruppengesetz durch eine sogenannte Verknüpfungstabelle angeben, in der man alle möglichen Produkte $x_i \star x_j$, $i = 1, ..., m$, $j = 1, ..., m$, einträgt:

$\star$	x_1	x_2	$\cdots$	x_m
x_1	$x_1 \star x_1$	$x_1 \star x_2$	$\cdots$	$x_1 \star x_m$
x_2	$x_2 \star x_1$	$x_2 \star x_2$	$\cdots$	$x_2 \star x_m$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_m	$x_m \star x_1$	$x_m \star x_2$	$\cdots$	$x_m \star x_m$

.

Benutzen wir die Bezeichnungen aus Aufgabe II.2.10, so stehen in der i-ten Zeile die Elemente $L_{x_i}(x)$, $x \in G$, und in der j-ten Spalte die Elemente $R_{x_j}(x)$, $x \in G$. Die Bijektivität der Abbildungen L_{x_i} und R_{x_j} sagt daher aus, dass in der i-ten Zeile bzw. j-ten Spalte jedes Element von G genau einmal vorkommt. Damit kann man bereits alle Gruppen mit wenigen Elementen bestimmen.

Für $G = \{ e, a, b \}$ folgt aus der Eigenschaft des Neutralements, dass die Verknüpfungstabelle wie folgt ausssieht:

$\star$	e	a	b
e	e	a	b
a	a	$?_1$	$?_2$
b	b	$?$	$?$

.

Das Element $?_1$ muss b sein, da andernfalls $?_2 = b$ gälte und somit b in der dritten Spalte zweimal vorkommen müsste. Damit füllt man die Verknüpfungstabelle auf zu

$\star$	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

.

Durch diese Tabelle wird in der Tat eine Gruppe definiert (Übung).

Für $G = \{ e, a, b, c \}$ erhält man zunächst

$\star$	e	a	b	c
e	e	a	b	c
a	a	$?_1$	$?$	$?$
b	b	$?$	$?$	$?$
c	c	$?$	$?$	$?$

.

Für $?_1$ können wir e, b oder c wählen. Nehmen wir einmal $?_1 = e$. Es ergibt sich

$\star$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	$?_2$	$?$
c	c	b	$?$	$?$

.

Für $?_2 = e$ bzw. $?_2 = a$ erhalten wir

$\star$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

bzw.

$\star$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	a	e
c	c	b	e	a

.

Man kann überprüfen, dass durch beide Verknüpfungstafeln Gruppen definiert werden. (Es muss nur das Assoziativgesetz überprüft werden, da die Existenz des Neutralelements und der inversen Elemente sofort abgelesen werden kann.)

In der ersten Gruppe gilt $x \star x = e$ für jedes Element $x \in G$. Diese Gruppe nennt man die *Kleinsche*[8] *Vierergruppe*. In der zweiten Gruppe gilt $b^1 = b$, $b^2 = a$, $b^3 = c$ und $b^4 = e$. Jedes Element ist also eine Potenz von b. Diese Gruppe heißt die *zyklische Gruppe der Ordnung* 4. Die beiden Gruppen sind offenbar grundlegend verschieden. Für die restlichen Wahlmöglichkeiten für $?_1$ und $?_2$ erhält man nach geeigneter Umbenennung von a, b und c wieder die zweite Tabelle. Es gibt also in einem gewissen Sinne genau zwei Gruppen mit vier Elementen. Die Behandlung von Gruppen mit mindestens fünf Elementen ist bereits komplexer, da man hier durch das Assoziativgesetz zusätzliche Einschränkungen erhält, die sehr unübersichtlich werden.

Untergruppen. — In der Theorie der linearen Gleichungssysteme haben uns die Eigenschaften der Lösungsmengen homogener Gleichungssysteme zum Begriff des linearen Teilraums geführt (s. Definition I.2.5). Für Gruppen ist das entsprechende Konzept in der folgenden Definition enthalten.

Definition II.2.7. Es sei G eine Gruppe. Eine Teilmenge $H \subseteq G$ wird *Untergruppe von* G genannt, wenn sie folgende Eigenschaften hat:

a) $e \in H$;

b) Für alle $x \in H$ gilt auch $x^{-1} \in H$;

c) Für alle $x, y \in H$ hat man ebenfalls $xy \in H$.

Bemerkungen und Beispiele II.2.8. i) Auf Grund von Eigenschaft c) hat man die Abbildung

$$\begin{array}{rcl} \cdot : H \times H & \longrightarrow & H \\ (x, y) & \longmapsto & xy. \end{array}$$

Aus a) und b) folgt sofort, dass H mit dieser Multiplikation selber zur Gruppe wird.

ii) Man kann a) - c) umformulieren, und zwar ist $H \subseteq G$ genau dann eine Untergruppe, wenn A) $H \neq \varnothing$ und B) $xy^{-1} \in H$ aus $x \in H$ und $y \in H$ folgt. Dies sei dem Leser als Übung überlassen.

iii) Für jede Gruppe G sind $\{e\}$ und G Untergruppen.

iv) Sind H und H' Untergruppen von G, dann ist auch $H \cap H'$ eine Untergruppe von G (Übung).

v) Es seien K ein Körper, $m, n \geq 1$ natürliche Zahlen und A eine $(m \times n)$-Matrix mit Einträgen in K. Dann ist $\mathbb{L}(A, 0)$ eine Untergruppe von $(K^n, +)$ (s. Lemma I.2.4).

[8] Felix Christian Klein (1849 - 1925), deutscher Mathematiker.

vi) Für $m \in \mathbb{N}$ setzen wir

$$m\mathbb{Z} := m \cdot \mathbb{Z} := \left\{ a \in \mathbb{Z} \,\middle|\, \exists k \in \mathbb{Z} : a = km \right\} = \left\{ 0, \pm m, \pm 2m, \pm 3m, \ldots \right\}.$$

Man überprüft, dass $m\mathbb{Z}$ eine Untergruppe von $(\mathbb{Z}, +)$ ist. In der Tat sind alle Untergruppen von $\mathbb{Z}$ von dieser Form, wie der folgende Satz zeigen wird.

Satz II.2.9. *Es sei $H \subseteq \mathbb{Z}$ eine Untergruppe. Dann gibt es eine natürliche Zahl m, so dass $H = m\mathbb{Z}$.*

Beweis. Wenn $H = \{0\}$, dann ist die Aussage mit $m = 0$ richtig. Ansonsten sei $H \setminus \{0\} \neq \varnothing$. Da mit jedem Element $x \in H$ auch $-x$ in H liegt, folgt $H_{>0} := \{h \in H \,|\, h > 0\} \neq \varnothing$. Es sei m das kleinste Element von $H_{>0}$ (s. [8], Satz 1.3.22). Wir behaupten $H = m\mathbb{Z}$. Da H eine Untergruppe von $\mathbb{Z}$ ist, gilt auf jeden Fall schon einmal $m\mathbb{Z} \subseteq H$. Zu einem beliebigen Element $h \in H$ können wir ein Element $k \in \mathbb{Z}$ und ein Element $0 \leq r < m$ finden, so dass $h = k \cdot m + r$ (Division mit Rest; [9], Satz I.1.1). Da $h \in H$ und $k \cdot m \in H$, gehört auch das Element $r = h - k \cdot m$ der Untergruppe H an. Nun ist $r \geq 0$. Auf der anderen Seite ist m das kleinste Element von $H_{>0}$, so dass wegen $r < m$ nur $r = 0$ gelten kann. Damit ist $h = k \cdot m \in m\mathbb{Z}$ nachgewiesen. □

Aufgabe II.2.10.
Es seien $(G, \star)$ eine Gruppe und $y \in G$. Wir betrachten die *Linkstranslation um* y

$$\begin{array}{rccc} L_y \colon & G & \longrightarrow & G \\ & x & \longmapsto & y \star x \end{array}$$

sowie die *Rechtstranslation um* y

$$\begin{array}{rccc} R_y \colon & G & \longrightarrow & G \\ & x & \longmapsto & x \star y. \end{array}$$

Zeigen Sie, dass L_y und R_y bijektiv sind, und geben Sie die inversen Abbildungen an.

II.3 Ringe und Körper

Definitionen II.3.1. i) Ein *Ring* ist ein Tripel $(R, \oplus, \odot)$, das aus einer Menge R sowie Abbildungen

$$\begin{array}{rccc} \oplus \colon & R \times R & \longrightarrow & R \\ & (x, y) & \longmapsto & x \oplus y \end{array} \qquad \text{(\textit{Addition})}$$

und

$$\begin{array}{rccc} \odot \colon & R \times R & \longrightarrow & R \\ & (x, y) & \longmapsto & x \odot y \end{array} \qquad \text{(\textit{Multiplikation})}$$

besteht, so dass es zwei Elemente $\mathbb{0} \neq \mathbb{1}$ gibt und gilt:

a) $(R, \oplus)$ ist eine abelsche Gruppe mit Neutralelement $\mathbb{0}$.

b) Für alle $x, y, z \in R$ gilt $x \odot (y \odot z) = (x \odot y) \odot z$.

c) Für jedes Element $x \in R$ gilt $\mathbb{1} \odot x = x = x \odot \mathbb{1}$.

d) Es gelten die Distributivgesetze

$$\begin{aligned} x \odot (y \oplus z) &= x \odot y \oplus x \odot z, \\ (x \oplus y) \odot z &= x \odot z \oplus y \odot z \end{aligned}$$

für alle $x, y, z \in R$.

Ist außerdem

e) $x \odot y = y \odot x$ für alle $x, y \in R$

erfüllt, so nennt man R einen *kommutativen Ring*.
Schreibweise. Üblicherweise werden wir Addition bzw. Multiplikation in einem Ring einfach als „+“ bzw. „·“ schreiben und „0“ bzw. „1“ anstatt „$\mathbb{0}$“ und „$\mathbb{1}$“.

ii) Ein *Körper* ist ein Ring $(K, \oplus, \odot)$, in dem zusätzlich gilt:

a) (*Nullteilerfreiheit*). Für alle $x, y \in K^* := K \setminus \{\mathbb{0}\}$ gilt $x \odot y \neq \mathbb{0}$.

b) $(K^*, \odot)$ mit der (nach a) existierenden) Verknüpfung

$$\begin{aligned} \odot\colon K^* \times K^* &\longrightarrow K^* \\ (x, y) &\longmapsto x \odot y \end{aligned}$$

ist eine **abelsche** Gruppe.

Bemerkungen II.3.2. i) In c) müssen wir $\mathbb{1} \odot x = x = x \odot \mathbb{1}$ fordern, da wir nicht den Trick aus der Theorie der Gruppen anwenden können (s. Satz II.2.4, i), der ja die Existenz von Inversen voraussetzt, die aber für die Multiplikation „$\odot$“ im Allgemeinen nicht gewährleistet ist.

ii) Lässt man die Möglichkeit $\mathbb{0} = \mathbb{1}$ zu, dann erhält man den *Nullring* $\{\mathbb{0}\}$, indem alle Gesetze wie oben gelten.

iii) In einem Ring gilt $0 \cdot x = 0$ für jedes Element $x \in R$. In der Tat hat man $0 \cdot x = (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x$. Addiert man $-(0 \cdot x)$ auf beiden Seiten, folgt die Behauptung.

Beispiele II.3.3. i) $(\mathbb{Z}, +, \cdot)$ ist ein kommutativer Ring aber kein Körper. Zwar erfüllt $\mathbb{Z}$ die zusätzliche Eigenschaft a) aber nicht b), denn z.B. 2 besitzt kein multiplikatives Inverses.

ii) $(\mathbb{Q}, +, \cdot)$ und $(\mathbb{R}, +, \cdot)$ sind Körper.

iii) Der kleinst mögliche Ring hat zwei Elemente; $\mathbb{F}_2 := \{0, 1\}$. Die einzig möglichen Verknüpfungstabellen für Addition und Multiplikation sind:

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

Man kann überprüfen, dass $(\mathbb{F}_2, +, \cdot)$ ein Körper ist. In ihm gilt die (auf den ersten Blick befremdliche) Regel $1 + 1 = 0$.

Bemerkung II.3.4. Man erkennt an Punkt iii) des obigen Beispiels, wie nützlich eine abstrakte Begriffsbildung ist. Obwohl sich die Objekte $(\mathbb{Q}, +, \cdot)$ und $(\mathbb{F}_2, +, \cdot)$ stark unterscheiden ($\mathbb{Q}$ hat z.B. unendlich viele Elemente, $\mathbb{F}_2$ nur zwei), gehorchen sie denselben fundamentalen Gesetzmäßigkeiten. Indem wir also die Theorie nur unter Verwendung der Körperaxiome entwickeln, erhalten wir Aussagen, die für alle Körper, also insbesondere $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{F}_2$, Gültigkeit haben. Dies trifft z.B. auf den Gauß-Algorithmus zu. Wollen wir ein Gleichungssystem über $\mathbb{F}_2$ lösen, so liefert er auch hier das Verfahren. In Abschnitt II.4 werden wir noch weitere endliche Körper kennenlernen. Endliche Körper mögen zunächst als Spielerei erscheinen, doch spielen sie in vielen mathematischen Gebieten, z.B. der Zahlentheorie und der algebraischen Geometrie, und auch in Anwendungen, z.B. der Kodierungstheorie (s. [7], Kapitel 12), eine wichtige Rolle.

Aufgabe II.3.5.
a) Zeigen Sie, dass $(\mathbb{R}^2, +, \cdot)$ mit

$$\begin{array}{rcl} +\colon \mathbb{R}^2 \times \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ \left(\begin{pmatrix} s \\ t \end{pmatrix}, \begin{pmatrix} u \\ v \end{pmatrix}\right) & \longmapsto & \begin{pmatrix} s+u \\ t+v \end{pmatrix} \end{array}$$

und

$$\begin{array}{rcl} \cdot\colon \mathbb{R}^2 \times \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ \left(\begin{pmatrix} s \\ t \end{pmatrix}, \begin{pmatrix} u \\ v \end{pmatrix}\right) & \longmapsto & \begin{pmatrix} su-tv \\ sv+tu \end{pmatrix} \end{array}$$

ein Körper ist. Es sei $\mathbb{1}$ das Neutralelement der Multiplikation. Zeigen Sie, dass es ein Element i gibt mit $i^2 = -\mathbb{1}$. Man nennt den eben erhaltenen Körper den *Körper der komplexen Zahlen* und bezeichnet ihn mit $\mathbb{C}$. Man schreibt

$$\begin{pmatrix} s \\ t \end{pmatrix} = s + i \cdot t.$$

b) Es sei $F = \{0, 1, x, y\}$ eine Menge mit genau vier Elementen. Wir nehmen an, es gäbe Verknüpfungen $+, \cdot : F \times F \longrightarrow F$, so dass $(F, +, \cdot)$ ein Körper ist. Stellen Sie unter Verwendung der Körperaxiome zunächst die Verknüpfungstafel für die Multiplikation und dann diejenige für die Addition auf.

Aufgabe II.3.6.
a) Die Abbildung

$$\begin{array}{rcl} \overline{}: \mathbb{C} & \longrightarrow & \mathbb{C} \\ z = x + i \cdot y & \longmapsto & \overline{z} := x - i \cdot y \end{array}$$

wird *Konjugation* genannt. Zeigen Sie:

i) Es gilt genau dann $z \in \mathbb{R}$, wenn $\overline{z} = z$.

ii) Für alle $z, z' \in \mathbb{C}$ hat man $\overline{z + z'} = \overline{z} + \overline{z'}$.

iii) Für alle $z, z' \in \mathbb{C}$ hat man $\overline{z \cdot z'} = \overline{z} \cdot \overline{z'}$.

iv) Für jedes Element $z \in \mathbb{C}^*$ gilt

$$z^{-1} = \frac{\overline{z}}{z \cdot \overline{z}}.$$

b) Schreiben Sie folgende komplexe Zahlen in der Form $s + i \cdot t$:

$$\text{i)}(5 - 7i) \cdot (1,5 + 12i); \quad \text{ii)}(2 - 3i)^4; \quad \text{iii)}(\sqrt{3} + 0,5i)^{-1}; \quad \text{iv)}\frac{1 + 5i}{2 - 7i}.$$

c) Zeigen Sie durch explizite Rechnungen, dass für jede komplexe Zahl z eine komplexe Zahl w mit $w^2 = z$ existiert.

Aufgabe II.3.7.
a) Es sei K ein Körper. Gegeben seien $f, g \in \mathrm{Abb}(\mathbb{N}, K)$ (s. Definition II.1.18). Wir schreiben $f(i) = a_i$, $g(i) = b_i$, $i \in \mathbb{N}$, und definieren

$$\begin{array}{rcl} f \oplus g: \mathbb{N} & \longrightarrow & K \\ k & \longmapsto & f(k) + g(k) = a_k + b_k. \end{array}$$

sowie

$$\begin{array}{rcl} f \odot g: \mathbb{N} & \longrightarrow & K \\ k & \longmapsto & \displaystyle\sum_{i,j \in \mathbb{N}: i+j=k} a_i \cdot b_j. \end{array}$$

Zeigen Sie, dass $(\mathrm{Abb}(\mathbb{N}, K), \oplus, \odot)$ ein kommutativer Ring ist.[9]
b) Wir setzen

$$\mathrm{Abb}'(\mathbb{N}, K) := \Big\{ f \in \mathrm{Abb}(\mathbb{N}, K) \,\Big|\, f(i) = 0 \text{ für alle bis auf endlich viele } i \in \mathbb{N} \Big\}.$$

Zeigen Sie, dass $\mathrm{Abb}'(\mathbb{N}, K)$ ein **Unterring**[10] von $K[\![t]\!]$ ist. Kommt Ihnen der Ring $(\mathrm{Abb}'(\mathbb{N}, K), +, \odot)$ bekannt vor?

II.4 Äquivalenzrelationen

Definitionen II.4.1. i) Es sei A eine Menge. Eine Teilmenge $R \subseteq A \times A$ wird eine *Relation* genannt.
Schreibweise. $x \sim_R y :\Longleftrightarrow (x, y) \in R$, $x, y \in A$.

ii) Eine Relation $R \subseteq A \times A$ ist eine *Äquivalenzrelation*, wenn sie die folgenden drei Bedingungen erfüllt.

a) (*Reflexivität*). Für jedes $x \in A$ gilt $(x, x) \in R$ ($x \sim_R x$).

b) (*Symmetrie*). Für alle $x, y \in A$ gilt: $(x, y) \in R \Longrightarrow (y, x) \in R$ ($x \sim_R y \Longrightarrow y \sim_R x$).

c) (*Transitivität*). Für alle $x, y, z \in A$ hat man: $(x, y) \in R \wedge (y, z) \in R \Longrightarrow (x, z) \in R$ ($x \sim_R y \wedge y \sim_R z \Longrightarrow x \sim_R z$).

iii) Es seien $R \subseteq A \times A$ eine Äquivalenzrelation und $x \in A$. Die *Äquivalenzklasse von* x ist die Menge

$$[x] := [x]_R := \Big\{ y \in A \,\Big|\, y \sim_R x \Big\}.$$

Die *Menge der Äquivalenzklassen* ist gegeben durch

$$A/R := A/\sim_R := \Big\{ [x]_R \,\Big|\, x \in A \Big\} \subseteq \mathcal{P}(A).$$

Eine Teilmenge $S \subset A$ ist ein *vollständiges Repräsentantensystem* für R, wenn gilt:

$$\forall M \in A/R \; \exists! x \in S : \quad x \in M.$$

[9] Man nennt diesen Ring den *Ring der formalen Potenzreihen (über K)* und bezeichnet ihn mit $K[\![t]\!]$, t eine Unbestimmte. Ein Element $f\colon \mathbb{N} \longrightarrow K$, $i \longmapsto a_i$, wird dann in der Form

$$f = \sum_{i=0}^{\infty} a_i \cdot t^i$$

geschrieben.
[10] Überlegen Sie sich, was damit gemeint ist.

Beispiele II.4.2. i) Es sei $R := \Delta_A = \{(x,x) \in (A \times A) \mid x \in A\}$. Dies ist eine Äquivalenzrelation, und es gilt $x \sim_R y \iff x = y$. Für jedes $x \in A$ hat man demnach $[x]_R = \{x\}$.[11]

ii) Die Menge $R = A \times A$ ist auch eine Äquivalenzrelation auf A. Hier gilt für alle $x, y \in A$ die Relation $x \sim_R y$ und folglich $[x]_R = A$ für alle $x \in A$.[12]

Eigenschaften II.4.3. i) *Für jedes Element* $M \in A/R$ *gilt* $M \neq \varnothing$.

ii) *Für jedes Element* $x \in A$ *gibt es ein* $M \in A/R$ *mit* $x \in M$.

iii) *Für Elemente* $M, M' \in A/R$ *gilt entweder* a) $M = M'$ *oder* b) $M \cap M' = \varnothing$.

Beweis. **Zu i).** Nach Definition gibt es ein $x \in A$, so dass $M = [x]_R$. Auf Grund der Reflexivität von R gilt $x \sim_R x$, i.e. $x \in [x]_R = M$ und somit $M \neq \varnothing$.

Zu ii). Wie wir schon gesehen haben, gilt $x \in [x]_R$.

Zu iii). Es seien $M = [x]_R$ und $M' = [x']_R$ für geeignete $x, x' \in A$. Wir nehmen an, dass $M \cap M' \neq \varnothing$. Es sei $z \in M \cap M'$. Wir werden $[x]_R \subseteq [x']_R$ zeigen. Da wir die Rollen von x und x' offenbar vertauschen können, folgt ebenso $[x']_R \subseteq [x]_R$ und daher $[x]_R = [x']_R$ (Bemerkung II.1.3). Es sei also $y \in [x]_R$, d.h. $y \sim_R x$. Ferner gilt $z \in [x]_R$, i.e. $z \sim_R x$, und wegen der Symmetrieeigenschaft hat man auch $x \sim_R z$. Die Transitivität von R liefert $y \sim_R z$. Weiterhin haben wir $z \in [x']_R$, also $z \sim_R x'$. Wiederum ergibt sich aus der Transitivität $y \sim_R x'$ und somit, wie gewünscht, $y \in [x']_R$. □

Schreibweise. Es seien A eine Menge, und A_i, $i \in I$, eine Familie von Teilmengen von A (s. Beispiel II.1.19, iv). Man schreibt $A = \bigsqcup_{i \in I} A_i$, wenn a) $A = \bigcup_{i \in I} A_i$ und b) $A_i \cap A_j = \varnothing$ für alle $i \neq j \in I$ gilt, und spricht von einer *disjunkten Vereinigung*.

Mit dieser Schreibweise ergibt sich aus II.4.3

$$A = \bigsqcup_{M \in A/R} M.$$

Beispiel II.4.4. Auf der Menge

$$M := \{0,1,2,3\} \times \{0,1,2,3\}$$

betrachten wir die Relation[13]

$$\forall (a,b),(a',b') \in M: \quad \big((a,b) \sim (a',b') \quad :\iff \quad a + b' = a' + b\big).$$

Es ist leicht nachzuprüfen, dass „$\sim$“ eine Äquivalenzrelation auf M ist (vgl. Aufgabe II.4.11). In Abbildung II.4 haben wir die zugehörigen Äquivalenzklassen farbig markiert. Ein vollständiges Repräsentatensystem wäre z.B. die Menge

[11] Nach Axiom a) ist dies die feinste Äquivalenzrelation.

[12] Dies ist die gröbste Äquivalenzrelation.

[13] Wir verwenden die Addition auf den natürlichen Zahlen.

$$M = \left\{ \begin{array}{cccc} (0,0), & (0,1), & (0,2), & (0,3), \\ (1,0), & (1,1), & (1,2), & (1,3), \\ (2,0), & (2,1), & (2,2), & (2,3), \\ (3,0), & (3,1), & (3,2), & (3,3) \end{array} \right\}$$

Abbildung II.4: Die Äquivalenzklassen der Relation „$\sim$“

$$\Big\{ (3,0), (2,0), (1,0), (0,0), (0,1), (0,2), (0,3) \Big\}$$

oder auch

$$\Big\{ (3,0), (3,1), (3,2), (3,3), (2,3), (1,3), (0,3) \Big\}.$$

Faktorgruppen. — Es seien G eine Gruppe und $H \subseteq G$ eine Untergruppe. Wir definieren

$$R_H := \Big\{ (x, xh) \in G \times G \,\Big|\, x \in G, h \in H \Big\}.$$

Schreibweise. „$x \sim_H y$“ anstatt „$x \sim_{R_H} y$“.

Lemma II.4.5. *Die Relation R_H ist eine Äquivalenzrelation.*

Beweis. Nach Definition gilt $x \sim_H y$ für $x, y \in G$ genau dann, wenn es ein Element $h \in H$ mit $y = xh$ gibt. Wir überprüfen die Axiome aus Definition II.4.1. **Zu a).** Für jedes $x \in G$ gilt $x = x \cdot e$. Weil e ein Element von H ist, folgt $x \sim_R x$. **Zu b).** Für $x, y \in G$ mit $x \sim_H y$ findet man ein Element $h \in H$, so dass $y = xh$. Damit erhält man $x = yh^{-1}$. Nun ist h^{-1} auch in H enthalten, so das $y \sim_H x$ gilt. **Zu c).** Für $x, y, z \in G$ mit $x \sim_H y$ und $y \sim_H z$ seien $h, j \in H$ so, dass $y = xh$ und $z = yj$. Zusammen ergibt sich $z = (xh)j = x(hj)$. Das Element hj liegt in H, und wir schließen $x \sim_H z$. □

Schreibweise. $xH := [x]_H := [x]_{R_H} = \{ xh \,|\, h \in H \}$; $G/H := G/R_H = \{ xH \,|\, x \in G \}$.

Satz II.4.6. *Es seien G eine* ***abelsche*** *Gruppe und $H \subseteq G$ eine Untergruppe. Dann wird auf G/H durch*

$$\begin{array}{rcl} \cdot : G/H \times G/H & \longrightarrow & G/H \\ \big([x], [y]\big) & \longmapsto & [xy] \end{array}$$

die Struktur einer Gruppe definiert.

Beweis. Wir müssen zeigen, dass die Abbildung „$\cdot$“ wohldefiniert ist, d.h. dass die angegebene Abbildungsvorschrift überhaupt Sinn macht. Für $M, M' \in G/H$ wird das Produkt $M \cdot M'$ durch die Auswahl sogenannter Repräsentanten $x \in M$ und $y \in M'$ definiert. Es ist deshalb nachzuweisen, dass das Ergebnis $M \cdot M'$

nicht von der Auswahl dieser Repräsentanten abhängt. In unserem Fall ist also zu überprüfen, dass für $x, x', y, y' \in G$ mit $[x] = [x']$ und $[y] = [y']$ auch $[xy] = [x'y']$ gilt. Unter der angegebenen Annahme finden wir Elemente $h, j \in H$ mit $x' = xh$ und $y' = yj$. Wir schließen

$$x'y' = (xh)(yj) = x(hy)j \overset{\textbf{G abelsch}}{=} x(yh)j = (xy)(hj).$$

Wegen $hj \in H$ folgt $[x'y'] = [xy]$. Das Neutralelement in G/H ist $[e]$, das inverse Element von $[x] \in G/H$ ist $[x^{-1}]$. Das Assoziativgesetz ist offenkundig erfüllt. □

Bemerkungen und Beispiele II.4.7. i) Die Bedingung, dass G abelsch ist, ging ganz wesentlich in unseren Beweis ein. Falls G nicht abelsch ist, dann muss H ein sogenannter **Normalteiler** sein. Dies bedeutet, dass für jedes $y \in G$ und jedes $h \in H$ ein Element $h' \in H$ mit $hy = yh'$ existiert. Unter dieser Voraussetzung funktioniert obiger Beweis fast genauso (s. [9], Satz II.9.4).

ii) Für $m > 1$, $G = \mathbb{Z}$ und $H = m\mathbb{Z}$, erhalten wir die Gruppe $\mathbb{Z}/m\mathbb{Z}$ mit der Addition $[x] + [y] = [x + y]$. Wir können die Multiplikation durch

$$\begin{aligned} \cdot\colon \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z} &\longrightarrow \mathbb{Z}/m\mathbb{Z} \\ ([x], [y]) &\longmapsto [xy] \end{aligned}$$

erklären. Wie im obigen Beweis verifiziert man, dass „$\cdot$“ wohldefiniert ist und dass $(\mathbb{Z}/m\mathbb{Z}, +, \cdot)$ sogar ein Ring ist (vgl [9], Beispiel III.1.3, vii).

Wir behaupten, dass die Menge $\{0, ..., m-1\}$ ein vollständiges Repräsentatensystem ist. Für $x, y \in \{0, ..., m-1\}$ gilt $-m < x-y < m$, so dass m die Differenz $x - y$ nicht teilen kann und x und y verschiedene Äquivalenzklassen repräsentieren. Nun sei $k \in \mathbb{Z}$. Auf Grund der Division mit Rest ([9], Satz I.1.1) gibt es eine ganze Zahl $l \in \mathbb{Z}$ und eine Zahl $r \in \{0, ..., m-1\}$ mit $k = l \cdot m + r$. Offenbar gilt $[k] = [r]$, d.h. die Äquivalenzklasse von k wird durch das Element $r \in \{0, ..., m-1\}$ repräsentiert.

Falls m keine Primzahl ist, finden wir natürliche Zahlen $1 < s < m$ und $1 < t < m$ mit $m = st$. Dabei gilt $[s] \neq [0]$ und $[t] \neq 0$ in $\mathbb{Z}/m\mathbb{Z}$, aber $[s][t] = [st] = [m] = [0]$. In diesem Fall ist also Axiom a) für einen Körper aus Definition II.3.1, ii), verletzt.

Satz II.4.8. *Ist $p \in \mathbb{N}$ eine Primzahl, dann ist $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$ mit den oben eingeführten Verknüpfungen ein Körper.*

Beweis. **Schritt 1.** Es seien $x, y \in \mathbb{F}_p \setminus \{0\}$. Wir können Zahlen $0 < a < p$ und $0 < b < p$ mit $x = [a]$ und $y = [b]$ finden. Falls $xy = [a][b] = [ab] = 0$ gälte, folgte, dass p die Zahl ab teilt. Weil p aber eine Primzahl ist, kann dies nur passieren, wenn p schon a oder b teilt ([9], Satz I.4.5). Wegen $0 < a < p$ bzw. $0 < b < p$ ist dies nicht möglich, so dass tatsächlich $xy = [ab] \neq 0$ gilt.

Schritt 2. Wir müssen die Existenz multiplikativer Inverser beweisen. Wir beginnen mit der

Behauptung. *Es sei* $0 < a < p$*. Dann gibt es ganze Zahlen* k *und* l *mit* $ak+pl = 1$*.*

Diese Behauptung sieht man wie folgt ein. Die Menge

$$H := a\mathbb{Z} + p\mathbb{Z} = \left\{ x \in \mathbb{Z} \,\middle|\, \exists k, l \in \mathbb{Z} : x = ak + pl \right\}$$

ist eine Untergruppe von $\mathbb{Z}$, die a und p enthält. Durch Satz II.2.9 wissen wir, dass es ein $m > 0$ gibt, so dass $H = m\mathbb{Z}$. Insbesondere gibt es ganze Zahlen s, t mit A) $p = ms$ und B) $a = mt$. Aus A) folgt $m = 1$ oder $m = p$ und wegen B) und $0 < a < p$ muss $m = 1$ zutreffen, also $H = \mathbb{Z}$. Daher gibt es Zahlen $k, l \in \mathbb{Z}$ mit $1 = ak + pl$. ✓

Ein Element $x \in \mathbb{F}_p \backslash \{0\}$ schreiben wir wieder in der Form $x = [a]$ mit $0 < a < p$. Gemäß der Behauptung wählen wir $k, l \in \mathbb{Z}$ mit $ak + pl = 1$. Wir sehen

$$[a][k] = [ak] \overset{[pl]=[0]}{=} [ak] + [pl] = [ak + pl] = [1],$$

so dass $y := [k]$ das zu $x = [a]$ multiplikative Inverse ist. □

Aufgabe II.4.9.
Wir kehren nun die Beobachtungen aus Eigenschaft II.4.3 um. Es seien M eine Menge, I eine Indezmenge und $(M_i)_{i\in I}$ eine Familie von nichtleeren Teilmengen von M, so dass

$$M = \bigsqcup_{i\in I} M_i.$$

Auf M führen wir folgende Relation ein:

$$\forall m, m' \in M : \quad \big(m \sim m' \quad :\Longleftrightarrow \quad \exists i \in I : m \in M_i \wedge m' \in M_i\big).$$

a) Zeigen Sie, dass „$\sim$“ eine Äquivalenzrelation auf M ist.
b) Weisen Sie nach, dass die Äquivalenzklassen von „$\sim$“ gerade die Teilmengen M_i, $i \in I$, sind.

Aufgabe II.4.10.
Es seien A, B Mengen und $f\colon A \longrightarrow B$ eine Abbildung. Auf A wird die Relation „$\sim_f$“ vermöge

$$\forall a, a' \in A : \quad a \sim_f a' \quad :\Longleftrightarrow \quad f(a) = f(a')$$

erklärt. Weisen Sie nach, dass „$\sim_f$“ eine Äquivalenzrelation ist.

Aufgabe II.4.11.
Für zwei Elemente $(m_1, n_1) \in \mathbb{N} \times \mathbb{N}$ und $(m_2, n_2) \in \mathbb{N} \times \mathbb{N}$ schreiben wir

$$(m_1, n_1) \sim (m_2, n_2) :\Longleftrightarrow m_1 + n_2 = m_2 + n_1.$$

a) Zeigen Sie, dass „$\sim$“ eine Äquivalenzrelation auf $\mathbb{N} \times \mathbb{N}$ definiert.
b) Es sei $Z := (\mathbb{N} \times \mathbb{N})/\sim$. Zeigen Sie, dass die Abbildungen

$$\begin{array}{rcl} \oplus\colon Z \times Z & \longrightarrow & Z, \\ ([m_1, n_1], [m_2, n_2]) & \longmapsto & [m_1 + m_2, n_1 + n_2] \end{array}$$

und

$$\begin{array}{rcl} \odot\colon Z \times Z & \longrightarrow & Z \\ ([m_1, n_1], [m_2, n_2]) & \longmapsto & [m_1 m_2 + n_1 n_2, m_1 n_2 + m_2 n_1] \end{array}$$

wohldefiniert sind.
c) Zeigen Sie, dass $(Z, \oplus, \odot)$ ein Ring ist. Zu welchem Ihnen bekannten Ring ist er isomorph? Geben Sie einen Isomorphismus an. (Für zwei Ringe R und R' ist ein *Isomorphismus* eine bijektive Abbildung $\varphi\colon R \longrightarrow R'$, so dass i) $\varphi(1) = 1$ und für alle $x, y \in R$ ii) $\varphi(x + y) = \varphi(x) + \varphi(y)$ und iii) $\varphi(xy) = \varphi(x)\varphi(y)$ gilt.)

Aufgabe II.4.12.
a) Stellen Sie die Verknüpfungstafel für die multiplikative Gruppe von $\mathbb{F}_{11}$ auf.
b) Bestimmen Sie für $p \in \{23, 53, 89\}$ die ganze Zahl $x \in \{0, ..., p-1\}$, für die

$$13 \cdot x \equiv 1 \bmod p$$

gilt.
c) Bestimmen Sie jeweils die ganze Zahl $x \in \mathbb{Z}$ mit den geforderten Eigenschaften:

- $5^3 \equiv x \bmod 3$, $0 \leq x \leq 2$,
- $4^5 \equiv x \bmod 5$, $0 \leq x \leq 4$,
- $11^7 \equiv x \bmod 7$, $0 \leq x \leq 6$,
- $9^{11} \equiv x \bmod 11$, $0 \leq x \leq 10$,
- $2^{12} \equiv x \bmod 13$, $0 \leq x \leq 12$.

d) Stellen Sie anhand von Teil c) eine Vermutung für eine allgemeine Gesetzmäßigkeit auf und benutzen Sie diese, um die ganze Zahl $x \in \{0, ..., 10\}$ zu bestimmen, die $6^{123} \equiv x \bmod 11$ erfüllt.

III

Vektorräume und lineare Abbildungen

Dieses Kapitel führt die Gegenstände der Linearen Algebra ein, nämlich Vektorräume und lineare Abbildungen. Der zugehörige Formalismus und die fundamentalen Eigenschaften werden ausführlich besprochen.

III.1 Vektorräume

Definition III.1.1. Es sei K ein Körper. Ein *Vektorraum über* K oder auch ein K-*Vektorraum* ist eine Menge V zusammen mit Abbildungen

$$\begin{aligned} +\colon V \times V &\longrightarrow V \quad (\textit{Addition}) \\ (v, w) &\longmapsto v + w, \\ \cdot\colon K \times V &\longrightarrow V \quad (\textit{Skalarmultiplikation}) \\ (\lambda, v) &\longmapsto \lambda \cdot v, \end{aligned}$$

so dass folgende Axiome erfüllt sind:

(A) $(V, +)$ ist eine abelsche Gruppe.
Schreibweise. 0 sei das Neutralelement, $-v$ sei das inverse Element zu $v \in V$.

(S1) $\forall \kappa, \lambda \in K$, $\forall v \in V$: $(\kappa \cdot \lambda) \cdot v = \kappa \cdot (\lambda \cdot v)$. (*Assoziativgesetz der Skalarmultiplikation.*)

(S2) $\forall v \in V : 1 \cdot v = v$.

(S3) $\forall \kappa, \lambda \in K$, $\forall v \in V$: $(\kappa + \lambda) \cdot v = \kappa \cdot v + \lambda \cdot v$.

(S4) $\forall \lambda \in K$, $\forall v, w \in V$: $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$.

((S3) und (S4) sind die *Distributivgesetze der Skalarmultiplikation*.)
Schreibweise. Wir werden wieder einfach von einem „Vektorraum V“ anstatt von einem „Vektorraum $(V, +, \cdot)$“ sprechen.

Vereinbarung. Für das Folgende wird ein Körper K fixiert.

Bemerkungen und Beispiele III.1.2. i) Nach Übung I.2.10 ist K^n ein K-Vektorraum. Im Hinblick auf Anwendungen bei linearen Gleichungssystemen ist er die Hauptmotivation für die obige Definition. Es wird sich aber zeigen, dass die Struktur eines Vektorraums auch in vielen anderen Zusammenhängen natürlich in Erscheinung tritt und daher wieder in voller Allgemeinheit untersucht werden sollte.

ii) Es seien A eine Menge und $\mathrm{Abb}(A, K)$ die Menge aller Abbildungen von A in den Körper K. Für zwei Abbildungen $f, g \in \mathrm{Abb}(A, K)$ definieren wir

$$\begin{array}{rcl} f + g \colon A & \longrightarrow & K \\ x & \longmapsto & f(x) + g(x). \end{array}$$

Für ein Element $\lambda \in K$ und eine Abbildung $f \in \mathrm{Abb}(A, K)$ definieren wir

$$\begin{array}{rcl} \lambda \cdot f \colon A & \longrightarrow & K \\ x & \longmapsto & \lambda \cdot f(x). \end{array}$$

Mit diesen Operationen wird $\mathrm{Abb}(A, K)$ zu einem K-Vektorraum (s. Aufgabe III.1.8). Die Axiome sind einfache Folgerungen aus den Gesetzmäßigkeiten im Körper K. Um z.B. zu überprüfen, dass für alle $\lambda \in K$ und alle $f, g \in \mathrm{Abb}(A, K)$

$$\lambda \cdot (f + g) = \lambda \cdot f + \lambda \cdot g$$

gilt, müssen wir nach Bemerkung II.1.17, i),

$$\forall x \in A : \quad \lambda \cdot (f(x) + g(x)) = \lambda \cdot f(x) + \lambda \cdot g(x)$$

nachweisen. Dies ist aber offensichtlich eine direkte Konsequenz aus dem Distributivgesetz im Körper K.

Dieses Beispiel können wir spezialisieren. Für die Menge $A := \{1, ..., m\} \times \{1, ..., n\}$ schreiben wir

$$\begin{array}{rcl} f \colon A & \longrightarrow & K \\ (i, j) & \longmapsto & a_{ij} \end{array}$$

in der Form $(a_{ij})_{\substack{i=1,\dots,m\\ j=1,\dots,n}}$, also als $(m \times n)$-Matrix. Damit definieren wir

$$\mathrm{Mat}(m, n; K) := \mathrm{Abb}(A, K),$$

den *Vektorraum der* $(m \times n)$*-Matrizen*. Falls $m = n$, so schreiben wir

$$\mathrm{Mat}(n; K) := \mathrm{Mat}(n, n; K).$$

iii) Der Körper $\mathbb{R}$ kann als Vektorraum über dem Körper $\mathbb{Q}$ der rationalen Zahlen aufgefasst werden.

Lemma III.1.3. *In einem* K*-Vektorraum* V *gelten u.A. die folgenden Rechenregeln:*

i) $\forall v \in V : 0 \cdot v = 0$.[1]

ii) $\forall v \in V : (-1) \cdot v = -v$.

iii) $\forall \lambda \in K : \lambda \cdot 0 = 0$.[2]

Beweis. Für i) benutzen wir denselben Trick wie bei Ringen. Es gilt

$$0 \cdot v = (0 + 0) \cdot v \overset{(S4)}{=} 0 \cdot v + 0 \cdot v.$$

Auf beiden Seiten addieren wir das zu $0 \cdot v$ in $(V, +)$ inverse Element und erhalten die Behauptung. Für ii) gehen wir wie folgt vor:

$$0 \overset{i)}{=} 0 \cdot v = \big(1 + (-1)\big) \cdot v \overset{(S4)}{=} 1 \cdot v + (-1) \cdot v \overset{(S2)}{=} v + (-1) \cdot v.$$

Punkt iii) kann man genauso wie i) zeigen. □

Definition III.1.4. Es sei V ein K-Vektorraum. Eine Teilmenge $W \subseteq V$ ist ein *Untervektorraum* oder *linearer Teilraum von* V, wenn gilt:

a) $0 \in W$.

b) $\forall v, w \in W$: $v + w \in W$.

c) $\forall \lambda \in K$, $\forall v \in W : \lambda \cdot v \in W$.

Bemerkungen III.1.5. i) Aus c) und Lemma III.1.3, ii), folgt, dass für jedes $v \in W$ auch $-v \in W$ gilt. Daher ist W eine Untergruppe von $(V, +)$ im Sinne von Definition II.2.7. Insbesondere induziert die Addition eine Abbildung $+ : W \times W \longrightarrow W$. Auf Grund von c) können wir auch $\cdot : K \times W \longrightarrow W$, $(\lambda, v) \longmapsto \lambda \cdot v$, bilden. Dabei ist $\lambda \cdot v$ die Skalarmultiplikation in V. Mit diesen Abbildungen erhält W selber die Struktur eines K-Vektorraums.

ii) Eine Teilmenge $W \subseteq V$ ist genau dann ein Untervektorraum, wenn gilt:

[1] Dabei ist links die Null in K und rechts die in V gemeint.

[2] Hier ist auf beiden Seiten die Null in V gemeint.

α) $W \neq \varnothing$.

β) $\forall \kappa, \lambda \in K$, $\forall v, w \in W : \kappa \cdot v + \lambda \cdot w \in W$.

Der einfache Beweis wird dem Leser als Übungsaufgabe überlassen.

Beispiele III.1.6. i) Für jeden Vektorraum V sind $\{0\}$ und V lineare Teilräume.

ii) Für $A \in \mathrm{Mat}(m, n; K)$ ist $\mathbb{L}(A, 0)$ nach Lemma I.2.4 ein linearer Teilraum von K^n.

iii) Wenn W_i, $i \in I$, eine Familie von Untervektorräumen von V ist, so ist ihr Durchschnitt[3]

$$U := \bigcap_{i \in I} W_i$$

ebenfalls ein Untervektorraum von V. **Zu a).** Da $0 \in W_i$ für alle $i \in I$, liegt 0 auch in U. **Zu b).** Es seien $v, w \in U$. Dies bedeutet $v, w \in W_i$ für alle $i \in I$. Da W_i ein Untervektorraum von V ist, gilt ebenfalls $v + w \in W_i$, $i \in I$. Daher haben wir $v + w \in U$. **Zu c).** Für $\lambda \in K$ und $v \in U$ hat man $v \in W_i$ und damit $\lambda \cdot v \in W_i$ für alle $i \in I$, so dass $\lambda \cdot v \in U$.

iv) Es seien A eine Menge und $B \subseteq A$ eine Teilmenge. Wir definieren

$$W := \Big\{ f \in \mathrm{Abb}(A, K) \,\Big|\, \forall x \in B : f(x) = 0 \Big\}.$$

Dies ist ein linearer Teilraum von $\mathrm{Abb}(A, K)$. Wir wenden dazu das Kriterium ii) aus Bemerkung III.1.5 an. **Zu α).** Offenbar liegt $0\colon A \longrightarrow K$, $x \longmapsto 0$, in W. Deshalb gilt $W \neq \varnothing$. Für $\kappa, \lambda \in K$, zwei Abbildungen $f, g \in W$ und $x \in B$ gilt:

$$\big(\kappa \cdot f + \lambda \cdot g\big)(x) = \kappa \cdot f(x) + \lambda \cdot g(x) = \kappa \cdot 0 + \lambda \cdot 0 = 0.$$

Daher liegt $\kappa \cdot f + \lambda \cdot g$ ebenfalls in W.

Aufgabe III.1.7.
a) Es seien K ein Körper, V ein K-Vektorraum und $W, W' \subseteq V$ zwei lineare Unterräume von V. Zeigen Sie, dass $W \cup W'$ genau dann ein linearer Unterraum von V ist, wenn $W \subseteq W'$ oder $W' \subseteq W$ gilt.
b) Es sei K ein Körper mit nur endlich vielen Elementen (z.B. $\mathbb{F}_2$). Zeigen Sie, dass K^n für jedes $n \geq 2$ Vereinigung von endlich vielen echten Teilräumen ist. (Dabei heißt „echt“, dass der entsprechende Teilraum nicht ganz K^n ist.) Erläutern Sie, warum für einen Körper mit unendlich vielen Elementen K^2 nicht die Vereinigung von endlich vielen echten Teilräumen sein kann.

[3] Als Menge ist dieser Durchschnitt vermöge Komprehension durch

$$\bigcap_{i \in I} W_i := \Big\{ v \in V \,\Big|\, \forall i \in I :\ v \in W_i \Big\}$$

gegeben.

Aufgabe III.1.8.
Es seien M eine Menge, V ein Vektorraum über dem Körper K und $\mathrm{Abb}(M,V)$ die Menge der Abbildungen von M nach V. Zeigen Sie, dass $\mathrm{Abb}(M,V)$ zusammen mit den Abbildungen

$$\begin{array}{rcl} +\colon \mathrm{Abb}(M,V)\times\mathrm{Abb}(M,V) & \longrightarrow & \mathrm{Abb}(M,V) \\ (f,g) & \longmapsto & f+g \end{array}$$

und

$$\begin{array}{rcl} \cdot\colon K\times\mathrm{Abb}(M,V) & \longrightarrow & \mathrm{Abb}(M,V) \\ (\lambda,f) & \longmapsto & \lambda\cdot f \end{array}$$

ein K-Vektorraum ist.

Aufgabe III.1.9.
Gegeben seien zwei K-Vektorräume V und W. Wir definieren folgende Abbildungen

$$\begin{array}{rcl} \cdot\colon K\times(V\times W) & \longrightarrow & V\times W \\ \big(\lambda,(v,w)\big) & \longmapsto & (\lambda v,\lambda w) \end{array}$$

und

$$\begin{array}{rcl} +\colon (V\times W)\times(V\times W) & \longrightarrow & V\times W \\ \big((v,w),(v',w')\big) & \longmapsto & (v+v',w+w'). \end{array}$$

Überprüfen Sie, dass durch diese Abbildungen auf $V\times W$ die Stuktur eines K-Vektorraumes eingeführt wird. Der so erhaltene Vektorraum heißt die *direkte Summe von V und W* und wird mit $V\oplus W$ bezeichnet.

III.2 Erzeugendensysteme

In diesem und dem folgenden Abschnitt führen wir die zentralen Begriffe ein, die es uns gestatten, die Struktur eines Vektorraums zu untersuchen und adäquat zu beschreiben.

Sprech- und Schreibweise. i) Es seien I und A Mengen und x_i, $i\in I$, eine durch I indizierte Familie von Elementen von A. Wir sagen, *fast alle Elemente x_i haben Eigenschaft $\mathcal{E}$*, wenn für alle bis auf endlich viele Indizes $i\in I$ das Element x_i die Eigenschaft $\mathcal{E}$ hat:

$$\#\big\{\, i\in I \,|\, x_i \text{ hat Eigenschaft } \mathcal{E} \textbf{ nicht} \,\big\} < \infty.$$

ii) Es seien V ein K-Vektorraum und v_i, $i \in I$, eine durch I indizierte Familie von Elementen aus V, die fast alle null sind. Dann setzen wir

$$\sum_{i\in I} v_i := \sum_{i\in I: v_i\neq 0} v_i.$$

Die rechte Seite ist als Summe von nur endlich vielen Elementen wohldefiniert.

Definitionen III.2.1. i) Es seien $v_1, ..., v_m$ Elemente des K-Vektorraums V. Eine *Linearkombination von* $v_1, ..., v_m$ ist ein Element $v \in V$, zu dem es Zahlen $\lambda_1, ..., \lambda_m \in K$ mit

$$v = \lambda_1 \cdot v_1 + \cdots + \lambda_m \cdot v_m$$

gibt.

ii) Es sei $M \subseteq V$ eine beliebige Teilmenge von V. Die *lineare Hülle* oder das *lineare Erzeugnis* von M ist die Menge

$$\begin{aligned}\langle M\rangle &:= \left\{ v \in V \,\middle|\, v \text{ ist Linearkombination v. endlich vielen Elementen aus } M \right\} \\ &= \left\{ \sum_{v\in M} \lambda_v \cdot v \,\middle|\, \lambda_v \in K,\ v \in M,\ \text{fast alle null} \right\}.\end{aligned}$$

Vereinbarungen. $\langle\varnothing\rangle := \{0\}$; $\langle v_1, ..., v_m\rangle := \langle\{v_1, ..., v_m\}\rangle$.

Spezialfall. Es sei W_i, $i \in I$, eine Familie von linearen Teilräumen von V. Dann heißt[4]

$$\begin{aligned}\sum_{i\in I} W_i &:= \left\langle \bigcup_{i\in I} W_i \right\rangle \\ &= \left\{ \sum_{i\in I} w_i \,\middle|\, w_i \in W_i,\ i \in I,\ \text{fast alle null} \right\}\end{aligned}$$

die *Summe der Unterräume* W_i, $i \in I$. Falls für alle Familien $w_i \in W_i$, $i \in I$, die fast alle null sind,

$$\sum_{i\in I} w_i = 0 \quad\Longrightarrow\quad \forall i \in I : w_i = 0$$

gilt, dann sprechen wir von der *direkten Summe der* W_i, $i \in I$, und schreiben

$$\bigoplus_{i\in I} W_i := \sum_{i\in I} W_i.$$

[4]Als Menge kann diese Vereinigung durch

$$\bigcup_{i\in I} W_i := \left\{ v \in V \,\middle|\, \exists i \in I :\ v \in W_i \right\}$$

beschrieben werden und ist daher durch Anwendung der Komprehension definiert.

iii) Eine Teilmenge $M \subseteq V$ heißt ein *Erzeugendensystem von* V, wenn

$$\langle M \rangle = V$$

gilt.

Die zentralen Eigenschaften der linearen Hülle werden in Aufgabe III.2.4 besprochen.

Beispiele III.2.2. i) V ist trivialerweise ein Erzeugendensystem für den K-Vektorraum V.

ii) Eine Menge $\left\{ \binom{a}{c}, \binom{b}{d} \right\}$ ist genau dann ein Erzeugendensystem für K^2, wenn $\delta(a, b, c, d) \neq 0$ gilt (s. Aufgabe I.2.9).

iii)

$$\left\langle \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ -5 \\ 0 \end{pmatrix} \right\rangle = \left\{ \begin{pmatrix} s \\ t \\ 0 \end{pmatrix} \,\middle|\, s, t \in \mathbb{R} \right\}.$$

iv) Wir setzen

$$e_i := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in K^n,$$

1 der i-te Eintrag, $i = 1, ..., n$. Dann gilt

$$\forall s = \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} \in K^n : \quad s = s_1 \cdot e_1 + \cdots + s_n \cdot e_n,$$

so dass $\langle e_1, ..., e_n \rangle = K^n$.

Zusammenhang mit linearen Gleichungssystemen. — Es seien

$$v_j := \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \in K^m, \quad j = 1, ..., n.$$

Wir definieren

$$A := (v_1 | \cdots | v_n) := \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in \mathrm{Mat}(m, n; K)$$

als die $(m \times n)$-Matrix mit den Spalten $v_1, ..., v_n$. Mit (I.5) sieht man, dass gilt:

$$b \in \langle v_1, ..., v_n \rangle \quad \Longleftrightarrow \quad \mathbb{L}(A, b) \neq \varnothing. \tag{III.1}$$

Umgekehrt seien $A \in \mathrm{Mat}(m, n; K)$ und $v_1, ..., v_n \in K^m$ die Spalten der Matrix A. Dann hat man

$$\langle v_1, ..., v_n \rangle = \left\{ b \in K^m \,|\, \mathbb{L}(A, b) \neq \varnothing \right\}.$$

Folgerung III.2.3. *Wenn $\{v_1, ..., v_n\}$ ein Erzeugendensystem für K^m ist, dann gilt $n \geq m$.*

Beweis. Wir nehmen an, es sei $n < m$. Mit dem Gauß-Algorithmus überführen wir die Matrix A in eine Matrix A' in Zeilenstufenform. Die Matrix A' hat $r \leq n < m$ Stufen, so dass die letzte Zeile von A' eine Nullzeile ist. Für $b' := e_m$ gilt daher $\mathbb{L}(A', b') = \varnothing$. Nun können wir die an A durchgeführten Zeilenoperationen wieder rückgängig machen, d.h. wir können die erweiterte Koeffizientenmatrix $(A'|b')$ durch Zeilenoperationen in eine Matrix der Form $(A|b)$ mit $b \in K^m$ überführen. Es gilt

$$\mathbb{L}(A, b) \overset{\text{Lemma I.4.1}}{=} \mathbb{L}(A', b') = \varnothing,$$

so dass nach (III.1) $b \notin \langle v_1, ..., v_n \rangle$. Dies befindet sich im Widerspruch zu unserer Annahme. □

Aufgabe III.2.4.
Es seien V ein K-Vektorraum und M, M' Teilmengen von V. Überprüfen Sie die folgenden Eigenschaften der linearen Hülle.

a) $\langle M \rangle$ ist ein linearer Teilraum von V, der M enthält.

b) Ein linearer Unterraum $U \subseteq V$, der M enthält, enthält auch die lineare Hülle $\langle M \rangle$. (Somit ist $\langle M \rangle$ der „kleinste“ lineare Teilraum von V, der M enthält.)

c) Es gilt

$$\langle M \rangle = \bigcap_{\substack{U \subseteq V \text{ UR:} \\ M \subseteq U}} U.$$

d) Gilt $M \subseteq M'$, dann gilt auch $\langle M \rangle \subseteq \langle M' \rangle$.

III.3 Linear unabhängige Teilmengen

Für eine Teilmenge M eines Vektorraums V haben wir den Unterraum $\langle M \rangle \subseteq V$ aller Linearkombinationen von Elementen aus M eingeführt. Nun wollen wir diejenigen Teilmengen M hervorheben, für die sich jedes Element $v \in \langle M \rangle$ in **eindeutiger Weise** als Linearkombination von Elementen aus M schreiben lässt.

Definition III.3.1. Es sei V ein K-Vektorraum. Eine Teilmenge $M \subseteq V$ heißt *linear unabhängig* oder *frei*, wenn für jede Familie λ_x, $x \in M$, von Elementen aus K, die fast alle null sind, aus

$$\sum_{x \in M} \lambda_x \cdot x = 0$$

bereits

$$\forall x \in M : \quad \lambda_x = 0$$

folgt.

Eine Teilmenge $M \subseteq V$ heißt *linear abhängig*, wenn sie nicht linear unabhängig ist.

Bemerkung III.3.2. Für jede Teilmenge $M \subseteq V$ gilt

$$0 = \sum_{x \in M} 0 \cdot x \in \langle M \rangle.$$

Wenn M linear unabhängig ist, ist dies bereits die einzige Möglichkeit, 0 als Linearkombination von Elementen aus M zu schreiben. Es ist leicht einzusehen (vgl. Lemma III.4.2), dass für jedes Element $v \in \langle M \rangle$ die Koeffizienten $\lambda_x \in K$, $x \in M$, mit

$$v = \sum_{x \in M} \lambda_x \cdot x$$

eindeutig bestimmt sind.

Beispiele III.3.3. i) Die leere Menge $\varnothing \subsetneq V$ ist linear unabhängig.

ii) Es sei $v \in V$. Die Menge $\{v\} \subseteq V$ ist genau dann linear unabhängig, wenn $v \neq 0$. Für $v = 0$ und $1 \in K$ gilt $1 \cdot v = 0$, so dass $\{v\}$ linear abhängig ist. Es sei umgekehrt $\alpha \in K^*$ gegeben, so dass $\alpha \cdot v = 0$. Dann gilt

$$v = 1 \cdot v = (\alpha^{-1} \cdot \alpha) \cdot v = \alpha^{-1} \cdot (\alpha \cdot v) = \alpha^{-1} \cdot 0 = 0.$$

iii) Es seien $e_i \in K^n$ die in Beispiel III.2.2, iv), definierten Vektoren, $i = 1, ..., n$. Die Menge $\{e_1, ..., e_n\}$ ist linear unabhängig (vgl. Teil v).

iv) Auch der Begriff der linearen Unabhängigkeit lässt sich im Zusammenhang mit linearen Gleichungssystemen verstehen. Dazu betrachten wir Vektoren

$$v_j := \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \in K^m, \quad j = 1, ..., n,$$

und definieren die zugehörige Matrix

$$A := (v_1 | \cdots | v_n) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

Es folgt

$$\{ v_1, ..., v_n \} \text{ ist linear unabhängig} \iff \mathbb{L}(A, 0) = \{0\}.$$

Als interessante Beobachtung ergibt sich aus Folgerung I.4.5, i), dass $\{ v_1, ..., v_n \}$ für $n > m$ linear abhängig ist.

v) Es sei M eine Menge. Für $x \in M$ definieren wir die Abbildung

$$\begin{array}{rcl} e_x \colon M & \longrightarrow & K \\ y & \longmapsto & \begin{cases} 1, & y = x \\ 0, & \text{sonst} \end{cases} . \end{array}$$

Die Menge $\{ e_x \mid x \in M \} \subseteq \mathrm{Abb}(M, K)$ (vgl. Beispiel III.1.2, ii) ist linear unabhängig. Dazu sei λ_x, $x \in M$, eine Familie von Elementen aus K, die fast alle null sind. Für alle $y \in M$ gilt

$$\Big(\sum_{x \in M} \lambda_x \cdot e_x \Big)(y) = \lambda_y. \tag{III.2}$$

Damit ist die Aussage evident.

Falls M unendlich viele Elemente enthält, ist $\{ e_x \mid x \in M \}$ **kein** Erzeugendensystem. Die konstante Abbildung

$$\begin{array}{rcl} k_1 \colon M & \longrightarrow & K \\ x & \longmapsto & 1 \end{array}$$

liegt dann nicht in der linearen Hülle $\langle e_x \mid x \in M \rangle$. Dies folgt aus (III.2).

vi) Es seien $M \subseteq V$ eine linear unabhängige Teilmenge und $M' \subseteq M$. Dann ist auch M' linear unabhängig.

vii) Es seien $M \subseteq V$ eine linear abhängige Teilmenge und $M \subseteq M'$. Dann ist M' ebenfalls linear abhängig.

Aufgabe III.3.4.
Stellen Sie fest, ob die folgenden Mengen von Vektoren im jeweiligen $\mathbb{R}^n$ linear unabhängig sind.

a) $v_1 := \begin{pmatrix} 1 \\ -1 \\ 5 \\ 2 \end{pmatrix}$, $v_2 := \begin{pmatrix} -2 \\ 0 \\ -1 \\ 1 \end{pmatrix}$ und $v_3 := \begin{pmatrix} 1 \\ 3 \\ 4 \\ 2 \end{pmatrix}$.

b) $v_1 := \begin{pmatrix} 1 \\ 2 \\ -4 \end{pmatrix}$, $v_2 := \begin{pmatrix} -3 \\ 5 \\ 0 \end{pmatrix}$ und $v_3 := \begin{pmatrix} -4 \\ 0 \\ -3 \end{pmatrix}$.

III.4 Basen

Definition III.4.1. Es sei V ein K-Vektorraum. Eine Teilmenge $B \subseteq V$ wird *Basis* genannt, wenn sie sowohl eine linear unabhängige Teilmenge als auch ein Erzeugendensystem ist.

Den wahren Nutzen des Begriffs einer Basis zeigt das folgende Resultat auf.

Lemma III.4.2. *Eine Teilmenge* $B \subseteq V$ *ist genau dann eine Basis, wenn es zu jedem Vektor* $v \in V$ *eindeutig bestimmte Zahlen* $\lambda_b \in K$, $b \in B$, *gibt, die fast alle null sind, so dass*

$$v = \sum_{b \in B} \lambda_b \cdot b. \tag{III.3}$$

Beweis. Nehmen wir zunächst an, die Teilmenge B habe die genannten Eigenschaften. Dann ist B offenbar ein Erzeugendensystem für V. Es seien $\lambda_b \in K$, $b \in B$, fast alle null, mit

$$\sum_{b \in B} \lambda_b \cdot b = 0 = \sum_{b \in B} 0 \cdot b$$

gegeben. Aus der Eindeutigkeit der λ_b in (III.3) folgt $\lambda_b = 0$ für alle $b \in B$, so dass B auch linear unabhängig ist.

Es sei umgekehrt B eine Basis für V. Dann ist B ein Erzeugendensystem. Jeder Vektor $v \in V$ besitzt daher eine Darstellung wie in (III.3). Es bleibt, die Eindeutigkeit der Koeffizienten $\lambda_b \in K$, $b \in B$, zu überprüfen. Dazu seien $v \in V$ und $\lambda_b \in K$ und $\lambda'_b \in K$, $b \in B$, zwei Sätze von Zahlen, so dass

$$\sum_{b \in B} \lambda_b \cdot b = v = \sum_{b \in B} \lambda'_b \cdot b.$$

Wir erhalten

$$0 = v - v = \sum_{b \in B} \lambda_b \cdot b - \sum_{b \in B} \lambda'_b \cdot b = \sum_{b \in B} (\lambda_b - \lambda'_b) \cdot b.$$

Da B linear unabhängig ist, folgt $\lambda_b - \lambda'_b = 0$, i.e. $\lambda_b = \lambda'_b$ für alle $b \in B$. □

Beispiele III.4.3. i) Es sei $M \subseteq V$ eine linear unabhängige Teilmenge. Dann ist M eine Basis für $\langle M \rangle$.

ii) Wenn B eine Basis für K^n ist, dann gilt $\#B = n$. Denn B ist ein Erzeugendensystem, so dass $\#B \geq n$ nach Folgerung III.2.3. Nach Beispiel III.3.3, iv), folgt $\#B \leq n$ aus der linearen Unabhängigkeit von B. Man beachte, dass $\{e_1, ..., e_n\}$ eine Basis für K^n ist. Allgemeiner gilt (Folgerung I.4.5, ii):

$$\{v_1, ..., v_n\} \text{ ist eine Basis für } K^n \iff \mathbb{L}(A, 0) = \{0\}, \quad A := (v_1 | \cdots | v_n).$$

iii) Für jede Matrix $A \in \mathrm{Mat}(m, n; K)$ liefert der Gauß-Algorithmus eine Basis für den Lösungsraum $\mathbb{L}(A, 0)$ (s. Satz I.4.2, ii), und I.4.4).

iv) Es sei M eine endliche Menge. Aus Beispiel III.3.3, v), folgt leicht, dass $\{e_x \mid x \in M\}$ eine Basis für $\mathrm{Abb}(M, K)$ ist. Dies können wir insbesondere auf

$M = \{1, ..., m\} \times \{1, ..., n\}$ anwenden (vgl. Beispiel III.1.2, ii). Für $(i, j) \in M$ erhalten wir in Matrix-Schreibweise

$$E_{ij} := e_{(i,j)} = \begin{array}{cc} \begin{array}{ccccccccc} & & & \text{j-te Spalte} & & & & & \\ & & & \downarrow & & & & & \end{array} & \\ \begin{pmatrix} & \mathbf{0} & & 0 & & & \mathbf{O} & & \\ & & & \vdots & & & & & \\ & & & 0 & & & & & \\ 0 & \cdots & 0 & 1 & 0 & \cdots & \cdots & \cdots & 0 \\ & & & 0 & & & & & \\ & \mathbf{0} & & \vdots & & & \mathbf{O} & & \\ & & & \vdots & & & & & \\ & & & \vdots & & & & & \\ & & & 0 & & & & & \end{pmatrix} & \begin{array}{l} \\ \\ \\ \leftarrow \text{i-te Zeile} \\ \\ \\ \\ \\ \\ \end{array} \end{array}.$$

Somit ist $\{E_{ij} \mid i = 1, ..., m,\ j = 1, ..., n\}$ eine Basis für $\mathrm{Mat}(m, n; K)$.

Der folgende Satz enthält weitere Charakterisierungen für Basen.

Satz III.4.4. *Es seien V ein K-Vektorraum und $B \subseteq V$ eine Teilmenge. Dann sind die folgenden Eigenschaften äquivalent.*

i) B *ist eine Basis für* V.

ii) B *ist ein minimales Erzeugendensystem für* V, *d.h. für jede echte Teilmenge* $B' \subsetneq B$ *gilt* $\langle B' \rangle \subsetneq V$.

iii) B *ist eine maximale linear unabhängige Teilmenge, d.h. jede Menge* B', *die* B *echt enthält (i.e.* $B \subsetneq B'$*), ist linear abhängig.*

Beweis. Wir zeigen zunächst die Implikation „i)$\Longrightarrow$iii)“. Nach Definition einer Basis ist B linear unabhängig. Es sei nun $B \subsetneq B'$. Wir wählen ein Element $v_0 \in B' \setminus B$. Da B eine Basis ist, finden wir Elemente $\lambda_b \in K$, $b \in B$, die fast alle null sind, so dass

$$v_0 = \sum_{b \in B} \lambda_b \cdot b.$$

Wir definieren nun für $x \in B'$

$$\kappa_x := \begin{cases} 1, & x = v_0 \\ -\lambda_x, & x \in B \\ 0, & \text{sonst} \end{cases}.$$

Die Elemente $\kappa_x \in K$, $x \in B'$, sind fast alle null, und es gilt

$$\sum_{x \in B'} \kappa_x \cdot x = v_0 + \sum_{b \in B} (-\lambda_b) \cdot b = v_0 - \sum_{b \in B} \lambda_b \cdot b = v_0 - v_0 = 0.$$

Da $\kappa_{v_0} \neq 0$, zeigt dies, dass B' linear abhängig ist.

Wir kommen nun zum Beweis von „iii)$\Longrightarrow$ii)“. Als erstes wird nachgewiesen, dass B ein Erzeugendensystem ist. Wenn das nicht so wäre, dann gäbe es ein Element $v_0 \in V \setminus \langle B \rangle$. Wir behaupten, dass in diesem Fall die Menge $B' := B \cup \{v_0\}$ immer noch linear unabhängig wäre. Dies stünde aber im Widerspruch zu der Maximalität von B. Betrachten wir also eine Relation

$$\lambda_0 \cdot v_0 + \sum_{b \in B} \lambda_b \cdot b = 0. \tag{III.4}$$

Wenn $\lambda_0 = 0$ ist, dann gilt $\sum_{b \in B} \lambda_b \cdot b = 0$ und daher $\lambda_b = 0$ für alle $b \in B$, denn B ist linear unabhängig. Falls $\lambda_0 \neq 0$, dann ergibt (III.4)

$$v_0 = \sum_{b \in B} (-\lambda_0^{-1} \cdot \lambda_b) \cdot b,$$

d.h. $v_0 \in \langle B \rangle$. Das steht im Widerspruch zu unserer Annahme, so dass B' in der Tat linear unabhängig ist.

Es bleibt noch zu überprüfen, dass B ein minimales Erzeugendensystem ist. Es sei andernfalls $B' \subsetneq B$ eine Teilmenge, die noch erzeugend ist. Es sei $v_0 \in B \setminus B'$. Da B' erzeugend ist, gibt es Zahlen $\lambda_b \in K$, $b \in B'$, fast alle null, so dass

$$v_0 = \sum_{b \in B'} \lambda_b \cdot b.$$

Definieren wir für $b \in B$

$$\kappa_b := \begin{cases} -1, & b = v_0 \\ \lambda_b, & b \in B' \\ 0, & \text{sonst} \end{cases},$$

dann erhalten wir

$$\sum_{b \in B} \kappa_b \cdot b = v_0 - v_0 = 0.$$

Dies widerspricht der linearen Unabhängigkeit von B.

Zu guter Letzt wird „ii)$\Longrightarrow$i)“ gezeigt. Offenbar ist B ein Erzeugendensystem. Wir betrachten eine Relation

$$\sum_{b \in B} \lambda_b \cdot b = 0.$$

Es sei $v_0 \in B$, so dass $\lambda_{v_0} \neq 0$. Dann gilt also

$$v_0 = \sum_{b \in B \setminus \{v_0\}} (-\lambda_{v_0}^{-1} \cdot \lambda_b) \cdot b,$$

d.h. $v_0 \in \langle B \setminus \{v_0\} \rangle$ und damit $B \subseteq \langle B \setminus \{v_0\} \rangle$ und schließlich

$$\langle B \rangle \subseteq \langle B \setminus \{v_0\} \rangle \subsetneq V. \tag{III.5}$$

Im letzten Schritt gilt die strikte Ungleichung, weil B ein minimales Erzeugendensystem ist. Relation (III.5) kann offensichtlich nicht wahr sein. □

Folgerung III.4.5. *Es sei V ein K-Vektorraum. Wenn es ein Erzeugendensystem M für V gibt, das nur endlich viele Elemente enthält, dann besitzt V eine Basis.*

Beweis. Wir entfernen solange Elemente aus M, bis wir ein minimales Erzeugendensystem und damit nach Satz III.4.4 eine Basis erhalten.

Formal ist dies eine vollständige Induktion über $n := \#M$. Für $n = 0$ gilt $V = \{0\}$, und $\varnothing$ ist eine Basis. Für den Schluß „$n \rightsquigarrow n+1$“ gehen wir folgendermaßen vor. Wenn M ein minimales Erzeugendensystem ist, dann ist M nach Satz III.4.4 bereits eine Basis. Ansonsten gibt es ein $v_0 \in M$, so dass $M' := M \setminus \{v_0\}$ ein Erzeugendensystem mit n Elementen ist. Nach Induktionsvoraussetzung besitzt V dann eine Basis. □

Satz III.4.6. *Es sei V ein K-Vektorraum. Wenn V eine endliche Basis B besitzt, dann ist jede andere Basis B' von V ebenfalls endlich, und es gilt*

$$\#B' = \#B.$$

Dieser Satz bildet die Grundlage für den Dimensionsbegriff für endlichdimensionale Vektorräume:

Definition III.4.7. Ein Vektorraum V heißt *endlichdimensional*, wenn es eine Basis B für V mit endlich vielen Elementen gibt.
Schreibweise. $\mathrm{Dim}_K(V) < \infty$.
In diesem Fall wird

$$\mathrm{Dim}_K(V) := \#B$$

die *Dimension von* V genannt, B eine Basis von V.

Ist V nicht endlichdimensional, so sagt man, V ist *unendlichdimensional*, und schreibt $\mathrm{Dim}_K(V) = \infty$.

Man beachte, dass Satz III.4.6 garantiert, dass die Begriffe „endlichdimensional“ und „Dimension“ sinnvoll sind.

Beispiel III.4.8. $\mathrm{Dim}_K(K^n) = n$; $\mathrm{Dim}_K\big(\mathrm{Mat}(m,n;K)\big) = m \cdot n$.

Bevor wir uns dem Beweis von Satz III.4.6 widmen können, benötigen wir noch einige Vorbereitungen.

Lemma III.4.9 (Austauschlemma)**.** *Es seien V ein K-Vektorraum, $B = \{b_1, ..., b_n\}$ eine Basis für V und $w = \sum_{b \in B} \lambda_b \cdot b \in V \setminus \{0\}$. Falls $\lambda_{b'} \neq 0$ für das Element $b' \in B$ gilt, dann ist*

$$B' := \big(B \setminus \{b'\}\big) \cup \{w\}$$

ebenfalls eine Basis für V.

Der Vektor b' kann also gegen den Vektor w „ausgetauscht“ werden.

Beweis. Indem wir gegebenenfalls neu nummerieren, können wir $b' = b_1$ annehmen. Wir schreiben $w = \lambda_1 \cdot b_1 + \lambda_2 \cdot b_2 + \cdots + \lambda_n \cdot b_n$ mit $\lambda_1 \neq 0$.

Wir zeigen nun, dass $B' = \{w, b_2, ..., b_n\}$ ein Erzeugendensystem ist. Es reicht zu zeigen, dass b_1 in $\langle w, b_2, ..., b_n \rangle$ liegt. Das ist richtig, weil

$$b_1 = \lambda_1^{-1} \cdot w + (-\lambda_1^{-1} \cdot \lambda_2) \cdot b_2 + \cdots + (-\lambda_1^{-1} \cdot \lambda_n) \cdot b_n$$

gilt.

Als nächstes betrachten wir eine Linearkombination

$$\begin{aligned} 0 &= \kappa_1 \cdot w + \kappa_2 \cdot b_2 + \cdots + \kappa_n \cdot b_n \\ &= \kappa_1 \cdot (\lambda_1 \cdot b_1 + \lambda_2 \cdot b_2 + \cdots + \lambda_n \cdot b_n) + \kappa_2 \cdot b_2 + \cdots + \kappa_n \cdot b_n \\ &= (\kappa_1 \cdot \lambda_1) \cdot b_1 + (\kappa_1 \cdot \lambda_2 + \kappa_2) \cdot b_2 + \cdots + (\kappa_1 \cdot \lambda_n + \kappa_n) \cdot b_n. \end{aligned}$$

Da B linear unabhängig ist, folgt

$$\kappa_1 \cdot \lambda_1 = \kappa_1 \cdot \lambda_2 + \kappa_2 = \cdots = \kappa_1 \cdot \lambda_n + \kappa_n = 0.$$

Aus $\lambda_1 \neq 0$ folgt $\kappa_1 = 0$ und daraus $\kappa_2 = \cdots = \kappa_n = 0$. Somit ist B' als linear unabhängig nachgewiesen. □

Satz III.4.10 (Steinitzscher[5] Austauschsatz)**.** *Es seien V ein K-Vektorraum, $B = \{b_1, ..., b_m\}$ eine Basis für V und $F \subseteq V$ eine linear unabhängige Teilmenge. Dann gilt:*

i) *F hat höchstens m Elemente.*

ii) *Es gibt eine Teilmenge $\widetilde{B} \subseteq B$ mit $m - \#F$ Elementen, so dass*

$$B' := \widetilde{B} \cup F$$

eine Basis von V ist, die m Elemente enthält.

Beweis. Es sei $A(n)$, $n \in \mathbb{N}$, die folgende Aussage:

Für jeden K-Vektorraum V, jede linear unabhängige Teilmenge $F \subseteq V$ mit n Elementen und jede Basis $B = \{b_1, ..., b_m\}$ mit $m \geq n$ Elementen gibt es eine Teilmenge $\widetilde{B} \subseteq B$ mit $m - n$ Elementen, so dass

$$B' := \widetilde{B} \cup F$$

eine Basis von V ist, die m Elemente enthält.

[5] Ernst Steinitz (1871 - 1928), deutscher Mathematiker.

Wir beweisen die Gültigkeit von $A(n)$ für alle $n \in \mathbb{N}$ durch vollständige Induktion. Für $n = 0$ muss $F = \varnothing$ gelten, und wir können $\widetilde{B} = B$ wählen. Offenbar ist dann $B' = B \cup \varnothing = B$ eine Basis mit m Elementen. Jetzt schließen wir von $A(n)$ auf $A(n+1)$. Für V, F und B wie in der Annahme suchen wir uns ein Element $v_0 \in F$ aus und setzen $F' := F \setminus \{v_0\}$. Nach Induktionsvoraussetzung gibt es eine Teilmenge $\overline{B} \subseteq B$ mit $m-n$ Elementen, so dass $B'' = \overline{B} \cup F'$ eine Basis von V mit m Elementen ist. Da F linear unabhängig ist, gilt $v_0 \notin \langle F' \rangle$. Schreiben wir

$$v_0 = \sum_{b \in \overline{B}} \lambda_b \cdot b,$$

dann existiert folglich ein Element $b_0 \in \overline{B}$ mit $\lambda_{b_0} \neq 0$. Nach dem Austauschlemma III.4.9 ist

$$B' := (\underbrace{\overline{B} \setminus \{b_0\}}_{=:\widetilde{B}}) \cup F' \cup \{v_0\} = \widetilde{B} \cup F$$

eine Basis für V. Offenbar gilt $\#\widetilde{B} = \#\overline{B} - 1 = m - n - 1 = m - (n+1)$, und wegen $\widetilde{B} \cap F = \varnothing$ gilt $\#B' = \#\widetilde{B} + \#F = m$.

Nun können wir i) beweisen. Die linear unabhängige Teilmenge F habe mindestens $m+1$ Elemente. Daher können wir eine Teilmenge $F' \subsetneq F$ mit genau m Elementen auswählen. Aus $A(m)$ folgt, dass F' eine Basis ist. Nach Satz III.4.4 ist F' eine maximale linear unabhängige Teilmenge. Daher muss F linear abhängig sein, im Widerspruch zur Annahme. Teil ii) folgt schließlich aus i) und $A(n)$, $n := \#F$. □

Beweis von Satz III.4.6. Es seien B und B' zwei Basen von V. Da B' eine linear unabhängige Teilmenge von V ist, impliziert Satz III.4.10, i), $\#B' \leq \#B$. Ebenso können wir die Basis B' zu Grunde legen und $\#B \leq \#B'$ für die linear unabhängige Teilmenge B schließen. □

Folgerung III.4.11. *Es seien V ein endlichdimensionaler K-Vektorraum und $W \subseteq V$ ein Unterraum. Dann ist auch W endlichdimensional mit $\mathrm{Dim}_K(W) \leq \mathrm{Dim}_K(V)$.*

Beweis. Nach Satz III.4.10, i), hat jede linear unabhängige Teilmenge von W höchstens $\mathrm{Dim}_K(V)$ Elemente. Es sei $F \subseteq W$ eine linear unabhängige Teilmenge mit der maximalen Anzahl von Elementen. Dann ist F offenkundig eine maximale linear unabhängige Teilmenge von W und somit eine Basis. Also ist W endlichdimensional und $\mathrm{Dim}_K(W) = \#F \leq \mathrm{Dim}_K(V)$. □

Zeilen- und Spaltenrang von Matrizen. — Mit Hilfe des Dimensionsbegriffs können wir nun die in Problem I.4.6, i), gesuchte Invariante von Matrizen erklären.

Definitionen III.4.12. i) Der *Vektorraum der Zeilenvektoren* ist

$$ZK^n := \mathrm{Mat}(1, n; K) = \Big\{ (s_1 \cdots s_n) \,|\, s_i \in K,\ i = 1, ..., n \Big\}.$$

ii) Es seien

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in \mathrm{Mat}(m, n; K)$$

eine $(m \times n)$-Matrix,

$$v_j := \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \in K^m$$

die j-te Spalte von A, $j = 1, ..., n$, und

$$z_i = (a_{i1} \cdots a_{in}) \in ZK^n$$

die i-te Zeile, $i = 1, ..., m$. Die Zahl

$$\text{S-Rg}(A) := \mathrm{Dim}_K \langle v_1, ..., v_n \rangle$$

ist der *Spaltenrang* von A und die Zahl

$$\text{Z-Rg}(A) := \mathrm{Dim}_K \langle z_1, ..., z_m \rangle$$

der *Zeilenrang*.

Beispiel III.4.13. Es sei A eine Matrix in Zeilenstufenform wie in Satz I.3.3, i), beschrieben. Dann gilt

$$\text{S-Rg}(A) = r = \text{Z-Rg}(A).$$

Satz III.4.14. *Es seien $A, A' \in \mathrm{Mat}(m, n; K)$ zwei $(m \times n)$-Matrizen.*

i) *Die Matrix A' gehe aus der Matrix A durch eine Zeilenoperation vom Typ* (ZI)*,* (ZII) *oder* (ZIII) *(s. Definition I.3.1) hervor. Dann gilt*

$$\text{Z-Rg}(A') = \text{Z-Rg}(A).$$

ii) *Falls die Matrix A' aus der Matrix A durch eine Spaltenoperation vom Typ* (SI)*,* (SII) *oder* (SIII) *(s. Aufgabe I.4.10) hervorgeht, so gilt*

$$\text{S-Rg}(A') = \text{S-Rg}(A).$$

Beweis. Der Beweis wird der Leserin bzw. dem Leser als Übungsaufgabe überlassen (Aufgabe III.4.21). □

Folgerung III.4.15. *Für jede Matrix* $A \in \mathrm{Mat}(m, n; K)$ *gilt*

$$Z\text{-}\mathrm{Rg}(A) + \mathrm{Dim}_K\big(\mathbb{L}(A, 0)\big) = n.$$

Bemerkung III.4.16. Mit dem Zeilenrang haben wir die in Problem I.4.6, i), gesuchte Invariante gefunden.

Beispiel III.4.13 suggeriert das folgende

Problem III.4.17. Gilt $S\text{-}\mathrm{Rg}(A) = Z\text{-}\mathrm{Rg}(A)$ für jede Matrix A?

Die Frage lautet also, ob Zeilenumformungen den Spaltenrang einer Matrix ändern können. Sie lässt sich elegant mit Hilfe der Theorie linearer Abbildungen beantworten, der wir uns im nächsten Abschnitt widmen möchten.

Aufgabe III.4.18.
Bestimmen Sie die Dimension der linearen Hülle der folgenden Zeilenvektoren im ZK^4:

$$v_1 := (1, -3, 7, -1),\ v_2 := (2, 0, -4, 3),\ v_3 := (2, 6, -22, 8) \text{ und } v_4 := (3, 3, -15, 7).$$

Aufgabe III.4.19.
Geben Sie eine Basis für $\mathrm{Abb}'(\mathbb{N}, K)$ (s. Aufgabe II.3.7, b) an.

Aufgabe III.4.20.
Gegeben seien zwei endlichdimensionale K-Vektorräume V und W der Dimension m bzw. n. Welche Dimension hat der Vektorraum $V \oplus W$ aus Aufgabe III.1.9?

Aufgabe III.4.21.
Zeigen Sie: Es seien A, $A' \in \mathrm{Mat}(m, n; K)$. Dabei gehe A' aus A durch eine Folge von Zeilenoperationen vom Typ (ZI)-(ZIII) (bzw. Spaltenoperationen vom Typ (SI)-(SIII)) hervor. Dann folgt

$$Z\text{-}\mathrm{Rg}(A') = Z\text{-}\mathrm{Rg}(A) \quad \big(\text{bzw. } S\text{-}\mathrm{Rg}(A') = S\text{-}\mathrm{Rg}(A)\big).$$

III.5 Lineare Abbildungen

Vektorräume sind Mengen, die mit den Zusatzstrukturen „Addition“ und „Skalarmultiplikation“ ausgestattet sind. Die Theorie der Vektorräume wird durch diejenigen Abbildungen, die diese Zusatzstrukturen respektieren, entscheidend bereichert.

Definition III.5.1. Es seien V und W zwei K-Vektorräume. Eine Abbildung $f\colon V \longrightarrow W$ ist eine *lineare Abbildung* oder ein (*Vektorraum-*)*Homomorphismus, wenn sie folgende Eigenschaften aufweist:*

- $\forall \lambda \in K \forall v \in V\colon f(\lambda \cdot v) = \lambda \cdot f(v)$,

- $\forall v, v' \in V$: $f(v + v') = f(v) + f(v')$.

Beobachtung III.5.2. *Es seien V, W zwei K-Vektorräume und $f\colon V \longrightarrow W$ eine lineare Abbildung. Dann gilt $f(0) = 0$.*

Beweis. Es sei $v \in V$. Dann hat man nach Lemma III.1.3, i),

$$f(0) = f(0 \cdot v) = 0 \cdot f(v) = 0.$$

Dies zeigt die Behauptung. □

Beispiele III.5.3. i) Für jeden K-Vektorraum V ist $\mathrm{Id}_V\colon V \longrightarrow V$ eine lineare Abbildung.

ii) Es seien V, W zwei K-Vektorräume. Dann ist $0\colon V \longrightarrow W$, $v \longmapsto 0$, eine lineare Abbildung.

iii) Es seien V ein K-Vektorraum und $\alpha \in K$. Die zugehörige *Homothetie* ist die Abbildung

$$\begin{array}{rccc} h_\alpha\colon & V & \longrightarrow & V \\ & v & \longmapsto & \alpha \cdot v. \end{array}$$

Mit Hilfe der Vektorraum-Axiome (Definition III.1.1) überprüfen wir, dass h_α eine lineare Abbildung ist:

- Für $\lambda \in K$ und $v \in V$ haben wir

$$h_\alpha(\lambda \cdot v) = \alpha \cdot (\lambda \cdot v) = (\alpha \cdot \lambda) \cdot v = (\lambda \cdot \alpha) \cdot v = \lambda \cdot (\alpha \cdot v) = \lambda \cdot h_\alpha(v).$$

- Für $v, v' \in V$ gilt

$$h_\alpha(v + v') = \alpha \cdot (v + v') = \alpha \cdot v + \alpha \cdot v' = h_\alpha(v) + h_\alpha(v').$$

Eigenschaften III.5.4. *Es seien V, W zwei K-Vektorräume und $f\colon V \longrightarrow W$ eine lineare Abbildung.*

i) *Für jeden linearen Teilraum $U \subseteq V$ von V ist $f(U) \subseteq W$ ein linearer Teilraum von W. Insbesondere ist* $\mathrm{Bild}(f)$ *ein linearer Teilraum von W.*

ii) *Es sei $U \subseteq W$ ein linearer Teilraum von W. Dann ist $f^{-1}(U) \subseteq V$ ein linearer Teilraum von V.*

iii) *Für jede Teilmenge $M \subseteq V$ gilt*

$$f\big(\langle M \rangle\big) = \big\langle f(M) \big\rangle.$$

Beweis. Teil i) folgt aus Teil iii). (Man wähle etwa $M := U$.)

ii) Wir überprüfen die Axiome aus Definition III.1.4:

- Nach Beobachtung III.5.2 gilt $f(0) = 0 \in U$, so dass $0 \in f^{-1}(U)$.
- Es seien $\lambda \in K$ und $v \in f^{-1}(U)$. Dann hat man $f(\lambda \cdot v) = \lambda \cdot f(v) \in \lambda \cdot U \subseteq U$, so dass $\lambda \cdot v \in f^{-1}(U)$.
- Es seien $v, v' \in f^{-1}(U)$. Wir haben $f(v + v') = f(v) + f(v') \in U + U \subseteq U$, und es folgt $v + v' \in f^{-1}(U)$.

iii) Für $M = \varnothing$ haben wir $f(\varnothing) = \varnothing$. Nach Vereinbarung in Definition III.2.1, ii), gilt $\langle \varnothing \rangle = \{0\}$. Die Behauptung lautet in diesem Fall $f(\{0\}) = \{0\}$ und folgt aus Beobachtung III.5.2.

Für das Folgende dürfen wir $M \neq \varnothing$ voraussetzen. Wir zeigen zunächst $f(\langle M \rangle) \subseteq \langle f(M) \rangle$. Es sei $v \in \langle M \rangle$. Dann existieren $s \geq 1$, $\lambda_1, ..., \lambda_s \in K$ und $v_1, ..., v_s \in M$, so dass $v = \lambda_1 \cdot v_1 + \cdots + \lambda_s \cdot v_s$. Wir berechnen

$$f(v) = f(\lambda_1 \cdot v_1 + \cdots + \lambda_s \cdot v_s) = \lambda_1 \cdot f(v_1) + \cdots + \lambda_s \cdot f(v_s) \in \big\langle f(v_1), ..., f(v_s) \big\rangle.$$

Weiter gilt

$$\big\langle f(v_1), ..., f(v_s) \big\rangle \subseteq \big\langle f(w), w \in M \big\rangle = \big\langle f(M) \big\rangle.$$

Die Inklusion $f(\langle M \rangle) \supseteq \langle f(M) \rangle$ beweisen wir auf folgende Weise: Für einen Vektor $w \in \langle f(M) \rangle$ existieren $\lambda_1, ..., \lambda_t$, $w_1, ..., w_t \in f(M)$ mit $w = \lambda_1 \cdot w_1 + \cdots + \lambda_t \cdot w_t$ sowie $v_1, ..., v_t \in M$ mit $f(v_i) = w_i$, $i = 1, ..., t$. Für $v := \lambda_1 \cdot v_1 + \cdots + \lambda_t \cdot v_t \in \langle M \rangle$ finden wir

$$f(v) = f(\lambda_1 \cdot v_1 + \cdots + \lambda_t \cdot v_t) = \lambda_1 \cdot f(v_1) + \cdots + \lambda_t \cdot f(v_t) = \lambda_1 \cdot w_1 + \cdots + \lambda_t \cdot w_t = w,$$

i.e. $w \in f(\langle M \rangle)$. □

Definition III.5.5. Es seien V, W zwei K-Vektorräume und $f\colon V \longrightarrow W$ eine lineare Abbildung. Die Teilmenge

$$\mathrm{Ker}(f) := f^{-1}(0) = \big\{ v \in V \,|\, f(v) = 0 \big\}$$

heißt der *Kern* von f.

Bemerkung III.5.6. Nach Eigenschaft III.5.4, ii), ist $\mathrm{Ker}(f)$ ein linearer Teilraum von V.

Lemma III.5.7. *Es seien V, W zwei K-Vektorräume und $f\colon V \longrightarrow W$ eine lineare Abbildung. Dann ist f genau dann injektiv, wenn $\mathrm{Ker}(f) = \{0\}$ gilt.*

Beweis. Die Abbildung f sei injektiv. Nach Beobachtung III.5.2 gilt $0 \in \mathrm{Ker}(f)$ und daher $\mathrm{Ker}(f) = \{0\}$. Jetzt setzen wir $\mathrm{Ker}(f) = \{0\}$ voraus. Es seien $v, v' \in V$ mit $f(v) = f(v')$. Dann gilt

$$0 = f(v) - f(v') = f(v - v').$$

Wir schließen $v - v' = 0$, also $v = v'$. □

Satz III.5.8. *Es seien V, W zwei K-Vektorräume und $f\colon V \longrightarrow W$ eine **bijektive** lineare Abbildung. Dann ist auch die Umkehrabbildung*[6] $f^{-1}\colon W \longrightarrow V$ *linear.*

Beweis. **Schritt 1.** Es seien $w \in W$, $\lambda \in K$ und $v := f^{-1}(w)$. Da f linear ist, gilt $\lambda \cdot w = \lambda \cdot f(v) = f(\lambda \cdot v)$. Darauf wenden wir f^{-1} an und finden

$$f^{-1}(\lambda \cdot w) = f^{-1}\big(f(\lambda \cdot v)\big) = (f^{-1} \circ f)(\lambda \cdot v) = \mathrm{Id}_V(\lambda \cdot v) = \lambda \cdot v = \lambda \cdot f^{-1}(w).$$

Schritt 2. Es seien $w, w' \in W$, $v := f^{-1}(w)$ und $v' = f^{-1}(w')$. Es gilt $w + w' = f(v) + f(v') = f(v + v')$, denn f ist linear. Wir schließen

$$f^{-1}(w + w') = v + v' = f^{-1}(w) + f^{-1}(w').$$

Damit ist alles gezeigt. □

Bemerkung III.5.9. Eine lineare Abbildung $f\colon V \longrightarrow W$ ist genau dann bijektiv, wenn eine **lineare** Abbildung $g\colon W \longrightarrow V$ existiert, so dass $g \circ f = \mathrm{Id}_V$ und $f \circ g = \mathrm{Id}_W$.

Definition III.5.10. Es seien V, W zwei K-Vektorräume und $f\colon V \longrightarrow W$ eine bijektive lineare Abbildung. Man nennt f einen (*linearen*) *Isomorphismus*.

Matrizen und lineare Abbildungen. — Lineare Abbildungen zwischen K^m und K^n lassen sich elegant mit Matrizen beschreiben. Hier treffen wir auf Bekannte aus Kapitel I.

Definition III.5.11. Es sei

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

eine $(m \times n)$-Matrix. Wir setzen

$$\begin{array}{ccc} f_A\colon K^n & \longrightarrow & K^m \\ \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} & \longmapsto & \begin{pmatrix} t_1 \\ \vdots \\ t_m \end{pmatrix}, \quad t_i := \sum_{j=1}^{n} a_{ij} \cdot s_j, \quad j = 1, \ldots, m. \end{array}$$

Bemerkung III.5.12. Es ist leicht zu überprüfen, dass f_A eine lineare Abbildung ist (s. Aufgabe III.5.44).

[6] s. Definition II.1.26

Definition III.5.13. Es sei I eine Menge. Das *Kronecker*[7]*-Delta* auf der Menge I ist die Abbildung

$$\begin{aligned} \delta\colon I \times I &\longrightarrow \{0,1\} \\ (i,j) &\longmapsto \delta_{ij} := \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}. \end{aligned}$$

Bemerkung III.5.14. Wir können $\{0,1\}$ als Teilmenge des Körpers K auffassen und somit δ als Abbildung von I nach K.

Beispiele III.5.15. i) Es sei $n \geq 1$ eine natürliche Zahl. Das Kronecker-Delta auf der Menge $I = \{1, ..., n\}$ ist die sogenannte *Einheitsmatrix* $\mathbb{E}_n \in \mathrm{Mat}(n; K)$;

$$\mathbb{E}_n = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}.$$

Wir haben $f_{\mathbb{E}_n} = \mathrm{Id}_{K^n}$.

ii) Für

$$A = \begin{pmatrix} 1 & -1 & 0 & 2 \\ 2 & 1 & 3 & -1 \\ 1 & 1 & 0 & 1 \\ -1 & 3 & -1 & -1 \end{pmatrix} \in \mathrm{Mat}(4; \mathbb{R})$$

gilt

$$f_A \begin{pmatrix} 1 \\ 2 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -3 \\ 4 \\ 6 \end{pmatrix}.$$

Satz III.5.16. *Es gilt*

$$\mathrm{Dim}_K\big(\mathrm{Bild}(f_A)\big) = \text{S-Rg}(A).$$

Beweis. Es sei

$$v_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \in K^m$$

die j-te Spalte von A, $j = 1, ..., n$. Die Abbildungsvorschrift zeigt

$$f_A(e_j) = v_j, \quad j = 1, ..., n. \tag{III.6}$$

Mit Eigenschaft III.5.4, iii), finden wir

$$f_A(K^n) = f_A\big(\langle e_1, ..., e_n \rangle\big) = \big\langle f(e_1), ..., f(e_n) \big\rangle = \langle v_1, ..., v_n \rangle.$$

Die Dimension des rechten Vektorraums ist nach Definition III.4.12, ii), der Spaltenrang von A. □

[7] Leopold Kronecker (1823 - 1891), deutscher Mathematiker.

Dieser Zusammenhang motiviert, eine entsprechende Invariante für lineare Abbildungen einzuführen.

Definition III.5.17. Es seien V, W endlichdimensionale K-Vektorräume und $f\colon V \longrightarrow W$ eine lineare Abbildung. Die Zahl

$$\mathrm{Rg}(f) := \mathrm{Dim}_K\big(\mathrm{Bild}(f)\big)$$

heißt der *Rang* von f.

Bemerkung III.5.18. Es besteht ein enger Zusammenhang zwischen den oben eingeführten linearen Abbildungen und linearen Gleichungssystemen. Genauer gilt

$$\mathrm{Bild}(f_A) = \big\{\, b \in K^n \,|\, \mathbb{L}(A,b) \neq \varnothing \,\big\} \quad \text{und} \quad \mathrm{Ker}(f_A) = \mathbb{L}(A,0).$$

Beispiel III.5.19. Es sei $\alpha \in [0, 2\pi)$. Die Drehung $D_\alpha\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ um den Winkel α ist eine lineare Abbildung. Wie man Abbildung III.1 entnimmt, gilt

$$D_\alpha(e_1) = \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix} \quad \text{und} \quad D_\alpha(e_2) = \begin{pmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{pmatrix}.$$

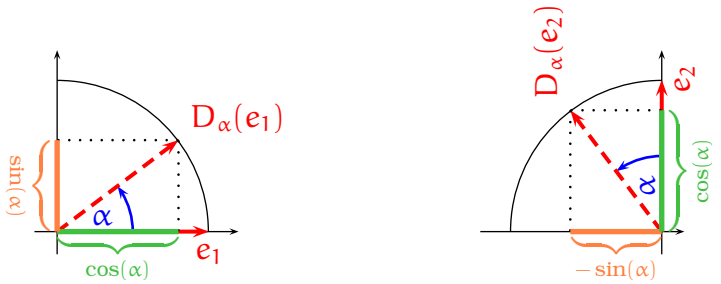


Abbildung III.1: Drehungen als lineare Abbildungen

Für die Matrix

$$A := \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

gilt

$$D_\alpha = f_A.$$

Lineare Abbildungen und Basen. — Mit Hilfe von Basen kann man lineare Abbildungen zwischen endlichdimensionalen K-Vektorräumen explizit und effizient beschreiben. Mit dieser Beschreibung werden wir hier beginnen und sie ausführlich in Kapitel IV entwickeln.

Definition III.5.20. Es seien V, W zwei K-Vektorräume. Man setzt

$$\mathrm{Hom}_K(V,W) := \left\{ f\colon V \longrightarrow W \,\middle|\, f \text{ ist linear} \right\}.$$

In Aufgabe III.5.47 werden Abbildungen

$$\begin{array}{rccc} +\colon \mathrm{Hom}_K(V,W) \times \mathrm{Hom}_K(V,W) & \longrightarrow & \mathrm{Hom}_K(V,W) \\ (f,g) & \longmapsto & f+g \end{array}$$

und

$$\begin{array}{rccc} \cdot\colon K \times \mathrm{Hom}_K(V,W) & \longrightarrow & \mathrm{Hom}_K(V,W) \\ (\lambda,f) & \longmapsto & \lambda \cdot f \end{array}$$

definiert, so dass $(\mathrm{Hom}_K(V,W),+,\cdot)$ ein K-Vektorraum ist.

Satz III.5.21. *Es seien V, W zwei K-Vektorräume und $B \subseteq V$ eine Basis für V. Dann ist*

$$\begin{array}{rccc} \Psi_B\colon \mathrm{Hom}_K(V,W) & \longrightarrow & \mathrm{Abb}(B,W) & \\ f & \longmapsto & f_{|B}\colon & \begin{array}{ccc} B & \longrightarrow & W \\ b & \longmapsto & f(b) \end{array} \end{array}$$

ein linearer Isomorphismus.

Die Injektivität besagt, dass eine lineare Abbildung durch ihre Werte auf den Basisvektoren eindeutig bestimmt ist. Die Surjektivität besagt, dass man für beliebige vorgeschriebene Werte der Basisvektoren eine lineare Abbildung von V nach W finden kann, die auf den Basisvektoren die vorgeschriebenen Werte annimmt.

Beweis von Satz III.5.21. Die Linearität der angegebenen Abbildung folgt direkt aus der Definition der Vektorraum-Strukturen auf $\mathrm{Hom}_K(V,W)$ und $\mathrm{Abb}(B,W)$ (s. Aufgabe III.1.8 und III.5.47).

Zu jedem Vektor $v \in V$ existieren eindeutig bestimmte Zahlen λ_b, $b \in B$, so dass $v = \sum_{b\in B} \lambda_b \cdot b$. Für eine lineare Abbildung $f\colon V \longrightarrow W$ gilt

$$f(v) = \sum_{b\in B} \lambda_b \cdot f(b).$$

Diese Formel zeigt, wie man $f(v)$ aus den Werten $f(b)$, $b \in B$, von f auf den Basisvektoren berechnet. Es folgt die Injektivität.

Es sei $\widetilde{f}\colon B \longrightarrow W$ eine mengentheoretische Abbildung. Wir definieren eine Abbildung $f\colon V \longrightarrow W$ auf folgende Weise: Zu $v \in V$ seien $\lambda_b \in K$, $b \in B$, die eindeutig bestimmten Zahlen, für die $v = \sum_{b\in B} \lambda_b \cdot b$ gilt. Damit setzen wir

$$f(v) = \sum_{b\in B} \lambda_b \cdot \widetilde{f}(b).$$

Man vergewissert sich leicht, dass f linear ist und $f_{|B} = \widetilde{f}$ erfüllt (s. Aufgabe III.5.48). □

Folgerung III.5.22. *Es seien $m, n \in \mathbb{N}$ natürliche Zahlen und $f\colon K^n \longrightarrow K^m$ eine lineare Abbildung. Dann gibt es genau eine $(m \times n)$-Matrix $A \in \mathrm{Mat}(m, n; K)$, so dass*

$$f = f_A.$$

Beweis. Es sei $\{ e_j \mid j = 1, ..., n \}$ die Standardbasis von K^n. Wir schreiben

$$f(e_j) = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \in K^m, \quad j = 1, ..., n,$$

und definieren

$$A := \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in \mathrm{Mat}(m, n; K).$$

Die Abbildungen f und f_A sind linear, und nach Konstruktion gilt $f(e_j) = f_A(e_j)$, $j = 1, ..., n$. Die Injektivitätsaussage in Satz III.5.21 impliziert $f = f_A$. Die Eindeutigkeit ergibt sich, da $f_A(e_j)$ gerade die j-te Spalte der Matrix A ist, $j = 1, ..., n$. □

Der Endomorphismenring und die allgemeine lineare Gruppe eines Vektorraums. — In Abschnitt II.2 haben wir gesehen, dass die bijektiven Selbstabbildungen einer Menge bzgl. Verknüpfung eine Gruppe bilden. Dasselbe gilt für die bijektiven **linearen** Selbstabbildungen eines Vektorraums. Bevor wir darauf eingehen, werden wir erklären, dass die Menge **aller** linearen Selbstabbildungen eines Vektorraums ebenfalls eine interessante algebraische Struktur trägt.

Satz III.5.23. *Es seien U, V, W K-Vektorräume und $f\colon U \longrightarrow V$ und $g\colon V \longrightarrow W$ Abbildungen. Wenn f und g linear sind, dann ist auch die Verknüpfung $g \circ f\colon U \longrightarrow W$ linear.*

Beweis. Es seien $\lambda \in K$ und $u, u' \in U$.

- Nach Definition gilt $(g \circ f)(u + u') = g\big(f(u + u')\big)$. Da f linear ist, haben wir $g\big(f(u + u')\big) = g\big(f(u) + f(u')\big)$. Die Linearität von g ergibt $g\big(f(u) + f(u')\big) = g\big(f(u)\big) + g\big(f(u')\big) = (g \circ f)(u) + (g \circ f)(u')$.
- Analog erhalten wir $(g \circ f)(\lambda \cdot u) = g\big(f(\lambda \cdot u)\big) = g\big(\lambda \cdot f(u)\big) = \lambda \cdot g\big(f(u)\big) = \lambda \cdot (g \circ f)(u)$.

Somit ist $g \circ f$ eine lineare Abbildung. □

Satz III.5.24. *Es seien* U, V, W, X *Vektorräume über dem Körper* K, $f, f' \in \mathrm{Hom}_K(U, V)$, $g, g' \in \mathrm{Hom}_K(V, W)$, $h \in \mathrm{Hom}_K(W, X)$ *und* $\lambda \in K$. *Dann hat man folgende Identitäten:*

i) $g \circ (f + f') = g \circ f + g \circ f'$, $(g + g') \circ f = g \circ f + g' \circ f$.
ii) $g \circ (\lambda \cdot f) = (\lambda \cdot g) \circ f = \lambda \cdot (g \circ f)$.
iii) $(h \circ g) \circ f = h \circ (g \circ f)$.

Beweis. i) Für $u \in U$ berechnen wir:

$$\begin{aligned}
\big(g \circ (f + f')\big)(u) &= g\big((f + f')(u)\big) \\
&= g\big(f(u) + f'(u)\big) \\
&= g\big(f(u)\big) + g\big(f'(u)\big) \\
&= (g \circ f)(u) + (g \circ f')(u) \\
&= (g \circ f + g \circ f')(u), \\
\big((g + g') \circ f\big)(u) &= (g + g')\big(f(u)\big) \\
&= g\big(f(u)\big) + g'\big(f(u)\big) \\
&= (g \circ f)(u) + (g' \circ f)(u) \\
&= (g \circ f + g' \circ f)(u).
\end{aligned}$$

In den Zwischenschritten haben wir die Definition der Summe von Abbildungen und die Linearität der beteiligten Abbildungen verwendet. Die Behauptung ergibt sich, weil zwei Abbildungen von U nach W genau dann gleich sind, wenn sie auf allen Vektoren $u \in U$ dieselben Werte annehmen (Bemerkung II.1.17, i).

ii) Es sei $u \in U$. Wir haben

$$\big(\lambda \cdot (g \circ f)\big)(u) = \lambda \cdot \big((g \circ f)(u)\big) = \lambda \cdot g\big(f(u)\big) = (\lambda \cdot g)\big(f(u)\big) = \big((\lambda \cdot g) \circ f\big)(u).$$

Damit ergibt sich die erste Gleichung. Aus der Linearität von g folgt weiter

$$\big(g \circ (\lambda \cdot f)\big)(u) = g\big((\lambda \cdot f)(u)\big) = g\Big(\lambda \cdot \big(f(u)\big)\Big) = \lambda \cdot g\big(f(u)\big) = \lambda \cdot \big((g \circ f)(u)\big),$$

und man erhält die zweite Gleichung. Der dritte Teil wurde bereits in Satz II.1.25, ii), bewiesen. □

Definition III.5.25. Es sei V ein K-Vektorraum. Eine lineare Abbildung $f\colon V \longrightarrow V$ wird auch *Endomorphismus von* V genannt.
Schreibweise. Entsprechend setzt man $\mathrm{End}_K(V) := \mathrm{Hom}_K(V,V)$.

In Aufgabe III.1.8 und III.5.47 haben wir die Operation

$$\begin{aligned} +\colon \mathrm{End}_K(V) \times \mathrm{End}_K(V) &\longrightarrow \mathrm{End}_K(V) \\ (f,g) &\longmapsto f+g \end{aligned}$$

erklärt. Nach Satz III.5.23 ist auch die Operation

$$\begin{aligned} \circ\colon \mathrm{End}_K(V) \times \mathrm{End}_K(V) &\longrightarrow \mathrm{End}_K(V) \\ (f,g) &\longmapsto f\circ g \end{aligned}$$

definiert.

Folgerung III.5.26. *Es sei V ein Vektorraum, so dass $\mathrm{Dim}_K(V) \geq 1$.*[8] *Dann ist $(\mathrm{End}_K(V), +, \circ)$ ein Ring. Das neutrale Element der Addition ist die Nullabbildung $0\colon V \longrightarrow V$, $v \longmapsto 0$, und dasjenige der Multiplikation die identische Abbildung Id_V.*

Beweis. **Schritt 1.** Das Tripel $(\mathrm{End}_K(V), +)$ ist eine abelsche Gruppe mit Neutralelement 0. Dies wurde bereits in Aufgabe III.5.47 nachgewiesen.

Schritt 2. Wir wissen ebenfalls (Beispiel II.1.24, i), Satz II.1.25, ii), dass für $f,g,h \in \mathrm{End}_K(V)$ gilt:

$$\begin{aligned} f\circ \mathrm{Id}_V &= f = \mathrm{Id}_V \circ f, \\ (f\circ g)\circ h &= f\circ(g\circ h). \end{aligned}$$

Da $\mathrm{Dim}_K(V) \geq 1$, also $V \neq \{0\}$, angenommen wurde, gilt $0 \neq \mathrm{Id}_V$.

Schritt 3. Die Distributivgesetze sind in Satz III.5.24, i), enthalten. □

Bemerkung III.5.27. Für den Nullvektorraum $V = \{0\}$ gilt $0 = \mathrm{Id}_V$, und es ergibt sich der Nullring $\mathrm{End}_K(V) = \{0\}$ (vgl. Bemerkung II.3.2, ii), den wir in Definition II.3.1, i), ausgeschlossen hatten.

Beispiel III.5.28. Es gelte $\mathrm{Dim}_K(V) = 1$. Dann ist (s. Beispiel III.5.3, iii)

$$\begin{aligned} \varphi\colon K &\longrightarrow \mathrm{End}_K(V) \\ \alpha &\longmapsto h_\alpha \end{aligned}$$

ein Isomorphismus (vgl. Aufgabe II.4.11, c), d.h. eine bijektive Abbildung, so dass

$$\begin{aligned} \varphi(\alpha+\alpha') &= \varphi(\alpha)+\varphi(\alpha'), \\ \varphi(\alpha\cdot\alpha') &= \varphi(\alpha)\circ\varphi(\alpha'), \quad \alpha,\alpha' \in K, \\ \varphi(1) &= \mathrm{Id}_V. \end{aligned}$$

[8] Der Vektorraum V darf auch unendlichdimensional sein.

Satz III.5.29. *Es sei V ein Vektorraum über K, Dann ist*

$$\begin{aligned} GL(V) &:= \left\{ f \in \mathcal{S}(V) \,|\, f \textit{ ist linear} \right\} \\ &= \left\{ f \in \mathrm{End}_K(V) \,|\, f \textit{ ist bijektiv} \right\} \\ &= \left\{ f \in \mathrm{End}_K(V) \,|\, \exists g \in \mathrm{End}_K(V) : f \circ g = \mathrm{Id}_V = g \circ f \right\} \end{aligned}$$

eine Untergruppe von $\mathcal{S}(V)$.[9]

Definition III.5.30. Die Gruppe $GL(V)$ aus dem Satz ist die *allgemeine lineare Gruppe* des Vektorraums V.

Beweis von Satz III.5.29. Wir überprüfen die Axiome aus Definition II.2.7. Das neutrale Element der Gruppe $\mathcal{S}(V)$ ist die Identität. Es ist eine lineare Abbildung, gehört also $GL(V)$ an. Für eine bijektive lineare Abbildung $f \in GL(V)$, ist die Umkehrabbildung f^{-1} nach Satz III.5.8 ebenfalls linear, i.e. $f^{-1} \in GL(V)$. Schließlich seien $f, g \in GL(V)$. Die Verknüpfung $f \circ g$ ist nach Satz III.5.23 ebenfalls linear. Da sie auch bijektiv ist, gilt $f \circ g \in GL(V)$. □

Definition III.5.31. Es sei V ein endlichdimensionaler K-Vektorraum der Dimension n. Eine *durchnummerierte Basis* von V ist eine Basis B von V zusammen mit einer bijektiven Abbildung $\beta : \{1, ..., n\} \longrightarrow B$. Mit $b_i := \beta(i)$, $i = 1, ..., n$, werden wir eine durchnummerierte Basis einfach in der Form $B = \{b_1, ..., b_n\}$ schreiben.

Satz III.5.32. i) *Es seien V ein endlichdimensionaler K-Vektorraum und $B = \{b_1, ..., b_n\}$ eine durchnummerierte Basis von V. Dann gibt es genau einen linearen Isomorphismus*

$$\Phi_B : V \longrightarrow K^n$$

mit

$$\Phi_B(b_i) = e_i, \quad i = 1, ..., n.$$

ii) *Seien umgekehrt V ein K-Vektorraum und $\Phi : V \longrightarrow K^n$ ein linearer Isomorphismus. Dann ist*

$$B_\Phi := \left\{ \Phi^{-1}(e_1), ..., \Phi^{-1}(e_n) \right\}$$

eine durchnummerierte Basis von V.

iii) *In den Bezeichnungen von* i) *und* ii) *gilt:*

$$B_{\Phi_B} = B \quad \textit{und} \quad \Phi_{B_\Phi} = \Phi.$$

[9] Die Gruppe $\mathcal{S}(V)$ wurde in Bemerkung II.2.1 eingeführt.

Beweis. i) Nach Satz III.5.21 gibt es genau eine lineare Abbildung $\Phi_B\colon V \longrightarrow K^n$ mit $\Phi_B(b_i) = e_i$, $i = 1, ..., n$. Es ist zu überprüfen, dass Φ_B bijektiv ist. Für die Surjektivität verwenden wir Eigenschaft III.5.4, iii). Es gilt

$$\text{Bild}(\Phi_B) = \Phi_B(V) = \Phi_B\big(\langle B\rangle\big) = \big\langle \Phi_B(B)\big\rangle = \big\langle\, \Phi_B(b_1), ..., \Phi_B(b_n)\,\big\rangle = \langle\, e_1, ..., e_n\,\rangle = K^n.$$

Zum Nachweis der Injektivität verwenden wir das Kriterium aus Lemma III.5.7. Es sei $v = \sum_{i=1}^{n} \lambda_i \cdot b_i$ ein Vektor aus dem Kern von Φ_B. Dann gilt

$$0 = \Phi_B\left(\sum_{i=1}^{n} \lambda_i \cdot b_i\right) = \sum_{i=1}^{n} \lambda_i \cdot \Phi_B(b_i) = \sum_{i=1}^{n} \lambda_i \cdot e_i.$$

Es folgt $\lambda_1 = \cdots = \lambda_n = 0$ und damit $v = 0$.

ii) Wir zeigen zunächst, dass B_Φ ein Erzeugendensystem ist. Dazu verwenden wir den linearen Isomorphismus $\Phi^{-1}\colon K^n \longrightarrow V$. Wie zuvor sehen wir

$$\langle B_\Phi\rangle = \Big\langle\, \Phi^{-1}\big(\{\, e_1, ..., e_n\,\}\big)\,\Big\rangle = \Phi^{-1}\big(\langle\, e_1, ..., e_n\,\rangle\big) = \Phi^{-1}(K^n) = V.$$

Jetzt widmen wir uns der linearen Unabhängigkeit von B_Φ. Es seien $\lambda_1, ..., \lambda_n \in K$, so dass

$$0 = \sum_{i=1}^{n} \lambda_i \cdot \Phi^{-1}(e_i).$$

Auf diese Identität wenden wir Φ an und erhalten

$$0 = \Phi(0) = \Phi\left(\sum_{i=1}^{n} \lambda_i \cdot \Phi^{-1}(e_i)\right) = \sum_{i=1}^{n} \lambda_i \cdot \Phi\big(\Phi^{-1}(e_i)\big) = \sum_{i=1}^{n} \lambda_i \cdot e_i.$$

Es folgt wiederum $\lambda_1 = \cdots = \lambda_n = 0$. Die Behauptungen in Teil iii) ergeben sich unmittelbar aus den Definitionen. □

Definition III.5.33. Es seien V, W zwei K-Vektorräume. Wir sagen, V und W sind *isomorph*, wenn es einen linearen Isomorphismus $\Phi\colon V \longrightarrow W$ gibt.
Schreibweise. $V \cong W$.

Folgerung III.5.34. *Es seien V und W zwei endlichdimensionale K-Vektorräume. Dann sind V und W genau dann isomorph, wenn*

$$\text{Dim}_K(V) = \text{Dim}_K(W).$$

Somit ist ein n-dimensionaler K-Vektorraum zum Vektorraum K^n isomorph. Hier schließt sich ein Kreis: Wir haben in Definition I.2.1 mit dem Vektorraum K^n begonnen, um Lösungsmengen linearer Gleichungssysteme definieren zu können. Ausgehend von den Eigenschaften dieses Vektorraums (Aufgabe I.2.10) haben wir das abstrakte Konzept des Vektorraums eingeführt. Nun

stellt sich heraus, dass jeder „abstrakte“ **endlichdimensionale** Vektorraum im Wesentlichen ein K^n ist. Trotzdem hat sich der große Aufwand gelohnt. Ein Vektorraum besitzt nur selten eine ausgezeichnete Basis. (Man betrachte etwa die in Abschnitt III.6 konstruierten Quotientenvektorräume.) Daher empfiehlt es sich, die Theorie soweit wie möglich ohne die Verwendung von Basen zu entwickeln, d.h. ohne die Vektorräume mit K^n zu identifizieren. Auf der anderen Seite käme man kaum auf die Idee, für K^n eine andere Basis als die Standardbasis zu wählen. Dennoch wird sich in Kapitel IV und V zeigen, dass die geschickte Wahl einer Basis sich oftmals auszeichnet. Daher ist es wichtig, über einen Formalismus zu verfügen, mit dem man diese Wahl angemessen beschreiben kann.

Lineare Komplemente. — Eine Variante von Satz III.4.10, ii), führt zu einem weiteren Verfahren, mit dem man einen Vektorraum in kleinere Bestandteile zerlegen kann.

Definition III.5.35. Es seien V ein K-Vektorraum und $U \subseteq V$ ein linearer Teilraum. Ein linearer Teilraum $U' \subseteq V$ ist ein *lineares Komplement* zu U in V, wenn

$$V = U \oplus U'$$

gilt.

Nach Definition III.2.1, ii), bedeutet die Bedingung $V = U \oplus U'$, dass

$$V = U + U' = \{ u + u' \mid u \in U,\ u' \in U' \}$$

und

$$U \cap U' = \{0\},$$

also dass sich jeder Vektor aus V in eindeutiger Weise als Summe eines Vektors aus U und eines Vektors aus U' darstellen lässt.

Lemma III.5.36. *Es seien V ein endlichdimensionaler K-Vektorraum, $U \subseteq V$ ein linearer Teilraum und $U' \subseteq V$ ein lineares Komplement zu U. Dann gilt:*

$$\mathrm{Dim}_K(V) = \mathrm{Dim}_K(U) + \mathrm{Dim}_K(U').$$

Beweis. Die Vektorräume U und U' sind ebenfalls endlichdimensional (Folgerung III.4.11). Wir wählen eine Basis B für U und eine Basis B' für U'. Wir behaupten, dass $B \cup B'$ eine Basis für V ist. Zunächst zeigen wir, dass die Menge $B \cup B'$ den Vektorraum V erzeugt. Es sei $v \in V$. Es existieren Vektoren $u \in U$

und $u' \in U'$ mit $v = u + u'$. Weiter gibt es $\lambda_b \in K$, $b \in B$, und $\mu_{b'} \in K$, $b' \in B'$, mit $u = \sum_{b \in B} \lambda_b \cdot b$ und $u' = \sum_{b' \in B'} \mu_{b'} \cdot b'$. Es folgt

$$v = \sum_{b \in B} \lambda_b \cdot b + \sum_{b' \in B'} \mu_{b'} \cdot b' \in \langle B \cup B' \rangle.$$

Weiter seien $\lambda_b \in K$, $b \in B$, und $\mu_{b'} \in K$, $b' \in B'$, so dass

$$\sum_{b \in B} \lambda_b \cdot b + \sum_{b' \in B'} \mu_{b'} \cdot b' = 0.$$

Es folgt, dass

$$\underbrace{\sum_{b \in B} \lambda_b \cdot b}_{=:u \in U} = \underbrace{- \sum_{b' \in B'} \mu_{b'} \cdot b'}_{=:u' \in U'}.$$

Somit gilt $u, u' \in U \cap U'$, d.h. $u = u' = 0$. Da B eine Basis von U ist, haben wir $\lambda_b = 0$, $b \in B$. Ebenso folgt, dass $\mu_{b'} = 0$, $b' \in B'$. Wir haben nachgewiesen, dass die Menge $B \cup B'$ linear unabängig ist. □

Bemerkungen und Beispiele III.5.37. i) Es sei V ein K-Vektorraum. Für $U = \{0\}$ ist $U' = V$ das einzige lineare Komplement, und für $U = V$ ist $U' = \{0\}$ das einzige.

ii) Wir betrachten den unendlichdimensionalen K-Vektorraum $V := \mathrm{Abb}(\mathbb{N}, K)$ (s. Aufgabe II.3.7 und Beispiel III.1.2, ii) und definieren

$$\begin{aligned} U &:= \left\{ f \in \mathrm{Abb}(\mathbb{N}, K) \,\middle|\, \forall n \in \mathbb{N} \text{ ungerade} : f(n) = 0 \right\}, \\ U' &:= \left\{ f \in \mathrm{Abb}(\mathbb{N}, K) \,\middle|\, \forall n \in \mathbb{N} \text{ gerade} : f(n) = 0 \right\}. \end{aligned}$$

Dann gilt

$$V = U \oplus U'.$$

Man beachte, dass sowohl U als auch U' unendlichdimensional ist.

iii) Lineare Komplemente sind nicht eindeutig bestimmt. Es sei z.B. $U := \left\langle \binom{1}{0} \right\rangle \subseteq K^2$. Dann ist $U' := \left\langle \binom{a}{b} \right\rangle$ für jeden Vektor $\binom{a}{b}$ mit $b \neq 0$ ein lineares Komplement.

Satz III.5.38. *Es seien V ein **endlichdimensionaler** K-Vektorraum und $U \subseteq V$ ein linearer Teilraum. Dann existiert ein lineares Komplement $U' \subseteq V$ zu U.*

Beweis. Es sei F eine Basis für U. Dann können wir nach Satz III.4.10, ii), eine Teilmenge $F' \subseteq V$ finden, so dass $F \cup F'$ eine Basis von V ist. Man überprüft mühelos, dass $U' := \langle F' \rangle$ ein lineares Komplement zu U ist. □

Als kleine Anwendung notieren wir:

Lemma III.5.39. *Es seien V, W zwei K-Vektorräume, $f\colon V \longrightarrow W$ eine lineare Abbildung, und $U' \subseteq V$ ein lineares Komplement zu $\mathrm{Ker}(f)$. Dann ist*

$$\begin{array}{rcl} f_{|U'}\colon U' & \longrightarrow & \mathrm{Bild}(f) \\ u' & \longmapsto & f(u') \end{array}$$

ein linearer Isomorphismus.

Beweis. Die Abbildung $f_{|U'}$ ist offenbar linear. Sei $w \in \mathrm{Bild}(f)$. Es gibt also einen Vektor $v \in V$ mit $f(v) = w$. Ferner existieren Elemente $u \in \mathrm{Ker}(f)$ und $u' \in U'$ mit $v = u + u'$. Wir finden

$$w = f(v) = f(u + u') = f(u) + f(u') = f(u') = f_{|U'}(u').$$

Dies zeigt, dass $f_{|U'}$ surjektiv ist. Jetzt sei $u' \in \mathrm{Ker}(f_{|U'})$. Also $f(u') = f_{|U'}(u') = 0$, so dass $u' \in \mathrm{Ker}(f) \cap U' = \{0\}$. Nach Lemma III.5.7 ist $f_{|U'}$ injektiv. □

Die Dimensionsformel für lineare Abbildungen. — Die Dimension des Bilds und des Kerns einer linearen Abbildung sind über die folgende Formel miteinander verbunden. Damit kann man z.B. die in Problem III.4.17 gestellte Frage beantworten.

Satz III.5.40 (Dimensionsformel für lineare Abbildungen)**.** *Es seien V, W Vektorräume über K und $f\colon V \longrightarrow W$ eine lineare Abbildung. Der Vektorraum V sei endlichdimensional. Dann gilt*

$$\mathrm{Dim}_K\big(\mathrm{Ker}(f)\big) + \mathrm{Dim}_K\big(\mathrm{Bild}(f)\big) = \mathrm{Dim}_K(V).$$

Beweis. Wir wählen gemäß Satz III.5.38 ein lineares Komplement U' zum Unterraum $\mathrm{Ker}(f)$ von V. Dann gilt:

- $\mathrm{Dim}_K\big(\mathrm{Ker}(f)\big) + \mathrm{Dim}_K(U') = \mathrm{Dim}_K(V)$ (Lemma III.5.36);
- Der Unterraum U' ist isomorph zu $\mathrm{Bild}(f)$ (Lemma III.5.39), so dass insbesondere $\mathrm{Dim}_K(U') = \mathrm{Dim}_K\big(\mathrm{Bild}(f)\big)$.

Diese beiden Tatsachen ergeben zusammen die Behauptung des Satzes. □

Folgerung III.5.41. *Gegeben seien endlichdimensionale K-Vektorräume V, W und eine lineare Abbildung $f\colon V \longrightarrow W$.*

i) *Falls f injektiv ist, dann ergibt sich* $\mathrm{Dim}_K(V) \leq \mathrm{Dim}_K(W)$.

ii) *Falls f surjektiv ist, dann folgt* $\mathrm{Dim}_K(V) \geq \mathrm{Dim}_K(W)$.

iii) *Gilt* $\mathrm{Dim}_K(V) = \mathrm{Dim}_K(W)$, *so sind folgende Aussagen äquivalent:*

a) *Die Abbildung* f *is ein linearer Isomorphismus.*
b) *Die Abbildung* f *ist injektiv.*
c) *Die Abbildung* f *ist surjektiv.*
iv) *Es seien* $m, n \geq 1$ *natürliche Zahlen und* A *eine* $(m \times n)$*-Matrix. Dann gilt*

$$\text{S-Rg}(A) = \text{Z-Rg}(A).$$

Beweis. i) Hier gilt $\text{Ker}(f) = \{0\}$ und $V \cong \text{Bild}(f)$ (Lemma III.5.39). Da $\text{Bild}(f) \subseteq W$ ein linearer Teilraum ist (Eigenschaft III.5.4, i), haben wir $\text{Dim}_K(\text{Bild}(f)) \leq \text{Dim}_K(W)$ nach Folgerung III.4.11.

ii) Es gilt $W = \text{Bild}(f)$, so dass aus der Dimensionsformel

$$\text{Dim}_K(W) = \text{Dim}_K(V) - \text{Dim}_K\big(\text{Ker}(f)\big) \leq \text{Dim}_K(V)$$

folgt.

iii) Ein Isomorphismus ist nach Definition bijektiv, d.h. sowohl injektiv als auch surjektiv. Somit sind die Implikationen „a)$\Longrightarrow$b)“ und „a)$\Longrightarrow$c)“ offensichtlich. Für „b)$\Longrightarrow$a)“ beobachten wir, dass die Voraussetzung und die Dimensionsformel wegen $\text{Ker}(f) = \{0\}$

$$\text{Dim}_K\big(\text{Bild}(f)\big) = \text{Dim}_K(V) = \text{Dim}_K(W)$$

ergeben. Da $\text{Bild}(f)$ ein linearer Teilraum von W ist, impliziert die Gleichung $\text{Dim}_K(\text{Bild}(f)) = \text{Dim}_K(W)$, dass $\text{Bild}(f) = W$. Es bleibt „c)$\Longrightarrow$a)“ nachzuweisen. Hier besagen die Voraussetzungen

$$\text{Dim}_K\big(\text{Bild}(f)\big) = \text{Dim}_K(W) = \text{Dim}_K(V).$$

Die Dimensionsformel zeigt $\text{Dim}_K(\text{Ker}(f)) = 0$, d.h. $\text{Ker}(f) = \{0\}$. Nach Lemma III.5.7 ist f injektiv.

iv) Es gilt $\text{S-Rg}(A) = \text{Dim}_K(\text{Bild}(f_A))$. Nach der Dimensionsformel für f_A hat man

$$\text{Dim}_K\big(\text{Bild}(f_A)\big) = \text{Dim}_K(K^n) - \text{Dim}_K\big(\text{Ker}(f_A)\big).$$

Wegen $\text{Ker}(f_A) = \mathbb{L}(A, 0)$ (Bemerkung III.5.18) lautet die Behauptung also

$$\text{Z-Rg}(A) = n - \text{Dim}_K\big(\mathbb{L}(A, 0)\big)$$

und ist nach Folgerung III.4.15 wahr. □

Definition III.5.42. Es seien $m, n \geq 1$ natürliche Zahlen und A eine $(m \times n)$-Matrix. Die Zahl

$$\text{Rg}(A) := \text{S-Rg}(A)$$

ist der *Rang* der Matrix A.

Aufgabe III.5.43.
a) Es seien M eine Menge, V ein K-Vektorraum und $x \in M$. Zeigen Sie, dass die so genannte *Auswertungsabbildung*

$$\begin{array}{rcl} \mathrm{ev}_x \colon \mathrm{Abb}(M,V) & \longrightarrow & V \\ f & \longmapsto & f(x) \end{array}$$

linear ist.
b) Es seien M und N Mengen, und V und W seien K-Vektorräume. Ferner seien eine Abbildung $\varphi \colon M \longrightarrow N$ und eine lineare Abbildung $\psi \colon V \longrightarrow W$ gegeben. Zeigen Sie, dass es sich bei den Abbildungen

$$\begin{array}{rcl} \varphi^* \colon \mathrm{Abb}(N,V) & \longrightarrow & \mathrm{Abb}(M,V) \\ f & \longmapsto & f \circ \varphi \end{array}$$

und

$$\begin{array}{rcl} \psi_* \colon \mathrm{Abb}(M,V) & \longrightarrow & \mathrm{Abb}(M,W) \\ f & \longmapsto & \psi \circ f \end{array}$$

um lineare Abbildungen handelt. Benutzen Sie eine Abbildung vom Typ φ^*, um einen linearen Isomorphismus zwischen den Vektorräumen $\mathrm{Mat}(m,n;K)$ und $K^{m\cdot n}$ herzustellen.

Aufgabe III.5.44.
Es seien $A \in \mathrm{Mat}(m,n;K)$ eine $(m \times n)$-Matrix und $f_A \colon K^n \longrightarrow K^m$ die in Definition III.5.11 definierte zugehörige Abbildung. Rechnen Sie nach, dass f_A eine lineare Abbildung ist.

Aufgabe III.5.45.
Wir betrachten eine lineare Abbildung $f_A \colon K^3 \longrightarrow K^3$, $K := \mathbb{F}_7$. Berechnen Sie $\mathrm{Ker}(f_A)$, $f^{-1}(\{v\})$ und $f^{-1}(\langle v \rangle)$ zu den Daten

$$A := \begin{pmatrix} \overline{1} & \overline{1} & \overline{1} \\ \overline{-2} & \overline{-1} & \overline{-1} \\ \overline{2} & \overline{-1} & \overline{6} \end{pmatrix}, \quad v := \begin{pmatrix} \overline{1} \\ \overline{1} \\ \overline{7} \end{pmatrix}.$$

Aufgabe III.5.46.
Gegeben seien folgende Vektoren im $\mathbb{R}^4$:

$$v_1 := \begin{pmatrix} 1 \\ 5 \\ -1 \\ -1 \end{pmatrix}, v_2 := \begin{pmatrix} -2 \\ -7 \\ 3 \\ 3 \end{pmatrix}, v_3 := \begin{pmatrix} 3 \\ 11 \\ -6 \\ -3 \end{pmatrix}, v_4 := \begin{pmatrix} -5 \\ -14 \\ 12 \\ 6 \end{pmatrix}.$$

Gibt es eine lineare Abbildung $f\colon \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ mit

$$f(v_1) = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix},\ f(v_2) = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix},\ f(v_3) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \text{ und } f(v_4) = \begin{pmatrix} 4 \\ 8 \\ 0 \end{pmatrix}?$$

Aufgabe III.5.47.
Zeigen Sie, dass $\mathrm{Hom}_K(V,W)$ ein linearer Teilraum von $\mathrm{Abb}(V,W)$ ist (s. Aufgabe III.1.8).

Aufgabe III.5.48.
Es seien V, W Vektorräume über dem Körper K, $B \subseteq V$ eine Basis und $\widetilde{f}\colon B \longrightarrow W$ eine mengentheoretische Abbildung. Für[10] $v = \sum_{b\in B} \lambda_b \cdot b$ definieren wir

$$f(v) := \sum_{b\in B} \lambda_b \cdot \widetilde{f}(b).$$

Erläutern Sie, warum dieser Ausdruck wohldefiniert ist, und weisen Sie nach. dass

$$\begin{array}{rcl} f\colon V & \longrightarrow & W \\ v & \longmapsto & f(v) \end{array}$$

eine lineare Abbildung ist, die $f_{|B} = \widetilde{f}$ erfüllt.

Aufgabe III.5.49.
Es sei V ein K-Vektorraum, und U, U' seien Unterräume von V. Gemäß Aufgabe III.1.9 können wir die direkte Summe $U \oplus U'$ der Vektorräume U und U' bilden und, wie in Definition III.2.1, ii), beschrieben, die Summe $U + U'$ in V. Geben Sie eine „natürliche“ lineare Abbildung $f\colon U \oplus U' \longrightarrow V$ mit $\mathrm{Bild}(f) = U + U'$ an. Zeigen Sie, dass f genau dann ein Isomorphismus ist, wenn $V = U \oplus U'$ (im Sinne von Definition III.2.1, ii), gilt.[11] Bestimmen Sie allgemeiner den Kern von f und leiten Sie die *Dimensionsformel für Unterräume*

$$\mathrm{Dim}_K(U + U') = \mathrm{Dim}_K(U) + \mathrm{Dim}_K(U') - \mathrm{Dim}_K(U \cap U')$$

aus der Dimensionsformel für lineare Abbildungen ab.

Aufgabe III.5.50.
Gegeben seien die Unterräume

$$U := \left\langle \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} \right\rangle, \quad U' := \left\langle \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \right\rangle$$

[10] Wie üblich müssen fast alle Koeffizienten $\lambda_b \in K$, $b \in B$, null sein.
[11] Die Doppelbezeichnung für zwei an sich verschiedene Objekte ist also gerechtfertigt.

von $\mathbb{R}^3$. Berechnen Sie eine Basis für $U \cap U'$, und bestimmen Sie die Summe $U + U' \subseteq \mathbb{R}^3$ als Unterraum. Ermitteln Sie schließlich ein lineares Komplement zu $U \cap U'$ in $U + U'$.

III.6 Quotientenvektorräume

Definition III.6.1. Es seien V ein K-Vektorraum und $U \subseteq V$ ein linearer Teilraum. Für Vektoren $v, v' \in V$ schreiben wir

$$v \sim_U v' \quad :\Longleftrightarrow \quad v' - v \in U.$$

Falls $v \sim_U v'$, dann sind v, v' *kongruent modulo* U.

Bemerkungen III.6.2. i) Nach Lemma II.4.5 ist „$\sim_U$“ eine Äquivalenzrelation. Die Äquivalenzklassen sind von der Form

$$[v]_U = v + U = \{ v + u \,|\, u \in U \}, \quad v \in V.$$

ii) Es gilt genau dann $v \sim_U v'$, wenn es einen Vektor $u \in U$ mit $v' = v + u$ gibt, $v, v' \in V$.

Satz III.6.3. *Es seien V ein K-Vektorraum, $U \subseteq V$ ein linearer Teilraum und $W := V/U$ die Menge der Äquivalenzklassen bzgl. der Äquivalenzrelation „$\sim_U$“.*

i) *Die Abbildungen*

$$\begin{aligned} +\colon W \times W &\longrightarrow W \\ \big([v]_U, [v']_U\big) &\longmapsto [v + v']_U, \\ \cdot\colon K \times W &\longrightarrow W \\ \big(\lambda, [v]_U\big) &\longmapsto [\lambda \cdot v]_U \end{aligned}$$

sind wohldefiniert und definieren auf W die Struktur eines K-Vektorraums.

ii) *Die Zuordnung*

$$\begin{aligned} \pi\colon V &\longrightarrow W \\ v &\longmapsto [v]_U \end{aligned}$$

ist eine surjektive lineare Abbildung mit

$$\mathrm{Ker}(\pi) = U.$$

Falls V endlichdimensional ist, so gilt

$$\mathrm{Dim}_K(W) = \mathrm{Dim}_K(V) - \mathrm{Dim}_K(U).$$

Beweis. i) Wir wissen bereits, dass „+“ wohldefiniert ist und $(W,+)$ eine abelsche Gruppe (Satz II.4.6). Es ist nun die Wohldefiniertheit von „·“ zu verifizieren. Dazu seien $\lambda \in K$, $v \in V$, $u \in U$ und $v' = v + u$. Wir wollen

$$[\lambda \cdot v']_U = [\lambda \cdot v]_U$$

zeigen. Dazu berechnen wir

$$\lambda \cdot v' - \lambda \cdot v = \lambda \cdot (v' - v) = \lambda \cdot u.$$

Das Element $\lambda \cdot u$ gehört zu U, denn U ist ein linearer Teilraum. Die Gültigkeit der verbleidenden Axiome (S1) - (S4) ist angesichts der Definition von „+“ und „·“ eine unmittelbare Konsequenz der Gültigkeit der entsprechenden Axiome im Vektorraum V.

ii) Die Linearität und die Surjektivität von π sind offensichtlich. Es sei $u \in U$. Dann gilt $0 - u = -u \in U$, so dass $u \sim_U 0$ und $[u]_U = [0]_U$. Dies zeigt $u \in \mathrm{Ker}(\pi)$. Sei umgekehrt $v \in \mathrm{Ker}(\pi)$. Dies bedeutet

$$[v]_U = \pi(v) = 0 = [0]_U,$$

so dass $0 \sim_U v$, also $v = v - 0 \in U$. □

Wir können jetzt endlich Problem I.4.6, ii), lösen:

Folgerung III.6.4. *Es seien $n \geq 1$ eine natürliche Zahl und $U \subseteq K^n$ ein linearer Teilraum. Dann existieren eine natürliche Zahl $m \in \{1, ..., n\}$ und eine $(m \times n)$-Matrix A mit*

$$U = \mathbb{L}(A, 0).$$

Beweis. Wir wählen eine durchnummerierte Basis B für den K-Vektorraum $W := K^n/U$ und definieren mit den Bezeichnungen aus Satz III.5.32 und III.6.3 die lineare Abbildung

$$g\colon K^n \xrightarrow{\pi} W \xrightarrow{\Phi_B} K^m, \quad m := \#B.$$

Es gibt eine $(m \times n)$-Matrix A, so dass $g = f_A$ (Folgerung III.5.22). Nun erkennen wir mit Satz III.6.3, ii), und Bemerkung III.5.18

$$U = \mathrm{Ker}(\pi) = \mathrm{Ker}(g) = \mathrm{Ker}(f_A) = \mathbb{L}(A, 0).$$

Die Gleichheit der Kerne ergibt sich aus der Tatsache, dass Φ_B ein Isomorphismus ist. □

Satz III.6.5 (Homomorphiesatz)**.** *Es seien V, W Vektorräume über dem Körper K und $f\colon V \longrightarrow W$ eine lineare Abbildung. Dann ist die Abbildung*

$$\begin{array}{rcl} \overline{f}\colon V/\mathrm{Ker}(f) & \longrightarrow & \mathrm{Bild}(f) \\ [v] & \longmapsto & f(v) \end{array}$$

wohldefiniert und ein linearer Isomorphismus.

Beweis. Wir zeigen, dass $\overline{f}$ wohldefiniert ist. Es seien $v \in V$, $u \in \mathrm{Ker}(f)$ und $v' = v + u$. Wir finden

$$f(v') = f(v+u) = f(v) + f(u) = f(v) + 0 = f(v).$$

Die Linearität ergibt sich folgendermaßen: Es seien $\lambda \in K$ und $v, v' \in V$. Dann haben wir

$$\overline{f}\big([v] + [v']\big) = \overline{f}\big([v+v']\big) = f(v+v') = f(v) + f(v') = \overline{f}\big([v]\big) + \overline{f}\big([v']\big)$$

und ebenso

$$\overline{f}\big(\lambda \cdot [v]\big) = \overline{f}\big([\lambda \cdot v]\big) = f(\lambda \cdot v) = \lambda \cdot f(v) = \lambda \cdot \overline{f}\big([v]\big).$$

Schließlich ist zu zeigen, dass f bijektiv ist. Sei $w \in \mathrm{Bild}(f)$. Dann existiert ein Vektor $v \in V$ mit $f(v) = w$, und wir finden

$$\overline{f}\big([v]\big) = f(v) = w.$$

Damit gilt $w \in \mathrm{Bild}(\overline{f})$. Diese Überlegung liefert die Surjektivität von $\overline{f}$. Für die Injektivität zeigen wir, dass der Kern trivial ist. Sei $v \in V$ ein Element mit $\overline{f}([v]) = 0$. Wegen $f(v) = \overline{f}([v])$ gilt $v \in \mathrm{Ker}(f)$ und deshalb $[v] = 0$. □

Dieser Satz gibt einen ersten Eindruck davon, wie Quotientenvektorräume auf natürliche Weise auftreten. In der Linearen Algebra kann man anstatt mit Quotientenvektorräumen oft auch mit linearen Komplementen arbeiten (vgl. Aufgabe III.6.7). Wir haben in Beispiel III.5.37, iii), gesehen, dass lineare Komplemente im Allgemeinen nicht eindeutig bestimmt sind. Dagegen ist die Konstruktion des Quotientenvektorraums kanonisch, d.h. sie hängt nicht von zusätzlichen Wahlen ab. Die Konstruktion des Quotientenvektorraums funktioniert auch im unendlichdimensionalen Kontext problemlos. Dagegen erfordert der Beweis der Existenz eines linearen Komplements in dieser Situation das Auswahlaxiom. Man beachte weiter, dass für einen K-Vektorraum V, einen linearen Teilraum $U \subset V$ und ein lineares Komplement U' von U die induzierte Abbildung $\pi_{|U'}\colon U' \longrightarrow V/U$ ein Isomorphismus ist. Man kann den Quotientenvektorraum also als intrinsische Version des linearen Komplements ansehen. Allerdings ist der Quotientenvektorraum kein linearer Teilraum des ursprünglichen Vektorraums. Schließlich sei darauf hingewiesen, dass in anderen Situationen, wie etwa der Gruppentheorie, Komplemente im Allgemeinen nicht

existieren und man somit auf die Konstruktion etwa von Faktorgruppen (vgl. Satz II.4.6) angewiesen ist. Für eine Diskussion von Quotientenvektorräumen mit Blick auf „kategorielle“ Konzepte verweisen wir auf [2], Kapitel II.

Aufgabe III.6.6.
a) Es seien $v_1, ..., v_s \in K^n$ und $U := \langle v_1, ..., v_s \rangle \subseteq K^n$. Entwickeln Sie mit Hilfe der obigen Ergebnisse einen Algorithmus zum Auffinden einer $(m \times n)$-Matrix A, für die

$$\mathbb{L}(A, 0) = U$$

gilt. Hierbei ist $m = n - \mathrm{Dim}_K(U)$.
Hinweis. Geben Sie zunächst ein Verfahren an, dass aus den Vektoren $v_1, ..., v_s$ eine Basis für U auswählt.
b) Bestimmen Sie mit dem Verfahren aus a) eine Matrix A, für die

$$\mathbb{L}(A, 0) = \left\langle \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ -3 \\ 4 \\ -5 \end{pmatrix}, \begin{pmatrix} 3 \\ -2 \\ 9 \\ -4 \\ 15 \end{pmatrix}, \begin{pmatrix} 4,5 \\ 1 \\ 13,5 \\ 2 \\ 23,5 \end{pmatrix} \right\rangle \subseteq \mathbb{R}^5$$

gilt.

Aufgabe III.6.7.
Beweisen Sie Folgerung III.6.4 mit Hilfe von linearen Komplementen anstatt mit Quotientenvektorräumen.

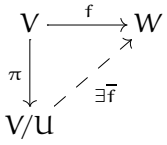
Aufgabe III.6.8.
Es seien V, W zwei K-Vektorräume, $f\colon V \longrightarrow W$ eine lineare Abbildung, $U \subseteq V$ ein linearer Teilraum und $\pi\colon V \longrightarrow V/U$ die Projektion aus Satz III.6.3, ii).
a) Beweisen Sie, dass es genau dann eine lineare Abbildung $\overline{f}\colon V/U \longrightarrow W$ mit

$$f = \overline{f} \circ \pi$$

gibt,[12] wenn

$$U \subseteq \mathrm{Ker}(f)$$

[12] Man kann sich diese Bedingung mit dem Diagramm



„veranschaulichen“ (vgl. Definition IV.1.21, iii).

gilt. Überprüfen Sie auch, dass $\overline{f}$ in diesem Fall eindeutig bestimmt ist.
b) Weisen Sie nach, dass im Fall $U \subseteq \mathrm{Ker}(f)$ die Gleichung

$$\mathrm{Ker}(\overline{f}) = \mathrm{Ker}(f)/U$$

erfüllt ist. Erläutern Sie dazu auch kurz, inwiefern $\mathrm{Ker}(f)/U$ ein Unterraum von V/U ist.
Zeigen Sie weiter

$$\mathrm{Bild}(\overline{f}) = \mathrm{Bild}(f).$$

Aufgabe III.6.9.
Es seien V, W endlichdimensionale K-Vektorräume und $f\colon V \longrightarrow W$ eine lineare Abbildung. Der Vektorraum

$$\mathrm{Koker}(f) := W/\mathrm{Bild}(f)$$

wird als *Kokern* von f angesprochen.
a) Beweisen Sie, dass die Abbildung f genau dann surjektiv ist, wenn $\mathrm{Koker}(f) = \{0\}$ gilt.
b) Zeigen Sie, dass folgende **Dimensionsformel** gilt:

$$\mathrm{Dim}_K\big(V/\mathrm{Ker}(f)\big) + \mathrm{Dim}_K\big(\mathrm{Koker}(f)\big) = \mathrm{Dim}_K(W).$$

Aufgabe III.6.10.
Wir betrachten die Abbildung $f_A\colon \mathbb{R}^3 \longrightarrow \mathbb{R}^4$ zu der Matrix

$$A := \begin{pmatrix} 1 & 0 & 1 \\ 2 & -1 & 3 \\ 0 & -1 & 1 \\ 4 & -2 & 6 \end{pmatrix} \in \mathrm{Mat}(4,3;\mathbb{R}).$$

Bestimmen Sie $\mathrm{Ker}(f)$, $\mathrm{Bild}(f)$, und geben Sie eine Basis für $\mathrm{Koker}(f)$ an.

IV

Matrizenrechnung

Im vorangegangenen Kapitel haben wir die Grundbegriffe der Linearen Algebra eingeführt. Dies sind Vektorräume und lineare Abbildungen zwischen Vektorräumen. Neben dem abstrakten Begriffsapparat und den theoretischen Aussagen spielen explizite Berechnungen eine große Rolle. Dazu verwendet man Matrizen. Wir haben bereits gesehen, wie man für gegebene natürliche Zahlen $m, n \geq 1$ lineare Abbildungen von K^n nach K^m mit Hilfe von Matrizen beschreiben kann, und verschiedene abstrakte Begriffe der Linearen Algebra mit Konzepten aus der Theorie linearer Gleichungssysteme verglichen. Die Beschreibung linearer Abbildungen durch Matrizen lässt sich auf endlichdimensionale Vektorräume **zusammen** mit einer durchnummerierten Basis ausdehnen. Geschickte Wahlen von Basen führen zu besonders einfachen Matrizen, an denen man die wichtigsten Daten der linearen Abbildung wie Kern, Bild und Rang sofort ablesen kann. Dies kann man auch anwenden, wenn die beteiligten Vektorräume von der Form K^n sind, d.h. es kann sinnvoll sein, auch mit anderen Basen als der Standardbasis zu arbeiten. In den Beweisen spielen die Resultate zu Zeilen- und Spaltenumformungen aus dem ersten Kapitel eine wichtige Rolle. Wir werden dazu Zeilen- und Spaltenumformungen durch das Matrixprodukt beschreiben. Schließlich werden wir die Theorie der Determinanten entwickeln, der wir bereits im ersten Kapitel für (2×2)-Matrizen begegnet sind. Sie wird für das folgende Kapitel unverzichtbar sein.

IV.1 Das Matrixprodukt

Beobachtung IV.1.1. *Es seien $m, n \geq 1$ natürliche Zahlen. Dann ist die Abbildung*

$$\begin{array}{rcl} \Psi_{m,n} \colon \mathrm{Mat}(m,n;K) & \longrightarrow & \mathrm{Hom}_K(K^n, K^m) \\ A & \longmapsto & f_A \end{array}$$

ein linearer Isomorphismus.

Beweis. In Folgerung III.5.22 wurde bereits nachgewiesen, dass $\Psi_{m,n}$ eine bijektive Abbildung ist. Wir müssen somit noch überprüfen, dass $\Psi_{m,n}$ linear ist. Dazu erinnern wir an folgende Tatsachen:

- Zwei lineare Abbildungen $f, g \colon K^n \longrightarrow K^m$ sind genau dann gleich, wenn $f(e_j) = g(e_j)$ für $j = 1, \ldots, n$ gilt (Satz III.5.21).
- Für eine Matrix $A \in \mathrm{Mat}(m,n;K)$ stimmt $f(e_j)$ mit der j-ten Spalte von A überein, $j = 1, \ldots, n$ (s. Definition III.5.11).

Es seien $A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$, $B = (b_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} \in \mathrm{Mat}(m,n;K)$ Matrizen und $\lambda \in K$ eine Zahl. Dann ist die j-te Spalte der Matriz $A + B$ durch

$$\begin{pmatrix} a_{1j} + b_{1j} \\ \vdots \\ a_{mj} + b_{mj} \end{pmatrix}$$

gegeben und die j-te Spalte von $\lambda \cdot A$ durch

$$\begin{pmatrix} \lambda \cdot a_{1j} \\ \vdots \\ \lambda \cdot a_{mj} \end{pmatrix}, \quad j = 1, \ldots, n.$$

Für $j \in \{1, \ldots, n\}$ folgt also

$$f_{A+B}(e_j) = \begin{pmatrix} a_{1j} + b_{1j} \\ \vdots \\ a_{mj} + b_{mj} \end{pmatrix} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} + \begin{pmatrix} b_{1j} \\ \vdots \\ b_{mj} \end{pmatrix} = f_A(e_j) + f_B(e_j)$$

und

$$f_{\lambda \cdot A}(e_j) = \begin{pmatrix} \lambda \cdot a_{1j} \\ \vdots \\ \lambda \cdot a_{mj} \end{pmatrix} = \lambda \cdot \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = \lambda \cdot f_A(e_j).$$

Auf Grund der oben genannten Eigenschaften sind f_{A+B} und $f_A + f_B$ bzw. $f_{\lambda \cdot A}$ und $\lambda \cdot f_A$ identisch. □

Definition IV.1.2. Es seien $l, m, n \geq 1$ natürliche Zahlen und $A \in \mathrm{Mat}(l, m; K)$, $B \in \mathrm{Mat}(m, n; K)$ Matrizen. Das *Produkt* von A und B ist die Matrix

$$A \cdot B := \Psi_{l,n}^{-1}(f_A \circ f_B) \in \mathrm{Mat}(l, n; K).$$

Bemerkungen IV.1.3. i) Das Matrixprodukt ist durch die Eigenschaft $f_{A\cdot B} = f_A \circ f_B$ charakterisiert.

ii) Man beachte, dass das Matrixprodukt nur definiert ist, wenn die Anzahl der Spalten von A mit der Anzahl der Zeilen von B übereinstimmt.

Wir wollen jetzt eine explizite Formel für das Produkt zweier Matrizen erarbeiten. In der Situation von Definition IV.1.2 seien

$$A = (a_{ij})_{\substack{i=1,\dots,l \\ j=1,\dots,m}} \quad \text{und} \quad B = (b_{jk})_{\substack{j=1,\dots,m \\ k=1,\dots,n}}.$$

Behauptung. *Es sei*

$$A \cdot B = (c_{ik})_{\substack{i=1,\dots,l \\ k=1,\dots n}}.$$

Dann gilt

$$c_{ik} = \sum_{j=1}^{m} a_{ij} \cdot b_{jk}, \quad i = 1, \dots, l, \ k = 1, \dots, n. \tag{IV.1}$$

Beweis. Für $k \in \{1, \dots, n\}$ hat man

$$\begin{pmatrix} c_{1k} \\ \vdots \\ c_{lk} \end{pmatrix} = f_{A\cdot B}(e_k) = f_A\big(f_B(e_k)\big) = f_A \begin{pmatrix} b_{1k} \\ \vdots \\ b_{mk} \end{pmatrix} = \sum_{j=1}^{m} b_{jk} \cdot f_A(e_j) = \sum_{j=1}^{m} b_{jk} \cdot \begin{pmatrix} a_{1j} \\ \vdots \\ a_{lj} \end{pmatrix}.$$

Daraus liest man die behauptete Formel ab. □

Bemerkung IV.1.4. Es seien $i \in \{1, \dots, l\}$ und $k \in \{1, \dots, n\}$. Um das Element c_{ik} in der Produktmatrix $A \cdot B$ zu ermitteln, benötigt man die i-te Zeile der Matrix A und die k-te Spalte der Matrix B. Wir multiplizieren komponentenweise, d.h. den ersten Eintrag der Zeile mit dem ersten Eintrag der Spalte, den zweiten Eintrag der Zeile mit dem zweiten Eintrag der Spalte usw., und summieren dann auf. Für den Anfang mag die Schreibweise die Berechnung in Abbildung IV.1 vereinfachen.

Beispiel IV.1.5. Wir berechnen das Produkt einer (3×3)-Matrix und einer (3×2)-Matrix mit rationalen Koeffizienten:

$$\begin{pmatrix} -1 & 0 & 3 \\ 2 & -4 & 1 \\ 1 & 5 & 0 \end{pmatrix} \cdot \begin{pmatrix} 2 & 3 \\ 1 & 5 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} (-1\cdot 2 + 0\cdot 1 + 3\cdot 3) & (-1\cdot 3 + 0\cdot 5 + 3\cdot 1) \\ (2\cdot 2 - 4\cdot 1 + 1\cdot 3) & (2\cdot 3 - 4\cdot 5 + 1\cdot 1) \\ (1\cdot 2 + 5\cdot 1 + 0\cdot 3) & (1\cdot 3 + 5\cdot 5 + 0\cdot 1) \end{pmatrix}$$
$$= \begin{pmatrix} 7 & 0 \\ 3 & -13 \\ 7 & 28 \end{pmatrix}.$$

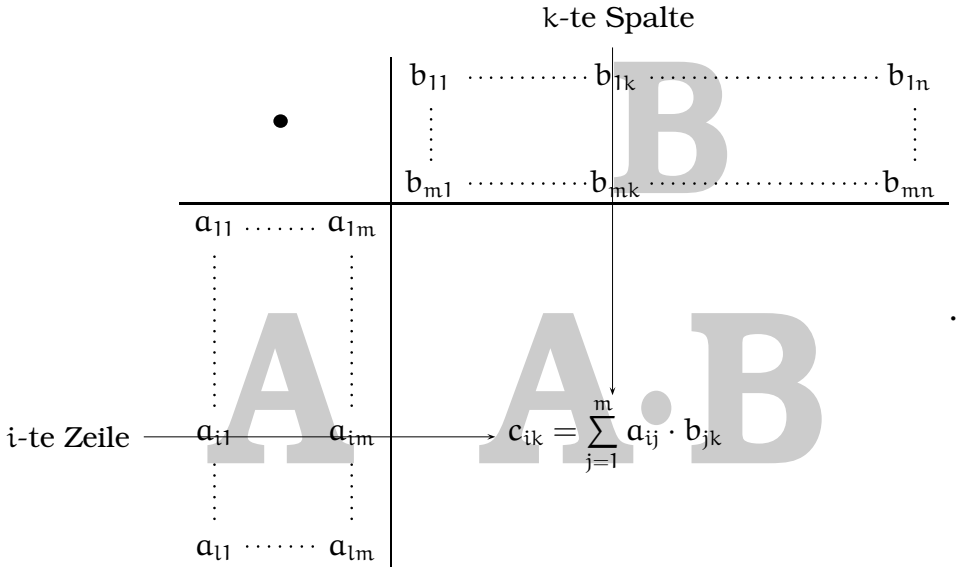


Abbildung IV.1: Zum Matrixprodukt

Satz IV.1.6. *Es seien* $l, m, n, p \geq 1$ *natürliche Zahlen,* $A, A' \in \mathrm{Mat}(l, m; K)$, $B, B' \in \mathrm{Mat}(m, n; K)$, $C \in \mathrm{Mat}(n, p; K)$ *Matrizen und* $\lambda \in K$. *Dann gelten folgende Rechenregeln:*

i) $A \cdot (B + B') = A \cdot B + A \cdot B'$, $(A + A') \cdot B = A \cdot B + A' \cdot B$.

ii) $A \cdot (\lambda \cdot B) = (\lambda \cdot A) \cdot B = \lambda \cdot (A \cdot B)$.

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

Beweis. Diese Regeln ergeben sich unmittelbar aus den entsprechenden Regeln für Abbildungen (s. Satz III.5.24). Für iii) beobachten wir z.B.

$$f_{(A\cdot B)\cdot C} = f_{A\cdot B} \circ f_C = (f_A \circ f_B) \circ f_C = f_A \circ (f_B \circ f_C) = f_A \circ f_{B\cdot C} = f_{A\cdot(B\cdot C)}.$$

Zusammen mit Beobachtung IV.1.1 folgt die behauptete Gleichheit. □

Folgerung IV.1.7. *Das Tupel* $(\mathrm{Mat}(n; K), +, \cdot)$ *ist ein Ring mit Einselement* $\mathbb{E}_n$.

Beispiel IV.1.8. Wir setzen

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

und berechnen

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

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

Insbesondere hat der Ring $\mathrm{Mat}(n; K)$ Nullteiler.

Zeilen- und Spaltenoperation. — Die in Definition I.3.1 und Aufgabe I.4.10 eingeführten Zeilen- und Spaltenumformungen lassen sich mit Hilfe von Matrixprodukten beschreiben.

Definitionen IV.1.9. Es sei $n \geq 1$ eine natürliche Zahl.

i) Für Indizes $k, l \in \{1, ..., n\}$ mit $k \neq l$ und eine Zahl $\lambda \in K$ sei die $(n \times n)$-Matrix $Q_{kl}(\lambda) = (q_{ij})_{\substack{i=1,...,n \\ j=1,...,n}}$ (s. Abbildung IV.2) definiert durch

$$q_{ij} := \begin{cases} 1, & i = j \\ \lambda, & i = k \wedge j = l \\ 0, & \text{sonst} \end{cases}, \quad i, j = 1, ..., n.$$

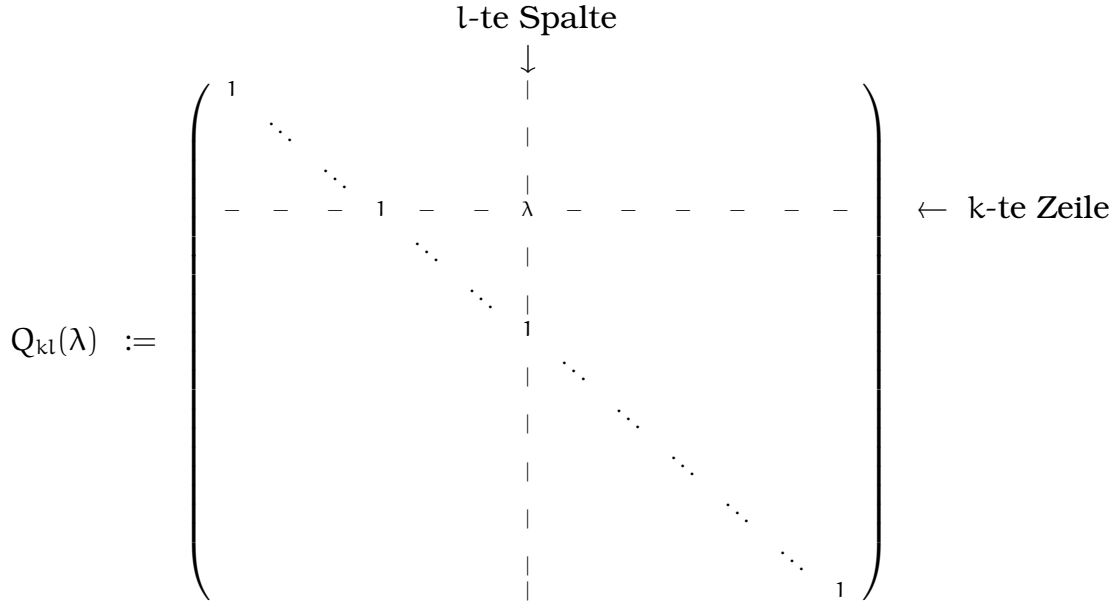


Abbildung IV.2: Die Elementarmatrix $Q_{kl}(\lambda)$

ii) Für Indizes $1 \leq k < l \leq n$ sei die $(n \times n)$-Matrix $P_{kl} = (p_{ij})_{\substack{i=1,...,n \\ j=1,...,n}}$ (s. Abbildung IV.3) gegeben durch

$$p_{ij} := \begin{cases} 1, & i = j \wedge i \notin \{k, l\} \\ 1, & i = k \wedge j = l \\ 1, & i = l \wedge j = k \\ 0, & \text{sonst} \end{cases}, \quad i, j = 1, ..., n.$$

$$
P_{kl} := \left(\begin{array}{ccccccccccc}
1 & & & | & & & & | & & & \\
& \ddots & & | & & & & | & & & \\
& & 1 & | & & & & | & & & \\
- & - & - & 0 & - & - & - & 1 & - & - & - \\
& & & | & 1 & & & | & & & \\
& & & | & & \ddots & & | & & & \\
& & & | & & & 1 & | & & & \\
- & - & - & 1 & - & - & - & 0 & - & - & - \\
& & & | & & & & | & 1 & & \\
& & & | & & & & | & & \ddots & \\
& & & | & & & & | & & & 1
\end{array}\right)
\begin{array}{l}
\\ \\ \\ \leftarrow \text{ k-te Zeile} \\ \\ \\ \\ \leftarrow \text{ l-te Zeile} \\ \\ \\ \\
\end{array}
$$

k-te Spalte l-te Spalte

Abbildung IV.3: Die Elementarmatrix P_{kl}

iii) Für einen Index $k \in \{1, ..., n\}$ und eine Zahl $\lambda \in K \setminus \{0\}$ sei $S_k(\lambda) = (s_{ij})_{\substack{i=1,...,n \\ j=1,...,n}}$ die $(n \times n)$-Matrix (s. Abbildung IV.4) mit

$$
s_{ij} := \begin{cases} \lambda, & i = j = k \\ 1, & i = j \neq k \\ 0, & \text{sonst} \end{cases}, \quad i, j = 1, ..., n.
$$

k-te Spalte

$$
S_k(\lambda) := \left(\begin{array}{ccccccc}
1 & & & | & & & \\
& \ddots & & | & & & \\
& & 1 & | & & & \\
- & - & - & \lambda & - & - & - \\
& & & | & 1 & & \\
& & & | & & \ddots & \\
& & & | & & & 1
\end{array}\right)
\begin{array}{l}
\\ \\ \\ \leftarrow \text{ k-te Zeile} \\ \\ \\ \\
\end{array}
$$

Abbildung IV.4: Die Elementarmatrix $S_k(\lambda)$

iv) Eine Matrix $A \in \mathrm{Mat}(n; K)$ nennt sich *Elementarmatrix*, wenn sie eine der Gestalten aus Teil i) - iii) annimmt.

Satz IV.1.10. *Es seien* $m, n \geq 1$, $A \in \mathrm{Mat}(m, n; K)$ *eine* $(m \times n)$*-Matrix und* $\lambda \in K$.

i) *Für* $1 \leq k \neq l \leq m$ *seien* $A^{(1)}$ *die Matrix, die aus* A *entsteht, wenn man das* λ*-Fache der* l*-ten Zeile zur* k*-ten addiert,* $A^{(2)}$ *die Matrix, die man erhält, indem man in* A *die* k*-te und die* l*-te Zeile vertauscht, und* $A^{(3)}$ *diejenige, die sich aus* A *durch Multiplikation der* k*-ten Zeile mit* λ *ergibt. Dann hat man folgende Formeln*[1]

$$A^{(1)} = Q_{kl}(\lambda) \cdot A, \quad A^{(2)} = P_{kl} \cdot A \quad \textit{und} \quad A^{(3)} = S_k(\lambda) \cdot A.$$

ii) *Es seien* $1 \leq k \neq l \leq n$, $A_{(1)}$ *die Matrix, die aus* A *entsteht, wenn man das* λ*-Fache der* k*-ten Spalte zur* l*-ten addiert,* $A_{(2)}$ *diejenige, die man erhält, indem man in* A *die* k*-te und die* l*-te Spalte vertauscht, und* $A_{(3)}$ *die Matrix, die sich aus* A *durch Multiplikation der* k*-ten Spalte mit* λ *ergibt. Dann hat man folgende Formeln*

$$A_{(1)} = A \cdot Q_{kl}(\lambda), \quad A_{(2)} = A \cdot P_{kl} \quad \textit{und} \quad A_{(3)} = A \cdot S_k(\lambda).$$

Beweis. Das ergibt sich durch direktes Nachrechnen. □

Folgerung IV.1.11. *Es sei* $A \in \mathrm{Mat}(m, n; K)$ *eine* $(m \times n)$*-Matrix vom Rang* r. *Dann existieren natürliche Zahlen* $s, t \in \mathbb{N}$ *und Elementarmatrizen* $B_1, \ldots, B_s \in \mathrm{Mat}(m; K)$ *und* $C_1, \ldots, C_t \in \mathrm{Mat}(n; K)$, *so dass*

$$B_1 \cdot \cdots \cdot B_s \cdot A \cdot C_1 \cdot \cdots \cdot C_t = \left(\begin{array}{c|c} \mathbb{E}_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

Beweis. Das ergibt sich aus dem vorangehenden Satz und Aufgabe I.4.10. □

Beispiele IV.1.12. i) Die folgenden Matrizen mit Einträgen im Körper $\mathbb{Q}$ illustrieren, wie Zeilen- bzw. Spaltenoperationen vom Typ (ZI) bzw. (SI) durch Links- bzw. Rechtsmultiplikation mit einer Elementarmatrix vom Typ $Q_{kl}(\lambda)$ realisiert werden können:

$$\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -2 & 7 & 0 \\ -1 & 5 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1-\lambda & -2+5\lambda & 7 & 2\lambda \\ -1 & 5 & 0 & 2 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ -1 & -1 & 0 \\ 2 & 0 & -2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ \lambda & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 \\ 1+\lambda & 1 & 1 \\ -1-\lambda & -1 & 0 \\ 2 & 0 & -2 \end{pmatrix}.$$

[1] Hier sind für $Q_{kl}(\lambda)$, P_{kl} und $S_k(\lambda)$ die entsprechenden $(m \times m)$-Matrizen einzusetzen.

ii) Die Vertauschung zweier Zeilen bzw. Spalten erhält man durch Links- bzw. Rechtsmultiplikation mit einer Matrix vom Typ P_{ij}:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & -2 & 7 & 0 \\ -1 & 5 & 0 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 5 & 0 & 2 \\ 1 & -2 & 7 & 0 \end{pmatrix},$$
$$\begin{pmatrix} 1 & -2 & 7 & 0 \\ -1 & 5 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 7 & -2 \\ -1 & 2 & 0 & 5 \end{pmatrix}.$$

iii) Mit Hilfe von Elementarmatrizen vom Typ $S_k(\lambda)$ kann man Zeilen- bzw. Spaltenoperationen vom Typ (ZIII) bzw. (SIII) beschreiben:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 & 3 \\ 2 & -4 & 1 \\ 1 & 5 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 3 \\ \lambda \cdot 2 & -\lambda \cdot 4 & \lambda \cdot 1 \\ 1 & 5 & 0 \end{pmatrix},$$
$$\begin{pmatrix} -1 & 0 & 3 \\ 2 & -4 & 1 \\ 1 & 5 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda \end{pmatrix} = \begin{pmatrix} -1 & 0 & 3 \cdot \lambda \\ 2 & -4 & 1 \cdot \lambda \\ 1 & 5 & 0 \end{pmatrix}.$$

Invertierbare Matrizen. — In Beobachtung IV.1.1 haben wir gesehen, wie lineare Abbildungen von K^n nach K^n mit $(n \times n)$-Matrizen beschrieben werden können. Matrizen, die zu Isomorphismen führen, verdienen besondere Beachtung.

Lemma IV.1.13. *Es seien $n \geq 1$ eine natürliche Zahl und $A \in \mathrm{Mat}(n; K)$ eine $(n \times n)$-Matrix. Dann sind folgende Bedingungen äquivalent:*

i) *Die lineare Abbildung $f_A \colon K^n \longrightarrow K^n$ ist ein Isomorphismus.*

ii) *Es gibt eine $(n \times n)$-Matrix B mit $A \cdot B = \mathbb{E}_n = B \cdot A$.*

iii) *Die Matrix hat Rang $\mathrm{Rg}(A) = n$.*

Beweis. **„i)⟹ii)".** Wenn f_A ein Isomorphismus ist, dann gibt es eine lineare Abbildung $g \colon K^n \longrightarrow K^n$ mit $g \circ f = \mathrm{Id}_{K^n} = f \circ g$ sowie eine $(n \times n)$-Matrix $B \in \mathrm{Mat}(n; K)$ mit $g = f_B$. Mit der Injektivität der Abbildung $\Psi_{n,n}$ (s. Beobachtung IV.1.1) folgert man aus

$$f_{B \cdot A} = f_B \circ f_A = \mathrm{Id}_{K^n} = f_{\mathbb{E}_n},$$

dass $B \cdot A = \mathbb{E}_n$. Die Gleichung $A \cdot B = \mathbb{E}_n$ erhält man genauso.

„ii)⟹i)". Es gilt

$$f_A \circ f_B = f_{A \cdot B} = f_{\mathbb{E}_n} = \mathrm{Id}_{K^n} \quad \text{und} \quad f_B \circ f_A = f_{B \cdot A} = f_{\mathbb{E}_n} = \mathrm{Id}_{K^n}.$$

Damit ist f_A nach Bemerkung III.5.9 bijektiv und somit ein Isomorphismus (Definition III.5.10).

„i)$\Longleftrightarrow$iii)". Nach Folgerung III.5.41, iii), ist f_A genau dann ein Isomorphismus, wenn f_A surjektiv ist. Letzteres ist äquivalent zu $\mathrm{Bild}(f_A) = K^n$ und somit zu $\mathrm{Dim}_K(\mathrm{Bild}(f_A)) = n$. Die Dimension des Bilds von f_A ist gerade der Rang von A (Definition III.4.12, ii), und III.5.42). □

Definition IV.1.14. Eine $(n \times n)$-Matrix $A \in \mathrm{Mat}(n; K)$ ist *invertierbar*, wenn es eine $(n \times n)$-Matrix $B \in \mathrm{Mat}(n; K)$ mit

$$A \cdot B = \mathbb{E}_n = B \cdot A$$

gibt. Die Matrix B heißt dann die zu A *inverse Matrix*.
Schreibweise. $A^{-1} := B$.

Bemerkung IV.1.15. Die inverse Matrix ist eindeutig bestimmt: Es seien A, B, $B' \in \mathrm{Mat}(n; K)$ Matrizen mit $A \cdot B = \mathbb{E}_n = B' \cdot A$. Dann gilt $B' = B' \cdot \mathbb{E}_n = B' \cdot (A \cdot B) = (B' \cdot A) \cdot B = \mathbb{E}_n \cdot B = B$.

Beispiele IV.1.16. i) Es seien $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ eine (2×2)-Matrix und $\mathrm{Det}(A) := \delta(a, b, c, d)$ die Determinante von A (Definition I.1.2). Mit Aufgabe I.1.6 und Lemma IV.1.13 schließt man, dass A genau dann invertierbar ist, wenn $\mathrm{Det}(A) \neq 0$. In diesem Fall hat man

$$A^{-1} = \frac{1}{\mathrm{Det}(A)} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

ii) Die Elementarmatrizen aus Definition IV.1.9 sind invertierbar, und für die inversen Matrizen gelten die Formeln

$$\begin{array}{rcll} Q_{kl}(\lambda)^{-1} & = & Q_{kl}(-\lambda), & k, l \in \{1, ..., n\},\ k \neq l,\ \lambda \in K, \\ P_{kl}^{-1} & = & P_{kl}, & 1 \leq k < l \leq n, \\ S_k(\lambda)^{-1} & = & S_k(\lambda^{-1}), & k \in \{1, ..., n\},\ \lambda \in K. \end{array}$$

Satz IV.1.17. *Es seien $n \geq 1$ eine natürliche Zahl und $A \in \mathrm{Mat}(n; K)$ eine **invertierbare** $(n \times n)$-Matrix.*

i) *Es gibt eine natürliche Zahl $u \geq 1$ und Elementarmatrizen $D_1, ..., D_u$, so dass*

$$A = D_1 \cdot \cdots \cdot D_u.$$

ii) *Die Matrix A kann durch eine Folge von Zeilenoperationen vom Typ* (ZI)*,* (ZII) *und* (ZIII) *in die Einheitsmatrix überführt werden.*

Beweis. i) Eine invertierbare Matrix hat Rang n (Lemma IV.1.13). Nach Folgerung IV.1.11 gibt es natürliche Zahlen $s, t \geq 0$ und Elementarmatrizen $B_1, ..., B_s$, $C_1, ..., C_t$ mit

$$B_1 \cdot \cdots \cdot B_s \cdot A \cdot C_1 \cdot \cdots \cdot C_t = \mathbb{E}_n.$$

Es ergibt sich

$$A = B_s^{-1} \cdot \cdots \cdot B_1^{-1} \cdot C_t^{-1} \cdot \cdots \cdot C_1^{-1}.$$

ii) Aus der gerade bewiesenen Darstellung von A leiten wir

$$C_1 \cdot \cdots \cdot C_t \cdot B_1 \cdot \cdots \cdot B_s \cdot A = \mathbb{E}_n$$

ab. Da die Linksmultiplikation mit einer Elementarmatrix einer Zeilenoperation entspricht, ergibt sich die Behauptung. □

Berechnung der inversen Matrix. — Teil i) des Satzes zeigt uns auch, wie wir die inverse Matrix ermitteln können. Dazu seien $n \geq 1$ eine natürliche Zahl und $A \in \mathrm{Mat}(n; K)$ eine **invertierbare** $(n \times n)$-Matrix. Wir bilden die $(n \times 2n)$-Matrix

$$M := (A|\mathbb{E}_n)$$

und führen sukzessiv Zeilenoperationen vom Typ (ZI) - (ZIII) aus, so dass A in die Einheitsmatrix überführt wird. Das Ergebnis sei

$$M' = (\mathbb{E}_n|B).$$

Dann gilt

$$B = A^{-1}.$$

Beispiel IV.1.18. Wir führen dieses Verfahren für eine (3×3)-Matrix mit rationalen Koeffizienten vor:

Zeilenoperation	
	$\left(\begin{array}{ccc\|ccc} 0 & 1 & -4 & 1 & 0 & 0 \\ 1 & 2 & -1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{array}\right)$
$I \leftrightarrow II$	$\left(\begin{array}{ccc\|ccc} 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{array}\right)$
$III - I$	$\left(\begin{array}{ccc\|ccc} 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & -1 & 3 & 0 & -1 & 1 \end{array}\right)$
$III + II$	$\left(\begin{array}{ccc\|ccc} 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & -1 & 1 \end{array}\right)$
$(-1) \cdot III$	$\left(\begin{array}{ccc\|ccc} 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & -1 \end{array}\right)$

$I - 2 \cdot II$	$\left(\begin{array}{ccc\|ccc} 1 & 0 & 7 & -2 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & -1 \end{array}\right)$
$I - 7 \cdot III$	$\left(\begin{array}{ccc\|ccc} 1 & 0 & 0 & 5 & -6 & 7 \\ 0 & 1 & -4 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & -1 \end{array}\right)$
$II + 4 \cdot III$	$\left(\begin{array}{ccc\|ccc} 1 & 0 & 0 & 5 & -6 & 7 \\ 0 & 1 & 0 & -3 & 4 & -4 \\ 0 & 0 & 1 & -1 & 1 & -1 \end{array}\right)$

.

Zur Probe berechnen wir:

$$\begin{pmatrix} 0 & 1 & -4 \\ 1 & 2 & -1 \\ 1 & 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 5 & -6 & 7 \\ -3 & 4 & -4 \\ -1 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Die allgemeine lineare Gruppe. — In Definition III.5.30 haben wir die allgemeine lineare Gruppe eines Vektorraums eingeführt. Wir können jetzt die entsprechende Gruppe für Matrizen definieren. Es sei $n \geq 1$ eine natürliche Zahl. Man definiere

$$\mathrm{GL}_n(K) := \left\{ A \in \mathrm{Mat}(n; K) \,|\, A \text{ ist invertierbar} \right\}.$$

Es ist leicht zu sehen (vgl. Aufgabe IV.1.33), dass

$$\forall A, B \in \mathrm{GL}_n(K): \quad A \cdot B \in \mathrm{GL}_n(K).$$

Somit ist die Multiplikation

$$\begin{array}{rcl} \cdot \colon \mathrm{GL}_n(K) \times \mathrm{GL}_n(K) & \longrightarrow & \mathrm{GL}_n(K) \\ (A, B) & \longmapsto & A \cdot B \end{array}$$

erklärt. Mit Hilfe von Satz IV.1.6, iii), sieht man ein:

Satz IV.1.19. *Die Menge* $\mathrm{GL}_n(K)$ *zusammen mit dem Matrixprodukt ist eine Gruppe.*

Definition IV.1.20. Die Gruppe $\mathrm{GL}_n(K)$ wird als *allgemeine lineare Gruppe* bezeichnet.

Darstellungsmatrizen. — Ein n-dimensionaler K-Vektorraum zusammen mit einer durchnummerierten Basis kann auf kanonische Weise mit K^n zusammen mit der Standardbasis identizifiert werden (Satz III.5.32, i). Mit dieser Beobachtung lässt sich die Beschreibung linearer Abbildungen durch Matrizen ausdehnen.

Definitionen IV.1.21. i) Es seien V, W zwei endlichdimensionale K-Vektorräume, $B = \{b_1, ..., b_n\}$ eine durchnummerierte Basis für V, $C = \{c_1, ..., c_m\}$ eine für W und $f\colon V \longrightarrow W$ eine lineare Abbildung. Die eindeutig bestimmte $(m \times n)$-Matrix A, für die, in den Bezeichnungen von Definition III.5.11 und Satz III.5.32, i), die Gleichung

$$f_A = \Phi_C \circ f \circ \Phi_B^{-1}$$

gilt, ist die *Darstellungsmatrix für* f bzgl. der durchnummerierten Basen B und C.

Schreibweise. $M_C^B(f) := A$.

ii) Es seien V, W, B, C wie in Teil i). Zusätzlich sei eine $(m \times n)$-Matrix A gegeben. Die lineare Abbildung

$$L_C^B(A) := \Phi_C^{-1} \circ f_A \circ \Phi_B$$

ist die zu A bzgl. B und C *assoziierte lineare Abbildung*.

iii) Es seien A, B, C, D Mengen und $f\colon A \longrightarrow B$, $g\colon B \longrightarrow D$, $h\colon A \longrightarrow C$, $l\colon C \longrightarrow D$ Abbildungen. Das Diagramm

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle g} \\ C & \xrightarrow{l} & D \end{array}$$

ist *kommutativ*, wenn die Bedingung

$$g \circ f = l \circ h$$

erfüllt ist.

Bemerkungen IV.1.22. i) Die Matrix $M_C^B(f)$ ist durch die Bedingung, dass das Diagramm

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ {\scriptstyle \Phi_B}\downarrow & & \downarrow{\scriptstyle \Phi_C} \\ K^n & \xrightarrow{f_{M_C^B(f)}} & K^m \end{array}$$

kommutieren soll, festgelegt.

ii) Entsprechend ist die lineare Abbildung $L_C^B(A)$ durch das kommutative Diagramm

$$\begin{array}{ccc} V & \xrightarrow{L_C^B(A)} & W \\ {\scriptstyle \Phi_B}\downarrow & & \downarrow{\scriptstyle \Phi_C} \\ K^n & \xrightarrow{f_A} & K^m \end{array}$$

bestimmt.

iii) Es seien V, W endlichdimensionale K-Vektorräume, $B = \{b_1, ..., b_n\}$ bzw. $C = \{c_1, ..., c_m\}$ durchnummerierte Basen für V bzw. W und $A \in \mathrm{Mat}(m, n; K)$ eine $(m \times n)$-Matrix. Dann findet man

$$\mathrm{Rg}\big(L_C^B(A)\big) = S\text{-}\mathrm{Rg}(A).$$

Dazu betrachte man das kommutative Diagramm

$$\begin{array}{ccc} V & \xrightarrow{L_C^B(A)} & W \\ {\scriptstyle \Phi_B}\downarrow & & \downarrow{\scriptstyle \Phi_C} \\ K^n & \xrightarrow{f_A} & K^m \end{array}.$$

Da Φ_B und Φ_C Isomorphismen sind, hat man

$$\mathrm{Dim}_K\Big(\mathrm{Bild}\big(L_C^B(A)\big)\Big) = \mathrm{Dim}_K(f_A) = S\text{-}\mathrm{Rg}(A).$$

Die letzte Gleichung haben wir bereits in Satz III.5.16 beobachtet.

Beobachtung IV.1.23. *Es seien V, W, B, C und f wie in Definition IV.1.21, i). Die Zahlen a_{ij}, $i = 1, ..., m$, $j = 1, ..., n$, werden durch die Formeln*

$$f(b_j) = \sum_{i=1}^{m} a_{ij} \cdot c_i, \quad j = 1, ..., n,$$

ermittelt. Dann gilt

$$M_C^B(f) = (a_{ij})_{\substack{i=1,...,m \\ j=1,...,n}}.$$

Beweis. Mit Satz III.5.32, i), und Definition IV.1.21, i), ergibt sich für einen Index $j \in \{1, ..., n\}$:

$$\begin{aligned} (\Phi_C \circ f \circ \Phi_B^{-1})(e_j) &= (\Phi_C \circ f)\big(\Phi_B^{-1}(e_j)\big) = (\Phi_C \circ f)(b_j) = \Phi_C\big(f(b_j)\big) \\ &= \Phi_C\Big(\sum_{i=1}^{m} a_{ij} \cdot c_i\Big) = \sum_{i=1}^{m} a_{ij} \cdot \Phi_C(c_i) \\ &= \sum_{i=1}^{m} a_{ij} \cdot e_i = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}. \end{aligned}$$

Daran lässt sich die Behauptung ablesen. □

Diese Beobachtung liefert ein Berechnungsverfahren für die Darstellungsmatrix.

Beobachtung IV.1.24. *Es seien* V, W, B, C *und* f *wie in Definition IV.1.21, i), und* $A = (a_{ij})_{\substack{i=1,\dots,m\\ j=1,\dots,n}}$ *eine* $(m \times n)$*-Matrix. Dann ist die lineare Abbildung* $L_C^B(A)$ *durch die Bedingung*

$$L_C^B(A)(b_j) = \sum_{i=1}^{m} a_{ij} \cdot c_i, \quad j = 1, \dots, n,$$

festgelegt.

Beweis. Diese einfache Verifikation überlassen wir der Leserin bzw. dem Leser als Übung. □

Satz IV.1.25. *Es seien* V, W, B, C *und* f *wie in Definition IV.1.21, i). Dann sind die Abbildungen*

$$\begin{aligned} M_C^B \colon \mathrm{Hom}_K(V,W) &\longrightarrow \mathrm{Mat}(m,n;K) \\ f &\longmapsto M_C^B(f) \end{aligned}$$

und

$$\begin{aligned} L_C^B \colon \mathrm{Mat}(m,n;K) &\longrightarrow \mathrm{Hom}_K(V,W) \\ A &\longmapsto L_C^B(A) \end{aligned}$$

zueinander inverse lineare Isomorphismen.

Ferner gilt

$$\forall f, g \in \mathrm{End}_K(V): \quad M_B^B(f \circ g) = M_B^B(f) \cdot M_B^B(g)$$

und

$$\forall A, D \in \mathrm{Mat}(n;K): \quad L_B^B(A \cdot D) = L_B^B(A) \circ L_B^B(D).$$

Beweis. Dass die angegebenen Abbildungen zueinander invers sind, kann man sich mit Hilfe der in Bemerkung IV.1.22, i) und ii), angegebenen Diagramme verdeutlichen. Formal argumentieren wir folgendermaßen: Für eine lineare Abbildung $f\colon V \longrightarrow W$ ist $L_C^B(M_C^B(f))$ die lineare Abbildung

$$\Phi_C^{-1} \circ f_{M_C^B(f)} \circ \Phi_B = \Phi_C^{-1} \circ (\Phi_C \circ f \circ \Phi_B^{-1}) \circ \Phi_B = (\Phi_C^{-1} \circ \Phi_C) \circ f \circ (\Phi_B^{-1} \circ \Phi_B) = \mathrm{Id}_W \circ f \circ \mathrm{Id}_V = f.$$

Sei umgekehrt eine $(m \times n)$-Matrix $A \in \mathrm{Mat}(m,n;K)$ gegeben. Dann ermitteln wir $M_C^B(L_C^B(A))$ durch

$$f_{M_C^B(L_C^B(A))} = \Phi_C \circ L_C^B(A) \circ \Phi_B^{-1} = \Phi_C \circ (\Phi_C^{-1} \circ f_A \circ \Phi_B) \circ \Phi_B^{-1} = f_A.$$

Aus der Injektivitätsaussage in Beobachtung IV.1.1 folgt $M_C^B(L_C^B(A)) = A$.

Wir überprüfen, dass M_C^B linear ist. Es seien $f, g\colon V \longrightarrow W$ lineare Abbildungen. Wir haben

$$\begin{aligned} f_{M_C^B(f+g)} &= \Phi_C \circ (f+g) \circ \Phi_B^{-1} = \Phi_C \circ (f \circ \Phi_B^{-1} + g \circ \Phi_B^{-1}) = \\ &= \Phi_C \circ f \circ \Phi_B^{-1} + \Phi_C \circ g \circ \Phi_B^{-1} = f_{M_C^B(f)} + f_{M_C^B(g)} = f_{M_C^B(f)+M_C^B(g)}. \end{aligned}$$

Hier haben wir zunächst die Distributivgesetze aus Satz III.5.24 verwendet. Im letzten Schritt haben wir die Linearität von $\Psi_{m,n}$ benutzt (s. Beobachtung IV.1.1). Die Injektivitätsaussage in dieser Beobachtung zeigt $M_C^B(f+g) = M_C^B(f) + M_C^B(g)$. Entsprechend weist man nach, dass $M_C^B(\lambda \cdot f) = \lambda \cdot M_C^B(f)$ für $\lambda \in K$ und $f \in \mathrm{Hom}_K(V,W)$ gilt.

Für die Überprüfung der Linearität von L_C^B kann man analog vorgehen. Man kann aber auch benutzen, dass L_C^B als Umkehrabbildung einer linearen Abbildung linear ist (Satz III.5.8).

Schließlich gilt nach Definition IV.1.21, i):

$$\begin{aligned} f_{M_B^B(f\circ g)} &= \Phi_B \circ f \circ g \circ \Phi_B^{-1} = \Phi_B \circ f \circ (\Phi_B^{-1} \circ \Phi_B) \circ g \circ \Phi_B^{-1} = \\ &= (\Phi_B \circ f \circ \Phi_B^{-1}) \circ (\Phi_B \circ g \circ \Phi_B^{-1}) = f_{M_B^B(f)} \circ f_{M_B^B(g)} = f_{M_B^B(f)\cdot M_B^B(g)}. \end{aligned}$$

Die letzte Gleichung folgt dabei aus Bemerkung IV.1.3, i). Für L_B^B geht man analog vor. □

Transformationsverhalten der Darstellungsmatrix unter Basiswechsel. — Die Darstellungsmatrix einer linearen Abbildung hängt von den gewählten Basen ab. Wir können den Zusammenhang der Darstellungsmatrizen einer linearen Abbildung zu verschiedenen Wahlen der Basen jedoch genau beschreiben.

Definition IV.1.26. Es seien V ein endlichdimensionaler K-Vektorraum und $B = \{b_1, ..., b_n\}$, $B' = \{b_1', ..., b_n'\}$ zwei durchnummerierte Basen von V. In der Notation von Satz III.5.32, i), setzen wir

$$T_{B'}^B := \Phi_{B'} \circ \Phi_B^{-1} \colon K^n \longrightarrow K^n,$$

und $S_{B'}^B \in \mathrm{Mat}(n;K)$ sei die eindeutig bestimmte $(n \times n)$-Matrix mit

$$f_{S_{B'}^B} = T_{B'}^B.$$

Bemerkung IV.1.27. In der Notation von Definition IV.1.21, i), ist $S_{B'}^B$ die Darstellungsmatrix der Identität bzgl. der Basen B und B'.

Eigenschaften IV.1.28. *Es seien V ein K-Vektorraum und B, B', B'' drei durchnummerierte Basen von V. Es gelten die Formeln:*

i) $T_{B''}^{B'} \circ T_{B'}^B = T_{B''}^B$.

ii) $T_B^{B'} = (T_{B'}^B)^{-1}$.

Beweis. i) Hier berechnen wir

$$\begin{aligned} T_{B''}^{B'} \circ T_{B'}^B &= (\Phi_{B''} \circ \Phi_{B'}^{-1}) \circ (\Phi_{B'} \circ \Phi_B^{-1}) = \Phi_{B''} \circ (\Phi_{B'}^{-1} \circ \Phi_{B'}) \circ \Phi_B^{-1} = \\ &= \Phi_{B''} \circ \mathrm{Id}_{K^n} \circ \Phi_B^{-1} = \Phi_{B''} \circ \Phi_B^{-1} = T_{B'}^B. \end{aligned}$$

ii) In diesem Fall erhalten wir

$$(T_{B'}^{B})^{-1} = (\Phi_{B'} \circ \Phi_{B}^{-1})^{-1} = (\Phi_{B}^{-1})^{-1} \circ \Phi_{B'}^{-1} = \Phi_{B} \circ \Phi_{B'}^{-1} = T_{B}^{B'}.$$

Dies ist die behauptete Gleichung. □

Satz IV.1.29. *Es seien V, W endlichdimensionale K-Vektorräume, B, B' durchnummerierte Basen von V, C, C' durchnummerierte Basen von W und $f\colon V \longrightarrow W$ eine lineare Abbildung. Dann sind die Darstellungsmatrizen von f bzgl. B und C und bzgl. B' und C' durch die Transformationsformel*

$$M_{C'}^{B'}(f) = S_{C'}^{C} \cdot M_{C}^{B}(f) \cdot (S_{B'}^{B})^{-1} = S_{C'}^{C} \cdot M_{C}^{B}(f) \cdot S_{B}^{B'}$$

verbunden.

Beweis. Vermöge Satz IV.1.25 lässt sich diese Aussage über Matrizen in eine Aussage über lineare Abbildungen übersetzen. Dazu berechnen wir

$$\Phi_{C'} \circ f \circ \Phi_{B'}^{-1} = (\Phi_{C'} \circ \Phi_{C}^{-1}) \circ (\Phi_{C} \circ f \circ \Phi_{B}^{-1}) \circ (\Phi_{B} \circ \Phi_{B'}^{-1}) = T_{C'}^{C} \circ (\Phi_{C} \circ f \circ \Phi_{B}^{-1}) \circ T_{B}^{B'}.$$

Dies entspricht der Formel

$$M_{C'}^{B'}(f) = S_{C'}^{C} \cdot M_{C}^{B}(f) \cdot S_{B}^{B'}.$$

Dazu wende man wie angekündigt Satz IV.1.25 an. □

Folgerung IV.1.30. *Es seien V, W Vektorräume über dem Körper K und $f\colon V \longrightarrow W$ eine lineare Abbildung. Dann existieren durchnummerierte Basen B bzw. C von V bzw. W, so dass*

$$M_{C}^{B}(f) = \left(\begin{array}{c|c} \mathbb{E}_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

Dabei ist r der Rang der Abbildung f, d.h. die Dimension des Bilds von f (Definition III.5.17).

Beweis. Es seien B' bzw. C' durchnummerierte Basen für V bzw. W. Gemäß Folgerung IV.1.11 existieren invertierbare Matrizen $S \in GL_m(K)$ und $T \in GL_n(K)$, so dass

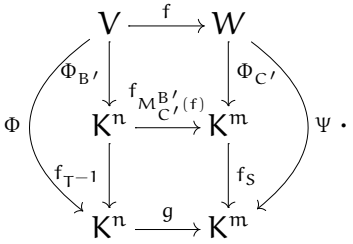
$$S \cdot M_{C'}^{B'}(f) \cdot T = \left(\begin{array}{c|c} \mathbb{E}_r & 0 \\ \hline 0 & 0 \end{array}\right). \tag{IV.2}$$

Dabei seien $m = \#C' = \mathrm{Dim}_K(W)$ und $n = \#B' = \mathrm{Dim}_K(V)$.

Mit

$$\Phi := f_{T^{-1}} \circ \Phi_{B'}, \quad \Psi := f_S \circ \Phi_{C'} \quad \text{und} \quad g := f_S \circ f_{M_{C'}^{B'}(f)} \circ f_T$$

kommutiert das Diagramm



Es existieren durchnummerierte Basen B bzw. C von V bzw. W, so dass $\Phi = \Phi_B$ und $\Psi = \Phi_C$ (Satz III.5.32, ii). Dementsprechend gilt

$$g = f_{M_C^B(f)}.$$

Ferner entnehmen wir dem Diagramm

$$g = \Phi_C \circ f \circ \Phi_B^{-1} = f_S \circ (\Phi_{C'} \circ f \circ \Phi_{B'}^{-1}) \circ (f_{T^{-1}})^{-1}.$$

Dazu gehört die Gleichung

$$M_C^B(f) = S \cdot M_{C'}^{B'}(f) \cdot (T^{-1})^{-1} = S \cdot M_{C'}^{B'}(f) \cdot T$$

von Matrizen. Mit (IV.2) ergibt sich die Behauptung. □

Beispiel IV.1.31. Es seien $V = \mathbb{Q}^3$, $W = \mathbb{Q}^2$, $B = \{e_1, e_2, e_3\}$,

$$B' = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\},$$

$C = \{e_1, e_2\}$, $C' = \left\{ \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$. Die Abbildung $\Phi_{B'}^{-1} \colon \mathbb{Q}^3 \longrightarrow \mathbb{Q}^3$ erfüllt

$$e_1 \longmapsto \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 \longmapsto \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad e_3 \longmapsto \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Weiter gilt $\Phi_B = \mathrm{Id}_{\mathbb{Q}^3}$. Damit haben wir

$$S_B^{B'} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Eine analoge Betrachtung liefert

$$S_C^{C'} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Es folgt

$$S_{C'}^{C} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Es seien

$$A := \begin{pmatrix} 1 & 3 & 2 \\ 0 & 1 & 2 \end{pmatrix}$$

und $f_A \colon \mathbb{Q}^3 \longrightarrow \mathbb{Q}^2$ die zugehörige lineare Abbildung. Dann haben wir $M_C^B(f_A) = A$ und

$$\begin{aligned} M_{C'}^{B'}(f_A) &= S_{C'}^{C} \cdot M_C^B(f_A) \cdot S_B^{B'} \\ &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 3 & 2 \\ 0 & 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -1 & -2 \\ 1 & 3 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -1 & -3 \\ 1 & 4 & 6 \end{pmatrix}. \end{aligned}$$

Die Matrix $M_{C'}^{B'}(f_A)$ gibt an, dass der Vektor $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ unter f_A auf $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ abgebildet wird, $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ auf

$$-\begin{pmatrix} 0 \\ -1 \end{pmatrix} + 4 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$$

und $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ auf

$$-3 \cdot \begin{pmatrix} 0 \\ -1 \end{pmatrix} + 6 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \end{pmatrix}.$$

In der Tat gilt

$$\begin{pmatrix} 1 & 3 & 2 \\ 0 & 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 3 & 2 \\ 0 & 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$$

und

$$\begin{pmatrix} 1 & 3 & 2 \\ 0 & 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \end{pmatrix}.$$

Aufgabe IV.1.32.
Es seien $A = (a_{ij})_{\substack{i=1,\dots,l \\ j=1,\dots,m}} \in \mathrm{Mat}(l,m;K)$, $B = (b_{jk})_{\substack{j=1,\dots,m \\ k=1,\dots,n}} \in \mathrm{Mat}(m,n;K)$ und $A \cdot B = (c_{ik})_{\substack{i=1,\dots,l \\ k=1,\dots,n}} \in \mathrm{Mat}(l,n;K)$ mit

$$c_{ik} := \sum_{j=1}^{m} a_{ij} \cdot b_{jk}, \quad i = 1,\dots,l,\ k = 1,\dots,n. \tag{IV.3}$$

Benutzen Sie die Formel (IV.3), um die folgenden Regeln für das Matrix-Produkt zu überprüfen. Für alle $A, A' \in \mathrm{Mat}(l,m;K)$, alle $B, B' \in \mathrm{Mat}(m,n;K)$, alle $C \in \mathrm{Mat}(n,p;K)$ und alle $\lambda \in K$ gilt:

1. $A \cdot (B + B') = A \cdot B + A \cdot B'$,
2. $(A + A') \cdot B = A \cdot B + A' \cdot B$,
3. $A \cdot (\lambda \cdot B) = (\lambda \cdot A) \cdot B = \lambda \cdot (A \cdot B)$ und
4. $(A \cdot B) \cdot C = A \cdot (B \cdot C)$.

Aufgabe IV.1.33.
Es sei $(R, +, \cdot)$ ein nicht notwendigerweise kommutativer Ring (s. Definition II.3.1, i). Ein Element $r \in R$ ist eine *Einheit*, wenn es ein Element $s \in R$ mit

$$r \cdot s = \mathbb{1} = s \cdot r$$

gibt.[2] Wir setzen

$$R^\star := \{\, r \in R \mid r \text{ ist eine Einheit} \,\}.$$

a) Zeigen Sie, dass

$$\forall r, s \in R^\star : r \cdot s \in R^\star.$$

b) Nach Teil a) ist die Multiplikation

$$\begin{aligned} \cdot\colon R^\star \times R^\star &\longrightarrow R^\star \\ (r,s) &\longmapsto r \cdot s \end{aligned}$$

definiert. Überprüfen Sie, dass $(R^\star, \cdot)$ eine Gruppe ist.

Aufgabe IV.1.34.
Es sei $n \geq 1$ eine natürliche Zahl. Gegeben seien zwei Matrizen $A, B \in \mathrm{Mat}(n;K)$. Zeigen Sie: Wenn $A \cdot B$ invertierbar ist, dann ist sowohl A als auch B invertierbar.

[2] Vorsicht. Wenn der Ring R nicht kommutativ ist, folgt aus $r \cdot s = \mathbb{1}$ nicht unbedingt $s \cdot r = \mathbb{1}$.

Aufgabe IV.1.35.
a) Bestimmen Sie Elementarmatrizen $B_1, ..., B_s \in \mathrm{Mat}(4;\mathbb{Q})$ und $C_1, ..., C_t \in \mathrm{Mat}(5;\mathbb{Q})$, für die

$$B_1 \cdot \cdots \cdot B_s \cdot A \cdot C_1 \cdot \cdots \cdot C_t = \left(\begin{array}{c|c} \mathbb{E}_{\mathrm{Rg}(A)} & 0 \\ \hline 0 & 0 \end{array}\right)$$

gilt. Dabei sei

$$A := \begin{pmatrix} 0 & 2 & 8 & -5 & 3 \\ -2 & 3 & 7 & -\frac{13}{2} & \frac{3}{2} \\ 2 & -3 & -6 & \frac{11}{2} & -\frac{1}{2} \\ 1 & -1 & -\frac{3}{2} & 2 & 0 \end{pmatrix}.$$

b) Zeigen Sie, dass die reelle Matrix

$$A := \begin{pmatrix} 0 & \frac{1}{2} & -2 & 0 \\ -2 & 0 & 1 & -\frac{7}{2} \\ 2 & -1 & 4 & -\frac{1}{2} \\ -4 & 1 & -3 & -2 \end{pmatrix}$$

invertierbar ist, bestimmen Sie A^{-1}, und schreiben Sie A als Produkt von Elementarmatrizen.

Aufgabe IV.1.36.
Es sei $K := \mathbb{F}_5$. Wir definieren

$$b_1 := \begin{pmatrix} \overline{0} \\ \overline{1} \\ \overline{2} \\ \overline{-1} \end{pmatrix}, \quad b_2 := \begin{pmatrix} \overline{1} \\ \overline{-1} \\ \overline{-3} \\ \overline{0} \end{pmatrix}, \quad b_3 := \begin{pmatrix} \overline{-1} \\ \overline{-1} \\ \overline{0} \\ \overline{2} \end{pmatrix} \quad \text{und} \quad b_4 := \begin{pmatrix} \overline{3} \\ \overline{8} \\ \overline{12} \\ \overline{-10} \end{pmatrix}$$

sowie

$$c_1 := \begin{pmatrix} \overline{1} \\ \overline{-1} \\ \overline{0} \end{pmatrix}, \quad c_2 := \begin{pmatrix} \overline{0} \\ \overline{1} \\ \overline{2} \end{pmatrix} \quad \text{und} \quad c_3 := \begin{pmatrix} \overline{-2} \\ \overline{1} \\ \overline{0} \end{pmatrix}.$$

a) Schreiben Sie die Vektoren $b_1, b_2, b_3, b_4, c_1, c_2$ und c_3 mit Einträgen der Form $\overline{a}$, $0 \leq a < 5$, und zeigen Sie, dass $B = \{b_1, b_2, b_3, b_4\}$ und $C = \{c_1, c_2, c_3\}$ durchnummerierte Basen für K^4 bzw. K^3 sind.
b) Es sei

$$A := \begin{pmatrix} \overline{1} & \overline{3} & \overline{0} & \overline{2} \\ \overline{0} & \overline{4} & \overline{2} & \overline{1} \\ \overline{2} & \overline{0} & \overline{4} & \overline{3} \end{pmatrix}.$$

Berechnen Sie $M_C^B(f_A)$. Schreiben Sie die Koeffizienten dabei wieder in der Form $\overline{a}$, $0 \leq a < 5$.

IV.2 Dualräume

Definitionen IV.2.1. i) Es sei V ein K-Vektorraum. Dann nennt man den K-Vektorraum

$$V^\vee := \mathrm{Hom}_K(V, K)$$

den *Dualraum* von V.

ii) Es seien V, W Vektorräume über dem Körper K und $f\colon V \longrightarrow W$ eine lineare Abbildung. Die Abbildung

$$\begin{array}{rcl} f^t\colon W^\vee & \longrightarrow & V^\vee \\ l & \longmapsto & l \circ f \end{array}$$

heißt die zu f *duale* oder *transponierte Abbildung* .

Lemma IV.2.2. i) *Es seien V, W zwei K-Vektorräume und $f\colon V \longrightarrow W$ eine lineare Abbildung. Dann ist die transponierte Abbildung $f^t\colon W^\vee \longrightarrow V^\vee$ ebenfalls linear.*

ii) *Für K-Vektorräume U, V, W und lineare Abbildungen $f\colon U \longrightarrow V$, $g\colon V \longrightarrow W$ gilt*

$$(g \circ f)^t = f^t \circ g^t.$$

iii) *Für K-Vektorräume und einen linearen* ***Isomorphismus*** *$f\colon V \longrightarrow W$ hat man*

$$(f^t)^{-1} = (f^{-1})^t.$$

Beweis. i) Es seien $\lambda \in K$ und $l, l'\colon W \longrightarrow K$ lineare Abbildungen. Man überprüft leicht (vgl. Beweis von Satz III.5.24)

$$(\lambda \cdot l) \circ f = \lambda \cdot (l \circ f) \quad \text{und} \quad (l + l') \circ f = l \circ f + l' \circ f.$$

ii) Für $l\colon W \longrightarrow K$ gilt

$$(g \circ f)^t(l) = l \circ (g \circ f) = (l \circ g) \circ f = f^t(l \circ g) = f^t\big(g^t(l)\big) = (f^t \circ g^t)(l).$$

iii) Offenbar hat man die Formel $\mathrm{Id}_V^t = \mathrm{Id}_{V^\vee}$. Somit findet man mit ii)

$$\mathrm{Id}_{V^\vee} = (\mathrm{Id}_V)^t = (f \circ f^{-1})^t = (f^{-1})^t \circ f^t.$$

Daraus ergibt sich die Behauptung. □

Es seien V ein endlichdimensionaler K-Vektorraum und $B = \{b_1, ..., b_n\}$ eine durchnummerierte Basis für V. Nach Satz III.5.21 ist

$$\begin{array}{rcl} \psi\colon V^\vee & \longrightarrow & \mathrm{Abb}(B, K) \\ l & \longmapsto & \begin{array}{rcl} l_{|B}\colon B & \longrightarrow & K \\ b & \longmapsto & l(b) \end{array} \end{array}$$

ein linearer Isomorphismus. Weiter bilden die Abbildungen

$$\begin{array}{rcl} f_i\colon B & \longrightarrow & K \\ b & \longmapsto & \begin{cases} 1, & b = b_i \\ 0, & \text{sonst} \end{cases}, \quad i = 1, \ldots, n, \end{array}$$

nach Beispiel III.4.3, iv), eine Basis für $\mathrm{Abb}(B, K)$. Somit ist $B^t = \{b^1, \ldots, b^n\}$ mit $b^i := \psi^{-1}(f_i)$, $i = 1, \ldots, n$, eine durchnummerierte Basis für $V^\vee$.

Bemerkung IV.2.3. Die Basis $B^t = \{b^1, \ldots, b^n\}$ ist durch die Bedingung

$$b^i(b_j) = \delta_{ij}, \quad i, j = 1, \ldots, n,$$

charakterisiert.

Definitionen IV.2.4. i) Die Basis B^t von $V^\vee$ ist die zu B *duale Basis* von $V^\vee$.

ii) Es sei $A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} \in \mathrm{Mat}(m, n; K)$ eine $(m \times n)$-Matrix. Die Matrix

$$A^t = (a_{ji})_{\substack{j=1,\ldots,n \\ i=1,\ldots,m}} \in \mathrm{Mat}(n, m; K)$$

ist die *transponierte Matrix* von A.

Beispiele IV.2.5. i) Für jede Matrix A hat man $(A^t)^t = A$.

ii) Für

$$A = \begin{pmatrix} 1 & 0 & -2 \\ 7 & 5 & -13 \\ -2 & 1 & 3 \end{pmatrix}$$

ergibt sich

$$A^t = \begin{pmatrix} 1 & 7 & -2 \\ 0 & 5 & 1 \\ -2 & -13 & 3 \end{pmatrix}.$$

iii) Die transponierte Matrix von

$$A = \begin{pmatrix} -1 & 2 & 5 & 0 \\ 2 & 7 & -10 & 8 \end{pmatrix}$$

ist

$$A^t = \begin{pmatrix} -1 & 2 \\ 2 & 7 \\ 5 & -10 \\ 0 & 8 \end{pmatrix}.$$

Satz IV.2.6. *Es seien V, W endlichdimensionale K-Vektorräume, B bzw. C durchnummerierte Basen von V bzw. W und B^t bzw. C^t die dazu dualen durchnummerierten Basen von $V^\vee$ bzw. $W^\vee$. Dann gilt*

$$M_{B^t}^{C^t}(f^t) = \big(M_C^B(f)\big)^t$$

für jede lineare Abbildung $f\colon V \longrightarrow W$.

Beweis. Es sei $M_C^B(f) = (a_{ij})_{\substack{i=1,\dots,m\\ j=1,\dots,n}}$. Für $j = 1, \dots, m$ und $i = 1, \dots, n$ berechnen wir

$$f^t(c^j)(b_i) = c^j\big(f(b_i)\big) = c^j\left(\sum_{k=1}^{m} a_{ki} \cdot c_k\right) = \sum_{k=1}^{m} a_{ki} \cdot c^j(c_k) = a_{ji}.$$

Somit erkennen wir

$$f^t(c^j) = \sum_{k=1}^{n} a_{jk} \cdot b^k.$$

Damit ist die j-te Spalte von $M_{B^t}^{C^t}(f^t)$ durch

$$\begin{pmatrix} a_{j1} \\ \vdots \\ a_{jn} \end{pmatrix}$$

gegeben, und wir erhalten die Behauptung. □

Folgerung IV.2.7. i) *Für natürliche Zahlen $l, m, n \geq 1$ und Matrizen $A \in \mathrm{Mat}(l, m; K)$, $B \in \mathrm{Mat}(m, n; K)$ hat man*

$$(A \cdot B)^t = B^t \cdot A^t.$$

ii) *Für eine natürliche Zahl $n \geq 1$ und eine **invertierbare** Matrix $A \in \mathrm{Mat}(n; K)$ gilt die Formel*

$$(A^t)^{-1} = (A^{-1})^t.$$

Eine weitere, interessante Anwendung von Satz IV.2.6 verbindet Eigenschaften einer linearen Abbildung und ihrer dualen. Der Rang einer Abbildung wurde dabei in Definition III.5.17 eingeführt.

Satz IV.2.8. *Für endlichdimensionale K-Vektorräume V, W und eine lineare Abbildung $f\colon V \longrightarrow W$ hat man*

$$\mathrm{Rg}(f^t) = \mathrm{Rg}(f).$$

Beweis. Wir wenden Folgerung IV.1.30 an. Es gibt durchnummerierte Basen B bzw. C von V bzw. W, so dass

$$M_C^B(f) = \left(\begin{array}{c|c} \mathbb{E}_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

Wie in Beispiel III.4.13 sieht man

$$\mathrm{Rg}(f) = r.$$

Weiter zeigt Satz IV.2.6[3]

$$M_{B^t}^{C^t}(f^t) = \left(M_C^B(f)\right)^t = \left(\begin{array}{c|c} \mathbb{E}_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

Hieran liest man $\mathrm{Rg}(f^t) = r$ ab. □

Bemerkung IV.2.9. Die Behauptung des Satzes ist äquivalent dazu, dass für natürliche Zahlen $m, n \geq 1$ und eine $(m \times n)$-Matrix $A \in \mathrm{Mat}(m, n; K)$

$$\text{S-Rg}(A) = \text{S-Rg}(A^t) = \text{Z-Rg}(A)$$

gilt. Damit haben wir einen weiteren Beweis für Folgerung III.5.41, iv), gefunden. Umgekehrt können wir den obigen Satz auch aus dieser Folgerung ableiten.

Aufgabe IV.2.10.
a) Zeigen Sie, dass die Vektoren

$$b_1 := \begin{pmatrix} -4 \\ 2{,}5 \\ -1 \end{pmatrix}, \quad b_2 := \begin{pmatrix} 0{,}5 \\ 0 \\ 4 \end{pmatrix} \quad \text{und} \quad b_3 := \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$$

eine Basis B' für $\mathbb{R}^3$ bilden.
b) Es seien $B = \{e_1, e_2, e_3\}$ bzw. und $C = \{1\}$ die Standardbasen für $\mathbb{R}^3$ bzw. $\mathbb{R}$. Es sei $B'^t = \{b^1, b^2, b^3\}$ die zu B' duale Basis von $\mathbb{R}^{3\vee}$. Bestimmen Sie die Darstellungsmatrizen $M_C^B(b^i)$, $i = 1, 2, 3$.

Aufgabe IV.2.11.
Es sei V ein nicht notwendigerweise endlichdimensionaler K-Vektorraum. Wir definieren

$$\begin{array}{rcl} \Phi\colon V & \longrightarrow & V^{\vee\vee} \\ v & \longmapsto & \varphi_v\colon V^\vee \longrightarrow K \\ & & \qquad l \longmapsto l(v) \end{array}.$$

[3]Falls $\mathrm{Dim}_K(V) \neq \mathrm{Dim}_K(W)$, haben die Matrizen $M_C^B(f)$ und $M_{B^t}^{C^t}(f^t)$ nicht dasselbe Format.

a) Zeigen Sie, dass Φ eine lineare Abbildung ist.
b) Beweisen Sie, dass Φ ein Isomorphismus ist, falls V endlichdimensional ist.
c) Es sei $V := \mathrm{Abb}'(\mathbb{N}, \mathbb{R})$ (s. Aufgabe II.3.7). Weisen Sie nach, dass es genau einen linearen Isomorphismus

$$\Psi\colon \mathrm{Abb}(\mathbb{N}, \mathbb{R}) \longrightarrow V^\vee$$

gibt, so dass

$$\forall i \in \mathbb{N} \forall (a_n)_{n\in\mathbb{N}} \in \mathrm{Abb}(\mathbb{N}, \mathbb{R}) :\ \Psi\big((a_n)_{n\in\mathbb{N}}\big)(e_i) = a_i.$$

Dabei ist e_i wie in Beispiel III.3.3, v), definiert, $i \in \mathbb{N}$.
d) In den Bezeichnungen von c) sei weiter $f = (f_i)_{i\in\mathbb{N}}$ die Folge mit $f_i = 1$, $i \in \mathbb{N}$. Zeigen Sie, dass die Teilmenge $F := \{f\} \cup \{\, e_i \mid i \in \mathbb{N}\,\}$ linear unabhängig ist.
e) Nehmen Sie an, es gebe eine Basis B, die die Teilmenge F aus Teil d) enthält.[4] Schließen Sie, dass unter dieser Voraussetzung die Abbildung Φ für den Vektorraum $V = \mathrm{Abb}'(\mathbb{N}, \mathbb{R})$ injektiv aber nicht surjektiv ist.

IV.3 Die symmetrische Gruppe

Definition IV.3.1. Es seien $n \geq 1$ eine natürliche Zahl und $A := \{1, ..., n\}$. Die Gruppe

$$\mathcal{S}_n := \mathcal{S}(A) := \Big\{\, \sigma\colon A \longrightarrow A \;\Big|\; \sigma \text{ ist bijektiv} \Big\}$$

heißt die *symmetrische Gruppe* von A. Ein Element $\sigma \in \mathcal{S}_n$ wird als *Permutation* bezeichnet.
Schreibweise. Eine Permutation $\sigma \in \mathcal{S}_n$ wird in der Form

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

aufgeschrieben.

Die Definition von $\mathcal{S}(A)$ ist in Bemerkung II.2.1 enthalten.

Beispiele IV.3.2. i) Die Permutation $\sigma\colon \{1,2,3\} \longrightarrow \{1,2,3\}$, $1 \longmapsto 1$, $2 \longmapsto 3$, $3 \longmapsto 2$, wird durch

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

dargestellt.

[4] Die Existenz einer solchen Basis lässt sich mit dem Zornschen Lemma ([10], S. 13) beweisen. (Max August Zorn (1906 - 1993), US-amerikanischer Mathematiker deutscher Abstammung)

ii) Beispiel II.1.24, ii), wird nun in folgender Form geschrieben:

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

Definitionen IV.3.3. i) Es seien $1 \leq i < j \leq n$ natürliche Zahlen. Die Permutation $\tau_{ij} \in \mathcal{S}_n$ mit

$$\tau_{ij}(k) = \begin{cases} j, & k = i \\ i, & k = j \\ k, & \text{sonst} \end{cases}$$

heißt die *Transposition* von i und j.

ii) Ein Permutation $\sigma \in \mathcal{S}_n$ ist eine *Transposition*, wenn es Indizes $1 \leq i < j \leq n$ mit $\sigma = \tau_{ij}$ gibt.

Bemerkungen und Beispiele IV.3.4. i) In der in Definition IV.3.1 eingeführten Schreibweise sieht eine Transposition folgendermaßen aus:

$$\tau_{ij} = \begin{pmatrix} 1 & 2 & \cdots & i-1 & i & i+1 & \cdots & j-1 & j & j+1 & \cdots & n \\ 1 & 2 & \cdots & i-1 & j & i+1 & \cdots & j-1 & i & j+1 & \cdots & n \end{pmatrix}, \quad 1 \leq i < j \leq n.$$

So hat man z.B.

$$\tau_{13} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix} \in \mathcal{S}_4.$$

ii) Für eine Transposition $\tau \in \mathcal{S}_n$ gilt $\tau^2 = e := \mathrm{Id}_{\{1,\ldots,n\}}$, so dass $\tau^{-1} = \tau$.

Lemma IV.3.5. i) *Die symmetrische Gruppe $\mathcal{S}_n$ ist genau dann abelsch, wenn $n \in \{1, 2\}$.*

ii) *Die Gruppe $\mathcal{S}_n$ umfasst $n!$ Elemente, $n \geq 1$.*

Beweis. i) Dass $\mathcal{S}_n$ für $n \geq 3$ nicht abelsch ist haben wir bereits in Bemerkung II.2.1 vermerkt. Für $n = 1$ bzw. $n = 2$ hat $\mathcal{S}_n$ ein bzw. zwei Elemente und ist daher abelsch.

ii) Die Aussage wird induktiv bewiesen. Für $n = 1$ und $n = 2$ ist sie richtig. Nun nehmen wir an, die Aussage sei für die natürliche Zahl n bekannt. Wir betrachten die Teilmengen

$$F_i := \left\{ \sigma \in \mathcal{S}_{n+1} \,|\, \sigma(n+1) = i \right\}, \quad i = 1, \ldots, n+1.$$

Offensichtlich gilt

$$\mathcal{S}_{n+1} = \bigsqcup_{i=1}^{n+1} F_i, \quad \text{so dass} \quad \#\mathcal{S}_{n+1} = \sum_{i=1}^{n+1} \#F_i.$$

Damit müssen wir noch zeigen:

Behauptung. $\#F_i = n!$, $i = 1, ..., n+1$.

Es gilt sicherlich $\#F_{n+1} = \#\mathcal{S}_n$, d.h. nach Induktionsvoraussetzung ist die Aussage für $i = n+1$ erfüllt. Die Abbildung

$$\begin{aligned} L_\sigma \colon \mathcal{S}_{n+1} &\longrightarrow \mathcal{S}_{n+1} \\ \rho &\longmapsto \sigma \cdot \rho \end{aligned}$$

ist für jede Permutation $\sigma \in \mathcal{S}_{n+1}$ bijektiv (Aufgabe II.2.10), und für $1 \leq i < j \leq n+1$ bildet $L_{\tau_{ij}}$ die Menge F_i auf F_j und die Menge F_j auf F_i ab. Wir folgern

$$\#F_1 = \cdots = \#F_{n+1} = n!,$$

und alles ist bewiesen. □

Satz IV.3.6. *Es sei $n \geq 2$ eine natürliche Zahl. Jede Permutation $\sigma \in \mathcal{S}_n$ lässt sich als Produkt von höchstens $n-1$ Transpositionen schreiben, d.h. es gibt eine natürliche Zahl $0 \leq k \leq n-1$ und Transpositionen $\tau_1, ..., \tau_k$, für die*

$$\sigma = \tau_1 \cdot \cdots \cdot \tau_k$$

gilt.[5]

Im Beweis verwenden wir „strukturerhaltende“ Abbildungen zwischen Gruppen.[6]

Definition IV.3.7. Es seien G, H Gruppen. Eine Abbildung $f \colon G \longrightarrow H$ ist ein (*Gruppen-*)*Homomorphismus*, wenn sie

$$\forall x, y \in G : \quad f(x \cdot y) = f(x) \cdot f(y)$$

erfüllt.[7]

Beispiele IV.3.8. i) Es seien V, W Vektorräume über dem Körper K und $f \colon V \longrightarrow W$ eine lineare Abbildung. Dann ist f ein Homomorphismus zwischen den entsprechenden additiven Gruppen $(V, +)$ und $(W, +)$.

ii) Es sei $n \geq 1$ eine natürliche Zahl. Dann ist

$$\begin{aligned} \iota_n \colon \mathcal{S}_n &\longrightarrow \mathcal{S}_{n+1} \\ \begin{pmatrix} 1 & \cdots & n \\ \sigma(1) & \cdots & \sigma(n) \end{pmatrix} &\longmapsto \begin{pmatrix} 1 & \cdots & n & n+1 \\ \sigma(1) & \cdots & \sigma(n) & n+1 \end{pmatrix} \end{aligned}$$

ein injektiver Gruppenhomomorphismus (Übung). In den Bezeichnungen des Beweises von Lemma IV.3.5, ii), bildet ι_n die Gruppe $\mathcal{S}_n$ bijektiv auf F_{n+1} ab.

[5] Das Produkt von null Transpositionen ist definitionsgemäß die Identität.

[6] Man erinnere sich, dass das entsprechende Konzept für Vektorräume dasjenige der linearen Abbildung (s. Definition III.5.1) ist.

[7] Dabei ist auf der linken Seite der Gleichung die Multiplikation in G und auf der rechten Seite diejenige in H gemeint.

Beobachtung IV.3.9. *Es seien* G, H *Gruppen und* $f\colon G \longrightarrow H$ *ein Gruppenhomomorphismus. Dann gilt:*

i) $f(e) = e$*;*

ii) $\forall x \in G$*:* $f(x^{-1}) = f(x)^{-1}$.

Beweis. i) Wir haben $f(e) = f(e \cdot e) = f(e) \cdot f(e)$. Durch Multiplikation dieser Gleichung mit $f(e)^{-1}$ gelangen wir zu $e = f(e)$.

ii) Hier beobachten wir $e = f(e) = f(x \cdot x^{-1}) = f(x) \cdot f(x^{-1})$, $x \in G$. □

Beweis von Satz IV.3.6. Wir beweisen die Aussage durch Induktion über n. In S_2 gibt es lediglich die Transposition τ_{12} und die Identität, welche Produkt von 0 Transpositionen ist.

Im Induktionsschritt $n \longrightarrow n+1$ verwenden wir den Homomorphismus ι_n aus Beispiel IV.3.8, ii). Er bildet Transpositionen auf Transpositionen ab:

$$\forall 1 \leq i < j \leq n : \quad \iota_n(\tau_{ij}) = \tau_{ij}.$$

Nun sei $\sigma \in S_{n+1}$ eine Permutation. Gilt $\sigma(n+1) = n+1$, so lässt sich σ nach Induktionsvoraussetzung als Produkt von höchstens $n-1$ Transpositionen schreiben. Andernfalls betrachten wir $\sigma' := \tau_{\sigma(n+1)n+1} \cdot \sigma$. Diese Permutation erfüllt $\sigma'(n+1) = n+1$. Es gibt daher eine natürliche Zahl $k \in \{0, ..., n-1\}$ und Transpositionen $\tau_1, ..., \tau_k$, so dass

$$\tau_{\sigma(n+1)n+1} \cdot \sigma = \sigma' = \tau_1 \cdot \dots \cdot \tau_k.$$

Aus dieser Gleichung folgt

$$\sigma = \tau_{\sigma(n+1)n+1} \cdot \tau_{\sigma(n+1)n+1} \cdot \sigma = \tau_{\sigma(n+1)\sigma} \cdot \tau_1 \cdot \dots \cdot \tau_k.$$

Damit ist alles gezeigt. □

Bemerkungen und Beispiele IV.3.10. i) Wir wenden das Verfahren aus dem Beweis auf die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix} \in S_4$$

an. Demnach bilden wir zuerst

$$\sigma' = \tau_{14} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix} = \tau_{12}.$$

Es folgt

$$\sigma = \tau_{14} \circ \tau_{12}.$$

Zur Probe berechnen wir

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

ii) Die Darstellung als Produkt von Transpositionen ist nicht eindeutig. Wir können ein gegebenes Produkt von Transpositionen z.B. mit $\tau_{12} \cdot \tau_{12}$ multiplizieren, ohne das Ergebnis zu ändern. Es gibt aber auch kompliziertere Phänomene. In $\mathcal{S}_4$ gilt z.B.

$$\begin{aligned} \tau_{12} \circ \tau_{23} \circ \tau_{12} &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix} = \tau_{13}. \end{aligned}$$

iii) Es seien V ein endlichdimensionaler K-Vektorraum und $B = \{b_1, ..., b_n\}$ eine durchnummerierte Basis von V. Durch

$$\begin{aligned} \varphi\colon \mathcal{S}_n &\longrightarrow \mathcal{S}(B) \\ \sigma &\longmapsto (b_i \longmapsto b_{\sigma(i)}, \quad i = 1, ..., n) \end{aligned}$$

ist ein bijektiver Gruppenhomomorphismus gegeben. Da $B \subset V$, können wir $\mathcal{S}_B$ als Teilmenge von $\mathrm{Abb}(B, V)$ auffassen. Wenden wir schließlich Ψ_B^{-1} aus Satz III.5.21 an, so erhalten wir eine Abbildung

$$\begin{aligned} \widetilde{\varphi}_B\colon \mathcal{S}_n &\longrightarrow \mathrm{End}_K(V) \\ \sigma &\longmapsto F_\sigma^B. \end{aligned}$$

Nach Konstruktion ist die lineare Abbildung $F_\sigma^B\colon V \longrightarrow V$ durch die Bedingung

$$\forall i \in \{1, ..., n\}\colon \quad F_\sigma^B(b_i) = b_{\sigma(i)}$$

charakterisiert. Man verifiziert nun:

- *Für jede Permutation $\sigma \in \mathcal{S}_n$ ist F_σ^B ein linearer Isomorphismus.*
- *Die resultierende Abbildung*

$$\begin{aligned} \varphi_B\colon \mathcal{S}_n &\longrightarrow \mathrm{GL}(V) \\ \sigma &\longmapsto F_\sigma^B \end{aligned}$$

ist ein injektiver Gruppenhomomorphismus.

Mit Satz IV.1.25 erhält man als direkte Folgerung:
Die Abbildung

$$\begin{aligned} \psi_B \colon \mathcal{S}_n &\longrightarrow \mathrm{GL}_n(K) \\ \sigma &\longmapsto M_B^B(F_\sigma^B) \end{aligned}$$

ist ein injektiver Gruppenhomomorphismus.
Man berechnet (in der Notation von Definition IV.3.1)

$$\forall 1 \leq i < j \leq n : \quad \psi_B(\tau_{ij}) = P_{ij}.$$

Satz IV.3.6 zeigt, dass ψ_B durch diese Bedingung eindeutig bestimmt ist. Es folgt, dass ψ_B nicht von V oder B abhängt. Die symmetrische Gruppe lässt sich somit als Untergruppe der allgemeinen linearen Gruppe auffassen.

Schreibt man eine Permutation als Produkt von Transpositionen, dann ist zwar die Anzahl der auftretenden Transpositionen nicht wohldefiniert, wohl aber deren Parität.[8] Das wird sich in Beobachtung IV.3.17 aus den allgemeinen Eigenschaften der Vorzeichenabbildung ergeben.

Definitionen IV.3.11. Es seien $n \geq 2$ eine natürliche Zahl und $\sigma \in \mathcal{S}_n$ eine Permutation.
i) Das *Vorzeichen* von σ ist definiert als

$$\mathrm{Sign}(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

ii) Ein *Fehlstand* von σ ist ein Paar (i, j) mit $1 \leq i < j \leq n$ und $\sigma(i) > \sigma(j)$.

Bemerkungen IV.3.12. i) Es seien $n \geq 2$, $\Delta = \{(i, i) \mid i \in \{1, ..., n\}\} \subset \{1, ..., n\} \times \{1, ..., n\}$ die Diagonale und

$$M := \big(\{1, ..., n\} \times \{1, ..., n\}\big) \setminus \Delta.$$

Dann ist

$$\begin{aligned} \sigma^{(2)} \colon M &\longrightarrow M \\ (i, j) &\longmapsto \big(\sigma(i), \sigma(j)\big) \end{aligned}$$

eine Bijektion. Weiter ist durch

$$(i, j) \sim (i', j') \quad :\Longleftrightarrow \quad (i = i' \wedge j = j') \vee (i = j' \wedge j = i')$$

[8] Die *Parität* einer natürlichen Zahl ist deren Eigenschaft, gerade oder ungerade zu sein.

eine Äquivalenzrelation gegeben. Die Äquivalenzklassen sind von der Form $\{(i,j),(j,i)\}$, $(i,j) \in M$. Es ist klar, dass das Bild einer Äquivalenzklasse unter $\sigma^{(2)}$ wieder eine Äquivalenzklasse ist. Da

$$R := \left\{ (i,j) \in M \,|\, i < j \right\}$$

ein vollständiges Repräsentantensystem für diese Äquivalenzrelation ist, ist auch $\sigma^{(2)}(R)$ eins. Diese langwierige Diskussion zeigt, dass

$$\prod_{1\le i<j\le n} \left|\sigma(j) - \sigma(i)\right| = \prod_{1\le i<j\le n} (j-i)$$

und somit

$$\mathrm{Sign}(\sigma) \in \{\pm 1\}.$$

ii) Zusammen mit der Verknüpfung

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

bildet die Menge $\{\pm 1\}$ eine Gruppe.

Lemma IV.3.13. *Es seien $n \ge 2$ eine natürliche Zahl, $\sigma \in \mathfrak{S}_n$ eine Permutation und*

$$m := \#\left\{ (i,j) \in \{1,\dots,n\} \times \{1,\dots,n\} \,\middle|\, (i,j) \text{ ist Fehlstand von } \sigma \right\}$$

die Anzahl der Fehlstände von σ. Dann gilt

$$\mathrm{Sign}(\sigma) = (-1)^m.$$

Beweis. Wir berechnen

$$\begin{aligned} \prod_{1\le i<j\le n} \frac{\sigma(j)-\sigma(i)}{j-i} &= (-1)^m \cdot \prod_{1\le i<j\le n} \frac{\left|\sigma(j)-\sigma(i)\right|}{j-i} \\ &= (-1)^m \cdot \frac{\displaystyle\prod_{1\le i<j\le n} \left|\sigma(j)-\sigma(i)\right|}{\displaystyle\prod_{i\le i<j\le n} (j-i)} = (-1)^m. \end{aligned}$$

Bei der letzten Umformung haben wir Bemerkung IV.3.12, i), verwandt. □

Satz IV.3.14. *Es sei $n \ge 2$ eine natürliche Zahl. Dann ist die Abbildung*

$$\begin{aligned} \mathrm{Sign}\colon \mathfrak{S}_n &\longrightarrow \{\pm 1\} \\ \sigma &\longmapsto \mathrm{Sign}(\sigma) \end{aligned}$$

ein Gruppenhomomorphismus.

Beweis. Für Permutationen $\sigma, \sigma' \in \mathcal{S}_n$ erhalten wir

$$\begin{aligned}\mathrm{Sign}(\sigma \cdot \sigma') &= \prod_{1\le i<j\le n} \frac{\sigma\big(\sigma'(j)\big) - \sigma\big(\sigma'(i)\big)}{j-i} = \\ &= \prod_{1\le i<j\le n} \frac{\sigma\big(\sigma'(j)\big) - \sigma\big(\sigma'(i)\big)}{\sigma'(j)-\sigma'(i)} \cdot \prod_{1\le i<j\le n} \frac{\sigma'(j)-\sigma'(i)}{j-i}.\end{aligned}$$

Das zweite Produkt ergibt nach Definition IV.3.11, i), $\mathrm{Sign}(\sigma')$. Die Gleichung

$$\prod_{1\le i<j\le n} \frac{\sigma\big(\sigma'(j)\big) - \sigma\big(\sigma'(i)\big)}{\sigma'(j)-\sigma'(i)} = \prod_{1\le i<j\le n} \frac{\sigma(j)-\sigma(i)}{j-i} = \mathrm{Sign}(\sigma)$$

ergibt sich wieder aus Bemerkung IV.3.12, i). □

Satz IV.3.15. *Es seien $n \ge 2$ eine natürliche Zahl und $\tau \in \mathcal{S}_n$ eine **Transposition**. Dann gilt*

$$\mathrm{Sign}(\tau) = -1.$$

Beweis. Für die Transposition τ_{12} folgt das sofort aus Lemma IV.3.13. Für $1 \le i < j \le n$ wählen wir eine Permutation $\sigma \in \mathcal{S}_n$ mit $\sigma(1) = i$ und $\sigma(2) = j$. Man überprüft

$$\tau_{ij} = \sigma \cdot \tau_{12} \cdot \sigma^{-1}.$$

Mit Satz IV.3.14 folgt

$$\begin{aligned}\mathrm{Sign}(\tau_{ij}) &= \mathrm{Sign}(\sigma)\cdot\mathrm{Sign}(\tau_{12})\cdot\mathrm{Sign}(\sigma^{-1}) = (-1)\cdot\mathrm{Sign}(\sigma)\cdot\mathrm{Sign}(\sigma^{-1}) = \\ &= (-1)\cdot\mathrm{Sign}(\sigma\cdot\sigma^{-1}) = (-1)\cdot\mathrm{Sign}(e) = -1.\end{aligned}$$

Damit ist alles gezeigt. □

Definitionen IV.3.16. Es sei $n \ge 2$ eine natürliche Zahl.

i) Eine Permutation $\sigma \in \mathcal{S}_n$ heißt *gerade* bzw. *ungerade*, wenn $\mathrm{Sign}(\sigma) = 1$ bzw. $\mathrm{Sign}(\sigma) = -1$.

ii) Wir setzen

$$A_n := \big\{\, \sigma \in \mathcal{S}_n \,|\, \sigma \text{ gerade} \,\big\} \quad \text{und} \quad U_n := \big\{\, \sigma \in \mathcal{S}_n \,|\, \sigma \text{ ungerade} \,\big\}.$$

Beobachtung IV.3.17. *Es seien $n \ge 2$ eine natürliche Zahl, $\tau_1, \ldots, \tau_k \in \mathcal{S}_n$ Transpositionen und $\sigma = \tau_1 \cdot \cdots \cdot \tau_k$. Dann ist σ genau dann gerade bzw. ungerade, wenn k gerade bzw. ungerade ist.*

Beweis. Nach Satz IV.3.14 und IV.3.15 gilt

$$\mathrm{Sign}(\sigma) = (-1)^k.$$

Dies impliziert die Behauptung. □

Lemma IV.3.18. *Es sei $n \geq 2$ eine natürliche Zahl. Dann ist A_n eine Untergruppe von S_n mit $n!/2$ Elementen.*

Beweis. Die Identität $e \in S_n$ ist offenbar eine gerade Permutation. Es sei $\sigma \in A_n$. Mit Beobachtung IV.3.9, ii), und Satz IV.3.14 berechnen wir $\mathrm{Sign}(\sigma^{-1}) = \mathrm{Sign}(\sigma)^{-1} = 1^{-1} = 1$. Es folgt $\sigma^{-1} \in A_n$. Schließlich seien $\sigma, \sigma' \in A_n$ zwei gerade Permutationen. Wir haben $\mathrm{Sign}(\sigma \cdot \sigma') = \mathrm{Sign}(\sigma) \cdot \mathrm{Sign}(\sigma') = 1 \cdot 1 = 1$, so dass $\sigma \cdot \sigma' \in A_n$.

Nach Definition gilt

$$S_n = A_n \sqcup U_n,$$

so dass insbesondere

$$n! = \#S_n = \#A_n + \#U_n.$$

Es sei $\tau := \tau_{12}$. Nach Aufgabe II.2.10 ist die Abbildung

$$\begin{array}{rccc} L_\tau \colon & S_n & \longrightarrow & S_n \\ & \sigma & \longmapsto & \tau \circ \sigma \end{array}$$

bijektiv. Sie bildet A_n auf U_n ab. Folglich gilt

$$\#A_n = \#U_n = \frac{\#S_n}{2},$$

und der zweite Teil der Behauptung ist ebenfalls bewiesen. □

Nach Bemerkung II.2.8, i), ist A_n selber eine Gruppe.

Definition IV.3.19. Es sei $n \geq 2$ eine natürliche Zahl. Die Gruppe A_n wird *alternierende Gruppe* genannt.

Beobachtung IV.3.20. *Die alternierende Gruppe A_n ist genau dann abelsch, wenn $n \in \{2, 3\}$.*

Beweis. Die Gruppe A_2 umfasst lediglich ein Element und ist damit trivialerweise abelsch. Nach Lemma IV.3.18 gilt $\#A_3 = 3$. Die Betrachtungen in Bemerkung II.2.6, iii), zeigen, dass jede Gruppe mit drei Elementen abelsch ist. Schließlich sei $n \geq 4$. Der in Beispiel IV.3.8, ii), betrachtete injektive Gruppenhomomorphismus

$$\iota_n \colon S_n \longrightarrow S_{n+1}$$

bildet A_n nach A_{n+1} ab. Es genügt daher nachzuweisen, dass A_4 nicht abelsch ist. Die Elemente

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 3 \end{pmatrix}$$

sind als Produkte von jeweils zwei Transpositionen gerade, d.h. in A_4 enthalten. Wir haben

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix},$$
$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}.$$

Die beiden angegebenen Elemente kommutieren also nicht miteinander. □

Aufgabe IV.3.21.
a) Berechnen Sie die Produkte σ_1^2, σ_2^2, $\sigma_1\sigma_2$ und $\sigma_2\sigma_1$ für

$$\sigma_1 := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 1 & 3 & 7 & 6 & 5 & 4 \end{pmatrix} \text{ und } \sigma_2 := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 3 & 2 & 5 & 6 & 4 & 7 \end{pmatrix}.$$

b) Stellen Sie die Permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 4 & 6 & 3 & 2 & 1 \end{pmatrix}$$

als Produkt von Transpositionen dar.

IV.4 Determinanten

Bereits in Abschnitt I.1 haben wir mit der Determinante einer (2×2)-Matrix gearbeitet. Jetzt möchten wir das analoge Konzept für $(n \times n)$-Matrizen für beliebige natürliche Zahlen $n \geq 1$ entwickeln.
Schreibweise. Es seien

$$z_i = (\ a_{i1} \ \cdots \ a_{in}\ ) \in ZK^n, \quad i = 1, ..., n,$$

Zeilenvektoren. Dann sei

$$\begin{pmatrix} \frac{z_1}{\vdots} \\ \frac{}{z_n} \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \in \mathrm{Mat}(n; K)$$

die $(n \times n)$-Matrix mit den Zeilen $z_1, ..., z_n$.

Definitionen IV.4.1. Es sei $n \geq 1$ eine natürliche Zahl.
i) Eine Abbildung

$$\mathrm{Det}\colon \mathrm{Mat}(n; K) \longrightarrow K$$

ist eine *Determinantenabbildung*, wenn sie folgende Eigenschaften aufweist:

1. Für $j \in \{1, ..., n\}$ ist Det **linear in der j-ten Zeile**, d.h. für $\lambda \in K$ und Zeilenvektoren $z_1, ..., z_n, z_j', z_j'' \in ZK^n$ hat man:

$$\text{a) } \operatorname{Det}\begin{pmatrix} z_1 \\ \vdots \\ z_{j-1} \\ \lambda \cdot z_j \\ z_{j+1} \\ \vdots \\ z_n \end{pmatrix} = \lambda \cdot \operatorname{Det}\begin{pmatrix} z_1 \\ \vdots \\ z_{j-1} \\ z_j \\ z_{j+1} \\ \vdots \\ z_n \end{pmatrix};$$

$$\text{b) } \operatorname{Det}\begin{pmatrix} z_1 \\ \vdots \\ z_{j-1} \\ z_j' + z_j'' \\ z_{j+1} \\ \vdots \\ z_n \end{pmatrix} = \operatorname{Det}\begin{pmatrix} z_1 \\ \vdots \\ z_{j-1} \\ z_j' \\ z_{j+1} \\ \vdots \\ z_n \end{pmatrix} + \operatorname{Det}\begin{pmatrix} z_1 \\ \vdots \\ z_{j-1} \\ z_j'' \\ z_{j+1} \\ \vdots \\ z_n \end{pmatrix}.$$

2. Die Abbildung Det ist *alternierend*, d.h. hat $A \in \operatorname{Mat}(n; K)$ zwei identische Zeilen, dann folgt $\operatorname{Det}(A) = 0$.

3. Die Abbildung Det ist *normiert*, d.h. $\operatorname{Det}(\mathbb{E}_n) = 1$.

ii) Es seien $\operatorname{Det}\colon \operatorname{Mat}(n; K) \longrightarrow K$ eine Determinantenabbildung und $A \in \operatorname{Mat}(n; K)$ eine $(n \times n)$-Matrix. Dann heißt die Zahl $\operatorname{Det}(A) \in K$ die *Determinante* von A.

Die folgenden Betrachtungen werden zeigen, dass man ausgehend von den obigen Axiomen die Determinante jeder Matrix berechnen kann. Daraus folgt erst einmal, dass es höchstens eine Determinantenabbildung gibt. Später (Satz IV.4.6) werden wir nachweisen, dass eine Determinantenabbildung existiert.

Satz IV.4.2. *Es seien $n \geq 1$ eine natürliche Zahl und $\operatorname{Det}\colon \operatorname{Mat}(n; K) \longrightarrow K$ eine Determinantenabbildung. Gegeben seien Zeilenvektoren $z_i = (\, a_{i1} \ \cdots \ a_{in} \,) \in ZK^n$, $i = 1, ..., n$, und damit die $(n \times n)$-Matrix*

$$A = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}.$$

i) *Zeilenoperationen vom Typ* (ZI) *ändern die Determinante nicht, d.h. für gegebene Indizes* $1 \leq i \neq j \leq n$ *und eine Zahl* $\lambda \in K$ *gilt*

$$\mathrm{Det}(A) = \mathrm{Det}\begin{pmatrix} z_1 \\ \hline \vdots \\ \hline z_{j-1} \\ \hline z_j + \lambda \cdot z_i \\ \hline z_{j+1} \\ \hline \vdots \\ \hline z_n \end{pmatrix}.$$

ii) *Bei Zeilenvertauschungen ändert sich das Vorzeichen der Determinante, d.h. für Indizes* $1 \leq i < j \leq n$ *hat man*

$$\mathrm{Det}(A) = -\mathrm{Det}\begin{pmatrix} z_1 \\ \hline \vdots \\ \hline z_{i-1} \\ \hline z_j \\ \hline z_{i+1} \\ \hline \vdots \\ \hline z_{j-1} \\ \hline z_i \\ \hline z_{j+1} \\ \hline \vdots \\ \hline z_n \end{pmatrix}.$$

iii) *Gibt es einen Index* $i \in \{1, ..., n\}$ *mit* $z_i = 0$*, so folgt* $\mathrm{Det}(A) = 0$*.*

Beweis. i) Hier berechnet man mit Eigenschaft 1) und 2) aus Definition IV.4.1:

$$\mathrm{Det}\begin{pmatrix} z_1 \\ \hline \vdots \\ \hline z_{j-1} \\ \hline z_j + \lambda \cdot z_i \\ \hline z_{j+1} \\ \hline \vdots \\ \hline z_n \end{pmatrix} = \mathrm{Det}\begin{pmatrix} z_1 \\ \hline \vdots \\ \hline \vdots \\ \hline \vdots \\ \hline z_n \end{pmatrix} + \mathrm{Det}\begin{pmatrix} z_1 \\ \hline \vdots \\ \hline z_{j-1} \\ \hline \lambda \cdot z_i \\ \hline z_{j+1} \\ \hline \vdots \\ \hline z_n \end{pmatrix} =$$

$$= \operatorname{Det}(A) + \lambda \cdot \underbrace{\operatorname{Det}\begin{pmatrix} z_1 \\ \vdots \\ z_{j-1} \\ z_i \\ z_{j+1} \\ \vdots \\ z_n \end{pmatrix}}_{=0,\ \text{Det alternierend}} = \operatorname{Det}(A).$$

iii) Wir wählen einen Index $j \neq i$ und wenden Teil i) an:

$$\operatorname{Det}\begin{pmatrix} z_1 \\ \vdots \\ z_{i-1} \\ 0 \\ z_{i+1} \\ \vdots \\ z_n \end{pmatrix} = \operatorname{Det}\begin{pmatrix} z_1 \\ \vdots \\ z_{i-1} \\ z_j \\ z_{i+1} \\ \vdots \\ z_n \end{pmatrix}.$$

Die Determinante auf der rechten Seite verschwindet, weil die Matrix zwei identische Zeilen hat (Definition IV.4.1, 2).

ii) Mit Teil iii) und den Eigenschaften 1) und 2) aus Definition IV.4.1 finden wir

$$0 = \operatorname{Det}\begin{pmatrix} z_1 \\ \vdots \\ z_{i-1} \\ z_i + z_j \\ z_{i+1} \\ \vdots \\ z_{j-1} \\ z_i + z_j \\ z_{j+1} \\ \vdots \\ z_n \end{pmatrix} = \operatorname{Det}\begin{pmatrix} z_1 \\ \vdots \\ z_{i-1} \\ z_i \\ z_{i+1} \\ \vdots \\ z_{j-1} \\ z_i + z_j \\ z_{j+1} \\ \vdots \\ z_n \end{pmatrix} + \operatorname{Det}\begin{pmatrix} z_1 \\ \vdots \\ z_{i-1} \\ z_j \\ z_{i+1} \\ \vdots \\ z_{j-1} \\ z_i + z_j \\ z_{j+1} \\ \vdots \\ z_n \end{pmatrix} =$$

$$= \underbrace{\operatorname{Det}\begin{pmatrix} z_1 \\ \vdots \\ z_{i-1} \\ z_i \\ z_{i+1} \\ \vdots \\ z_{j-1} \\ z_i \\ z_{j+1} \\ \vdots \\ z_n \end{pmatrix}}_{=0} + \underbrace{\operatorname{Det}\begin{pmatrix} z_1 \\ \vdots \\ z_{i-1} \\ z_i \\ z_{i+1} \\ \vdots \\ z_{j-1} \\ z_j \\ z_{j+1} \\ \vdots \\ z_n \end{pmatrix}}_{=\operatorname{Det}(A)} + \operatorname{Det}\begin{pmatrix} z_1 \\ \vdots \\ z_{i-1} \\ z_j \\ z_{i+1} \\ \vdots \\ z_{j-1} \\ z_i \\ z_{j+1} \\ \vdots \\ z_n \end{pmatrix} + \underbrace{\operatorname{Det}\begin{pmatrix} z_1 \\ \vdots \\ z_{i-1} \\ z_j \\ z_{i+1} \\ \vdots \\ z_{j-1} \\ z_j \\ z_{j+1} \\ \vdots \\ z_n \end{pmatrix}}_{=0}.$$

Daraus folgt die Behauptung. □

Mit diesem Satz können wir die Determinanten der Elementarmatrizen bestimmen.

Beispiel IV.4.3. Es seien $1 \leq i \neq j \leq n$ Indizes und $\lambda \in K$ eine Zahl. In den Bezeichnungen von Definition IV.1.9 ergibt sich mit Satz IV.4.2

$$\begin{array}{lrcrcl} a) & \operatorname{Det}\big(Q_{ij}(\lambda)\big) & = & \operatorname{Det}(\mathbb{E}_n) & = & 1, \\ b) & \operatorname{Det}(P_{ij}) & = & -\operatorname{Det}(\mathbb{E}_n) & = & -1, \\ c) & \operatorname{Det}\big(S_i(\lambda)\big) & = & \lambda \cdot \operatorname{Det}(\mathbb{E}_n) & = & \lambda. \end{array}$$

Dieses Beispiel kombinieren wir mit den Ergebnissen zu Elementarmatrizen aus Satz IV.1.10 und der folgenden Diskussion.

Satz IV.4.4. *Es sei $n \geq 1$ eine natürliche Zahl.*

i) *Für eine Determinantenabbildung* $\operatorname{Det}\colon \operatorname{Mat}(n;K) \longrightarrow K$ *und eine* $(n \times n)$*-Matrix* $A \in \operatorname{Mat}(n;K)$ *sind folgende Bedingungen äquivalent:*

a) $\operatorname{Det}(A) = 0$,

b) $\operatorname{Rg}(A) < n$.

ii) *Für eine Determinantenabbildung* $\operatorname{Det}\colon \operatorname{Mat}(n;K) \longrightarrow K$ *und* $(n \times n)$*-Matrizen* $A, B \in \operatorname{Mat}(n;K)$ *ist die Gleichung*

$$\operatorname{Det}(A \cdot B) = \operatorname{Det}(A) \cdot \operatorname{Det}(B)$$

erfüllt.

Beweis. Wir beginnnen mit dem Beweis der folgenden

Behauptung. *Für eine Determinantenabbildung* $\mathrm{Det}\colon \mathrm{Mat}(n;K) \longrightarrow K$, *eine* $(n \times n)$*-Matrix* $A \in \mathrm{Mat}(n;K)$ *und eine Elementarmatrix* T *gilt*

$$\mathrm{Det}(T \cdot A) = \mathrm{Det}(T) \cdot \mathrm{Det}(A).$$

Wir verwenden den Zusammenhang zwischen Zeilenoperationen und Multiplikation mit Elementarmatrizen (Satz IV.1.10), Definition IV.4.1, 1), Satz IV.4.2 und Beispiel IV.4.3.

Für Indizes $i \neq j \in \{1, ..., n\}$ und $\lambda \in K$ gilt

$$\mathrm{Det}\big(Q_{ij}(\lambda) \cdot A\big) = \mathrm{Det}(A) = \mathrm{Det}\big(Q_{ij}(\lambda)\big) \cdot \mathrm{Det}(A).$$

Für Indizes $i \neq j \in \{1, ..., n\}$ hat man

$$\mathrm{Det}(P_{ij} \cdot A) = -\mathrm{Det}(A) = \mathrm{Det}(P_{ij}) \cdot \mathrm{Det}(A).$$

Für einen Index $i \in \{1, ..., n\}$ und $\lambda \in K$ finden wir

$$\mathrm{Det}\big(S_i(\lambda) \cdot A\big) = \lambda \cdot \mathrm{Det}(A) = \mathrm{Det}\big(S_i(\lambda)\big) \cdot \mathrm{Det}(A).$$

Jetzt beweisen wir Teil i). Wir betrachten zunächst eine $(n \times n)$-Matrix A mit $\mathrm{Rg}(A) < n$. Nach Satz I.3.3 und IV.1.10 gibt es Elementarmatrizen $C_1, ..., C_k \in \mathrm{Mat}(n;K)$, so dass sich die Matrix

$$A' = C_1 \cdot \cdots \cdot C_k \cdot A$$

in Zeilenstufenform befindet. Dabei ist $r = \mathrm{Rg}(A') = \mathrm{Rg}(A)$ (Beispiel III.4.13 und Satz III.4.14) die Anzahl der Stufen. Wegen $r < n$ enthält A' eine Nullzeile. Satz IV.4.2, iii), besagt $\mathrm{Det}(A') = 0$. Mit der Behauptung finden wir

$$0 = \mathrm{Det}(A') = \underbrace{\mathrm{Det}(C_1) \cdot \cdots \cdot \mathrm{Det}(C_k)}_{\neq 0,\ \text{Beispiel IV.4.3}} \cdot \mathrm{Det}(A).$$

Nun widmen wir uns dem Fall einer $(n \times n)$-Matrix A mit $\mathrm{Rg}(A) = n$. Hier gibt es Elementarmatrizen $D_1, ..., D_u \in \mathrm{Mat}(n;K)$ mit $A = D_1 \cdot \cdots \cdot D_u$ (Satz IV.1.17, i), und die Behauptung liefert

$$\mathrm{Det}(A) = \mathrm{Det}(D_1) \cdot \cdots \cdot \mathrm{Det}(D_u) \neq 0.$$

ii) Wenn A Rang n hat, dann wählen wir Elementarmatrizen $D_1, ..., D_u \in \mathrm{Mat}(n;K)$ mit $A = D_1 \cdot \cdots \cdot D_u$ und berechnen unter Verwendung der Behauptung

$$\mathrm{Det}(A \cdot B) = \mathrm{Det}(D_1 \cdot \cdots \cdot D_u \cdot B) = \underbrace{\mathrm{Det}(D_1) \cdot \cdots \cdot \mathrm{Det}(D_u)}_{=\mathrm{Det}(A)} \cdot \mathrm{Det}(B).$$

Gilt $\mathrm{Rg}(A) < n$, dann haben wir auch $\mathrm{Rg}(A \cdot B) < n$ (Warum?). Mit i) folgt $0 = \mathrm{Det}(A \cdot B) = 0 \cdot \mathrm{Det}(B) = \mathrm{Det}(A) \cdot \mathrm{Det}(B)$. □

Folgerung IV.4.5. *Es seien $n \geq 1$ eine natürliche Zahl.*
i) *Es gibt höchstens eine Determinantenabbildung* $\mathrm{Det}\colon \mathrm{Mat}(n;K) \longrightarrow K$.
ii) *Für eine Determinantenabbildung* $\mathrm{Det}\colon \mathrm{Mat}(n;K) \longrightarrow K$ *gilt*

$$\mathrm{GL}_n(K) = \left\{ A \in \mathrm{Mat}(n;K) \,|\, \mathrm{Det}(A) \neq 0 \right\},$$

und

$$\mathrm{Det}\colon \mathrm{GL}_n(K) \longrightarrow K \setminus \{0\}$$

ist ein Gruppenhomomorphismus.

Beweis. Die obigen Betrachtungen beinhalten, dass eine Determinantenabbildung durch ihre Werte auf den Elementarmatrizen bestimmt ist. Diese haben wir in Beispiel IV.4.3 ermittelt. Dies zeigt Teil i). Teil ii) folgt unmittelbar aus Satz IV.4.4. □

Jetzt beschäftigen wir uns mit der Existenz von Determinantenabbildungen.

Satz IV.4.6. *Es sei $n \geq 1$ eine natürliche Zahl. Durch die* ***Leibnizformel***[9]

$$\begin{array}{rcl} \mathrm{Mat}(n;K) & \longrightarrow & K \\ A = (a_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}} & \longmapsto & \displaystyle\sum_{\sigma \in S_n} \mathrm{Sign}(\sigma) \cdot a_{1\sigma(1)} \cdot \cdots \cdot a_{n\sigma(n)} \end{array}$$

ist eine Determinantenabbildung definiert.

Beweis. Wir verifizieren die Axiome aus Definition IV.4.1, i). Wir beginnnen mit 1, a). Es seien $i \in \{1, \ldots, n\}$ ein Index und $\lambda \in K$ eine Zahl. Wir finden

$$\begin{aligned} \mathrm{Det}\begin{pmatrix} z_1 \\ \hline \vdots \\ \hline z_{i-1} \\ \hline \lambda \cdot z_i \\ \hline z_{i+1} \\ \hline \vdots \\ \hline z_n \end{pmatrix} &= \sum_{\sigma \in S_n} \mathrm{Sign}(\sigma) \cdot a_{1\sigma(1)} \cdot \cdots \cdot a_{i-1\sigma(i-1)} \cdot (\lambda \cdot a_{i\sigma(i)}) \cdot a_{i+1\sigma(i+1)} \cdot \cdots \cdot a_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} \lambda \cdot \mathrm{Sign}(\sigma) \cdot a_{1\sigma(1)} \cdot \cdots \cdot a_{i-1\sigma(i-1)} \cdot a_{i\sigma(i)} \cdot a_{i+1\sigma(i+1)} \cdot \cdots \cdot a_{n\sigma(n)} \\ &= \lambda \cdot \sum_{\sigma \in S_n} \mathrm{Sign}(\sigma) \cdot a_{1\sigma(1)} \cdot \cdots \cdot a_{i-1\sigma(i-1)} \cdot a_{i\sigma(i)} \cdot a_{i+1\sigma(i+1)} \cdot \cdots \cdot a_{n\sigma(n)} \\ &= \lambda \cdot \mathrm{Det}(A). \end{aligned}$$

[9] Gottfried Wilhelm Leibniz (1646 - 1716), deutscher Philosoph, Mathematiker uvm.

Für 1, b), seien $i \in \{1,...,n\}$, $z_i = (\ a_{i1}\ \cdots\ a_{in}\ )$, $z_i' = (\ a_{i1}'\ \cdots\ a_{in}'\ )$, $z_i'' = (\ a_{i1}''\ \cdots\ a_{in}''\ )$, Zeilenvektoren, so dass $z_i = z_i' + z_i''$, d.h. $a_{ij} = a_{ij}' + a_{ij}''$, $j = 1,...,n$. Mit der Leibniz-Formel berechnen wir

$$\begin{aligned}
&\mathrm{Det}\begin{pmatrix} \frac{z_1}{\vdots} \\ \hline z_n \end{pmatrix} \\
&= \sum_{\sigma\in S_n} \mathrm{Sign}(\sigma)\cdot a_{1\sigma(1)}\cdot\dots\cdot a_{i-1\sigma(i-1)}\cdot a_{i\sigma(i)}\cdot a_{i+1\sigma(i+1)}\cdot\dots\cdot a_{n\sigma(n)} \\
&= \sum_{\sigma\in S_n} \mathrm{Sign}(\sigma)\cdot a_{1\sigma(1)}\cdot\dots\cdot a_{i-1\sigma(i-1)}\cdot (a_{i\sigma(i)}' + a_{i\sigma(i)}'')\cdot a_{i+1\sigma(i+1)}\cdot\dots\cdot a_{n\sigma(n)} \\
&= \sum_{\sigma\in S_n} \Big(\mathrm{Sign}(\sigma)\cdot a_{1\sigma(1)}\cdot\dots\cdot a_{i-1\sigma(i-1)}\cdot a_{i\sigma(i)}'\cdot a_{i+1\sigma(i+1)}\cdot\dots\cdot a_{n\sigma(n)} + \\
&\qquad +\mathrm{Sign}(\sigma)\cdot a_{1\sigma(1)}\cdot\dots\cdot a_{i-1\sigma(i-1)}\cdot a_{i\sigma(i)}''\cdot a_{i+1\sigma(i+1)}\cdot\dots\cdot a_{n\sigma(n)}\Big) \\
&= \Big(\sum_{\sigma\in S_n} \mathrm{Sign}(\sigma)\cdot a_{1\sigma(1)}\cdot\dots\cdot a_{i-1\sigma(i-1)}\cdot a_{i\sigma(i)}'\cdot a_{i+1\sigma(i+1)}\cdot\dots\cdot a_{n\sigma(n)}\Big) + \\
&\quad +\Big(\sum_{\sigma\in S_n} \mathrm{Sign}(\sigma)\cdot a_{1\sigma(1)}\cdot\dots\cdot a_{i-1\sigma(i-1)}\cdot a_{i\sigma(i)}''\cdot a_{i+1\sigma(i+1)}\cdot\dots\cdot a_{n\sigma(n)}\Big) \\
&= \mathrm{Det}\begin{pmatrix} \frac{z_1}{\vdots} \\ \hline z_{i-1} \\ \hline z_i' \\ \hline z_{i+1} \\ \hline \vdots \\ \hline z_n \end{pmatrix} + \mathrm{Det}\begin{pmatrix} \frac{z_1}{\vdots} \\ \hline z_{i-1} \\ \hline z_i'' \\ \hline z_{i+1} \\ \hline \vdots \\ \hline z_n z_n \end{pmatrix}.
\end{aligned}$$

2) Es seien $1 \le i < j \le n$ Indizes, so dass die i-te und j-te Zeile übereinstimmen, d.h. $a_{ik} = a_{jk}$, $k = 1,...,n$. Man beachte, dass in den Bezeichnungen von Definition IV.3.16, ii), $U_n = A_n \cdot \tau_{ij}$ gilt. Wie im Beweis von Lemma IV.3.18 schreiben wir

$$S_n = A_n \sqcup U_n = A_n \sqcup A_n \cdot \tau_{ij}.$$

Dabei haben wir $A_n \cdot \tau_{ij} = \{\, \sigma\cdot\tau_{ij} \mid \sigma \in A_n \,\}$ gesetzt. Man beachte, dass $\mathrm{Sign}(\sigma) = 1$ und $\mathrm{Sign}(\sigma\cdot\tau_{ij}) = -1 = -\mathrm{Sign}(\sigma)$ für jede Permutation $\sigma \in A_n$ gilt. Wir erhalten

$$\sum_{\sigma\in S_n} \mathrm{Sign}(\sigma)\cdot a_{1\sigma(1)}\cdot\dots\cdot a_{n\sigma(n)}$$

$$= \sum_{\sigma\in A_n} \mathrm{Sign}(\sigma)\cdot a_{1\sigma(1)}\cdot\dots\cdot a_{n\sigma(n)} + \sum_{\sigma\in A_n} \mathrm{Sign}(\sigma\cdot\tau_{ij})\cdot a_{1\sigma(\tau_{ij}(1))}\cdot\dots\cdot a_{n\sigma(\tau_{ij}(n))}$$
$$= \sum_{\sigma\in A_n} \underbrace{a_{1\sigma(1)}\cdot\dots\cdot a_{n\sigma(n)} - a_{1\sigma(\tau_{ij}(1))}\cdot\dots\cdot a_{n\sigma(\tau_{ij}(n))}}_{=:D_\sigma}.$$

Für $l \notin \{i,j\}$ gilt $a_{l\sigma(\tau_{ij}(l))} = a_{l\sigma(l)}$. Weiter haben wir

$$\begin{aligned} a_{i\sigma(\tau_{ij}(i))} &= a_{i\sigma(j)} = a_{j\sigma(j)}, \\ a_{j\sigma(\tau_{ij}(j))} &= a_{j\sigma(i)} = a_{i\sigma(i)}. \end{aligned}$$

Mit diesen Formeln erkennen wir, dass $D_\sigma = 0$, $\sigma \in A_n$.

iii) Es gilt $\mathbb{E}_n = (\delta_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}}$ (Beispiel III.5.15, i). Für eine Permutation $\sigma \in S_n$ ist das Produkt $\delta_{1\sigma(1)}\cdot\dots\cdot\delta_{n\sigma(n)}$ genau dann ungleich null, wenn $i = \sigma(i)$, $i = 1,\dots,n$, also wenn $\sigma = e$. In diesem Fall hat das Produkt den Wert 1. □

Beispiele IV.4.7. i) Es sei $n = 2$. Wir haben $S_2 = \{e, \tau_{12}\}$. Damit ergibt die Leibnizformel

$$\mathrm{Det}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}\cdot a_{22} - a_{12}\cdot a_{21}.$$

Dies ist die Formel, die wir bereits in Definition I.1.2 erklärt hatten.

ii) Für $n = 3$ schreiben wir gemäß Definition IV.3.16 $S_3 = A_3 \sqcup U_3$. Dabei gilt $U_3 = \tau_{12}\cdot A_3$. Wir haben

$$\begin{aligned} A_3 &= \left\{ \begin{pmatrix} 1&2&3\\1&2&3 \end{pmatrix}, \begin{pmatrix} 1&2&3\\2&3&1 \end{pmatrix}, \begin{pmatrix} 1&2&3\\3&1&2 \end{pmatrix} \right\}, \\ U_3 = \tau_{12}\cdot A_3 &= \left\{ \begin{pmatrix} 1&2&3\\2&1&3 \end{pmatrix}, \begin{pmatrix} 1&2&3\\1&3&2 \end{pmatrix}, \begin{pmatrix} 1&2&3\\3&2&1 \end{pmatrix} \right\}. \end{aligned}$$

Die Leibnizformel ergibt

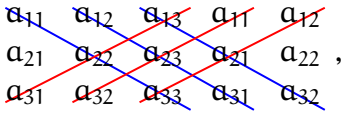
$$\begin{aligned} \mathrm{Det}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= a_{11}\cdot a_{22}\cdot a_{33} + a_{12}\cdot a_{23}\cdot a_{31} + a_{13}\cdot a_{21}\cdot a_{32} - \\ &\quad - a_{12}\cdot a_{21}\cdot a_{33} - a_{11}\cdot a_{23}\cdot a_{32} - a_{13}\cdot a_{22}\cdot a_{31}. \end{aligned}$$

Diese Formel kann man sich mit der **Regel von Sarrus**[10] merken: Gegeben eine (3×3)-Matrix

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

[10] Pierre Frédéric Sarrus (1798 - 1861), französischer Mathematiker.

so füge man die ersten beiden Spalten an,

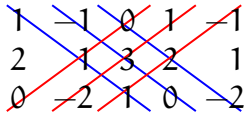


summiere die Produkte über die Diagonalen von links oben nach rechts unten und subtrahiere die Produkte über die Diagonalen von links unten nach rechts oben. Das Ergebnis ist die Determinante von A.

Für

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 2 & 1 & 3 \\ 0 & -2 & 1 \end{pmatrix}$$

bildet man also



und erhält

$$\mathrm{Det}(A) = 1 - (-6) - (-2) = 9.$$

Für $n \geq 4$ gibt es leider keine einfachen Merkregeln mehr.

Beobachtung IV.4.8. *Es sei $n \geq 1$ eine natürliche Zahl. Für jede $(n \times n)$-Matrix $A \in \mathrm{Mat}(n; K)$ gilt*

$$\mathrm{Det}(A^t) = \mathrm{Det}(A).$$

Beweis. Es seien $A = (a_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$ und $A^t = (a^t_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$. Nach Definition IV.2.4, ii), gilt $a^t_{ij} = a_{ji}$, $i, j = 1, \dots, n$. Die Leibnizformel liefert

$$\begin{aligned}
\mathrm{Det}(A^t) &= \sum_{\sigma \in S_n} \mathrm{Sign}(\sigma) \cdot a_{\sigma(1)1} \cdot \dots \cdot a_{\sigma(n)n} \\
&= \sum_{\sigma \in S_n} \mathrm{Sign}(\sigma) \cdot a_{1\sigma^{-1}(1)} \cdot \dots \cdot a_{n\sigma^{-1}(n)} \\
&= \sum_{\sigma \in S_n} \mathrm{Sign}(\sigma^{-1}) \cdot a_{1\sigma^{-1}(1)} \cdot \dots \cdot a_{n\sigma^{-1}(n)} \\
&= \sum_{\sigma^{-1} \in S_n} \mathrm{Sign}(\sigma^{-1}) \cdot a_{1\sigma^{-1}(1)} \cdot \dots \cdot a_{n\sigma^{-1}(n)} \\
&= \mathrm{Det}(A).
\end{aligned}$$

In der dritten Zeile haben wir die Formel $\mathrm{Sign}(\sigma^{-1}) = \mathrm{Sign}(\sigma)$ angewandt. □

Bemerkungen IV.4.9. i) Diese Beobachtung lässt sich auch folgendermaßen beweisen: Falls $\mathrm{Rg}(A) < n$, so gilt nach Folgerung III.5.41, iv), ebenfalls $\mathrm{Rg}(A^t) < n$, und Satz IV.4.4, i), zeigt $\mathrm{Det}(A^t) = 0 = \mathrm{Det}(A)$.

Mit Beispiel IV.4.3 erkennt man weiter, dass die Aussage für jede Elementarmatrix stimmt.

Wenn schließlich $\mathrm{Rg}(A) = n$ gilt, dann existieren nach Satz IV.1.17, i), eine natürliche Zahl $u \geq 1$ und Elementarmatrizen $D_1, ..., D_u$, so dass

$$A = D_1 \cdot \dots \cdot D_u,$$

und wir finden mit Folgerung IV.2.7, i), und Satz IV.4.4, ii),

$$\begin{aligned} \mathrm{Det}(A^t) &= \mathrm{Det}(D_u^t \cdot \dots \cdot D_1^t) = \mathrm{Det}(D_u^t) \cdot \dots \cdot \mathrm{Det}(D_1^t) = \mathrm{Det}(D_u) \cdot \dots \cdot \mathrm{Det}(D_1) \\ &= \mathrm{Det}(D_1) \cdot \dots \cdot \mathrm{Det}(D_u) = \mathrm{Det}(A). \end{aligned}$$

ii) Aus der Beobachtung folgt, dass eine Determinantenabbildung auch linear in jeder Spalte ist. Ferner impliziert das Verschwinden der Determinante für Matrizen, deren Rang nicht maximal ist, dass eine Matrix mit zwei identischen Spalten Determinante Null hat. Man hätte die Determinantenabbildung auch über diese beiden Eigenschaften und $\mathrm{Det}(\mathbb{E}_n) = 1$, $n \geq 1$, charakterisieren können.

Die Eigenschaften der Determinanten gestatten es uns, auch Determinanten für Endomorphismen zu definieren.

Lemma IV.4.10. *Es seien V ein endlichdimensionaler K-Vektorraum, $f\colon V \longrightarrow V$ ein Endomorphismus und B, B' zwei durchnummerierte Basen von V. Dann gilt*

$$\mathrm{Det}\big(M_B^B(f)\big) = \mathrm{Det}\big(M_{B'}^{B'}(f)\big).$$

Beweis. Nach Satz IV.1.29 und Eigenschaft IV.1.28, ii), existiert eine invertierbare Matrix $S \in \mathrm{GL}_n(K)$, so dass

$$M_{B'}^{B'}(f) = S \cdot M_B^B(f) \cdot S^{-1}.$$

Deshalb gilt

$$\begin{aligned} \mathrm{Det}\big(M_{B'}^{B'}(f)\big) &= \mathrm{Det}\big(S \cdot M_B^B(f) \cdot S^{-1}\big) = \mathrm{Det}(S) \cdot \mathrm{Det}\big(M_B^B(f)\big) \cdot \mathrm{Det}(S^{-1}) = \\ &= \mathrm{Det}(S) \cdot \mathrm{Det}(S^{-1}) \cdot \mathrm{Det}\big(M_B^B(f)\big) = \mathrm{Det}(S \cdot S^{-1}) \cdot \mathrm{Det}\big(M_B^B(f)\big) = \\ &= \mathrm{Det}(\mathbb{E}_n) \cdot \mathrm{Det}\big(M_B^B(f)\big) = \mathrm{Det}\big(M_B^B(f)\big), \end{aligned}$$

und das Lemma ist bewiesen. □

Mit diesem Lemma wird folgende Definition sinnvoll:

Definition IV.4.11. Es seien V ein endlichdimensionaler K-Vektorraum und $f\colon V \longrightarrow V$ ein Endomorphismus. Man wähle eine durchnummerierte Basis von V und setze

$$\mathrm{Det}(f) := \mathrm{Det}\big(M_B^B(f)\big).$$

Diese Zahl ist die *Determinante* von f.

Eigenschaften IV.4.12. *Es sei V ein endlichdimensionaler K-Vektorraum.*
i) *Die Abbildung*

$$\begin{array}{rccc} \mathrm{Det}\colon & \mathrm{GL}(V) & \longrightarrow & K \\ & f & \longmapsto & \mathrm{Det}(f) \end{array}$$

ist ein Gruppenhomomorphismus.
ii) *Es sei $f\colon V \longrightarrow V$ ein Endomorphismus. Die folgenden Aussagen sind äquivalent:*

- *f ist ein Isomorphismus.*
- $\mathrm{Det}(f) \neq 0$.

Beweis. Das ergibt sich mit Folgerung IV.4.5, ii), und Lemma IV.1.13. □

Die komplementäre Matrix. — Jeder Matrix kann eine sogenannte komplementäre Matrix zugewiesen werden. Für invertierbare Matrizen ist dies bis auf einen Faktor die inverse Matrix.

Definitionen IV.4.13. i) Es sei $A = (a_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}} \in \mathrm{Mat}(n;K)$ eine $(n \times n)$-Matrix. Für Indizes $i, j \in \{1, \dots, n\}$ sei

$$A_{ij} = \begin{pmatrix} a_{11} & \cdots & a_{1j-1} & 0 & a_{1j+1} & \cdots & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & & \vdots \\ a_{i-11} & \cdots & a_{i-1j-1} & 0 & a_{i-1j+1} & \cdots & \cdots & a_{i-1n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & \cdots & 0 \\ a_{i+11} & \cdots & a_{i+1j-1} & 0 & a_{i+1j+1} & \cdots & \cdots & a_{i+1n} \\ \vdots & & \vdots & \vdots & \vdots & & & \vdots \\ \vdots & & \vdots & \vdots & \vdots & & & \vdots \\ a_{n1} & \cdots & a_{nj-1} & 0 & a_{nj+1} & \cdots & \cdots & a_{nn} \end{pmatrix},$$

d.h. in der Matrix $A_{ij} = (a'_{kl})_{\substack{k=1,\dots,n\\ l=1,\dots,n}}$ gilt

$$a'_{kl} = \begin{cases} 0, & (k = i \wedge l \neq j) \vee (l = j \wedge k \neq i) \\ 1, & k = i \wedge l = j \\ a_{kl}, & \text{sonst} \end{cases}, \quad k, l = 1, \dots, n.$$

ii) Die Matrix $\widetilde{A} = (\widetilde{a}_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}} \in \mathrm{Mat}(n;K)$ mit[11]

$$\widetilde{a}_{ij} = \mathrm{Det}(A_{ji})$$

heißt die zu A *komplementäre Matrix*.

Satz IV.4.14. *Es seien* $A \in \mathrm{Mat}(n;K)$ *eine* $(n \times n)$*-Matrix und* $\widetilde{A}$ *ihre komplementäre Matrix. Dann gilt*

$$A \cdot \widetilde{A} = \widetilde{A} \cdot A = \mathrm{Det}(A) \cdot \mathbb{E}_n.$$

Beweis. Es seien $z_k := (\ a_{k1} \ \cdots \ a_{kn}\ )$ die k-te Zeile von A, $k = 1,\dots,n$, $e_l := (\ 0 \ \cdots \ 0 \ 1 \ 0 \ \cdots \ 0\ )$ der Zeilenvektor, bei dem die Eins an der l-ten Stelle steht, $l = 1,\dots,n$, und $i,j \in \{1,\dots,n\}$ Indizes. Wir bilden die Matrix

$$A'_{ij} = \begin{pmatrix} z_1 \\ \hline \vdots \\ \hline z_{i-1} \\ \hline e_j \\ \hline z_{i+1} \\ \hline \vdots \\ \hline z_n \end{pmatrix}.$$

Die Matrix A_{ij} in Definition IV.4.13, i), geht aus A'_{ij} durch Zeilenoperationen von Typ (ZI) hervor. Nach Satz IV.4.2, i), gilt

$$\mathrm{Det}(A_{ij}) = \mathrm{Det}(A'_{ij}).$$

Es sei $A \cdot \widetilde{A} = (\overline{a}_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}}$. Mit (IV.1) und Definition IV.4.13, i), haben wir

$$\begin{aligned} \overline{a}_{ij} &= \sum_{k=1}^{n} a_{ik} \cdot \widetilde{a}_{kj} = \sum_{k=1}^{n} a_{ik} \cdot \mathrm{Det}(A_{jk}) = \sum_{k=1}^{n} a_{ik} \cdot \mathrm{Det}(A'_{jk}) = \\ &= \mathrm{Det}\begin{pmatrix} z_1 \\ \hline \vdots \\ \hline z_{i-1} \\ \hline \sum\limits_{k=1}^{n} a_{ik} \cdot e_k \\ \hline z_{i+1} \\ \hline \vdots \\ \hline z_n \end{pmatrix} = \mathrm{Det}\begin{pmatrix} z_1 \\ \hline \vdots \\ \hline z_{i-1} \\ \hline z_i \\ \hline z_{i+1} \\ \hline \vdots \\ \hline z_n \end{pmatrix} = \delta_{ij} \cdot \mathrm{Det}(A), \quad i,j = 1,\dots,n. \end{aligned}$$

[11] Man beachte die Reihenfolge der Indizes auf der rechten Seite.

Ebenso könnten wir $\widetilde{A} \cdot A = \mathrm{Det}(A) \cdot \mathbb{E}_n$ nachrechnen. Wir argumentieren aber folgendermaßen: In Definition IV.4.13 gilt offenbar

$$\widetilde{A^t} = \widetilde{A}^t. \tag{IV.4}$$

Damit berechnen wir

$$\widetilde{A} \cdot A = (\widetilde{A} \cdot A)^{t^t} = (A^t \cdot \widetilde{A}^t)^t = (A^t \cdot \widetilde{A^t})^t = (\mathrm{Det}(A^t) \cdot \mathbb{E}_n)^t = (\mathrm{Det}(A) \cdot \mathbb{E}_n)^t = \mathrm{Det}(A) \cdot \mathbb{E}_n.$$

Bei diesen Umformungen haben wir Folgerung IV.2.7, i), (IV.4) und das bereits Bewiesene für die Matrix A^t benutzt. □

Definition IV.4.15. Es sei $A = (a_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,n}} \in \mathrm{Mat}(n; K)$ eine $(n \times n)$-Matrix. Für Indizes $i, j \in \{1, \dots, n\}$ sei

$$S_{ij}(A) = \begin{pmatrix} a_{11} & \cdots & a_{1j-1} & a_{1j+1} & \cdots & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & & \vdots \\ a_{i-11} & \cdots & a_{i-1j-1} & a_{i-1j+1} & \cdots & \cdots & a_{i-1n} \\ a_{i+11} & \cdots & a_{i+1j-1} & a_{i+1j+1} & \cdots & \cdots & a_{i+1n} \\ \vdots & & \vdots & \vdots & & & \vdots \\ \vdots & & \vdots & \vdots & & & \vdots \\ a_{n1} & \cdots & a_{nj-1} & a_{nj+1} & \cdots & \cdots & a_{nn} \end{pmatrix} \in \mathrm{Mat}(n-1; K),$$

d.h. entsteht aus A durch Streichung der i-ten Zeile und j-ten Spalte.

Lemma IV.4.16. *Es sei* $A = (a_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,n}} \in \mathrm{Mat}(n; K)$ *eine* $(n \times n)$*-Matrix. Für Indizes* $i, j \in \{1, \dots, n\}$ *gilt in den Bezeichnungen von Definition IV.4.13, i), und IV.4.15:*

$$\mathrm{Det}(A_{ij}) = (-1)^{i+j} \cdot \mathrm{Det}\big(S_{ij}(A)\big).$$

Beweis. Wir erinnern daran, dass die Vertauschung der k-ten und l-ten Zeile bzw. Spalte in einer $(n \times n)$-Matrix A durch Links- bzw. Rechtsmultiplikation mit der Matrix P_{kl} gegeben ist (s. Satz IV.1.10), $1 \le k < l \le n$. Folglich ändert die Determinante bei jeder Zeilen- oder Spaltenvertauschung das Vorzeichen (Beispiel IV.4.3 und Satz IV.4.4, ii). Nun formen wir die Matrix

$$A'_{ij} = \left(\begin{array}{ccc|c} & S_{ij}(A) & & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline 0 & \cdots & 0 & 1 \end{array}\right) = (a'_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,n}}.$$

Die Matrix A_{ij} geht aus A'_{ij} durch $(n-i)$ Zeilenvertauschungen und $(n-j)$ Spaltenvertauschungen hervor. Demnach haben wir

$$\begin{aligned}
\mathrm{Det}(A_{ij}) &= (-1)^{n-i+n-j} \cdot \sum_{\sigma\in S_n} \mathrm{Sign}(\sigma)\cdot a'_{1\sigma(1)}\cdot\dots\cdot a'_{n\sigma(n)} \\
&= (-1)^{i+j} \cdot \sum_{\sigma\in S_n:\sigma(n)=n} \mathrm{Sign}(\sigma)\cdot a'_{1\sigma(1)}\cdot\dots\cdot a'_{n\sigma(n)} \\
&\overset{a'_{nn}=1}{=} (-1)^{i+j} \cdot \sum_{\sigma\in S_{n-1}} \mathrm{Sign}(\sigma)\cdot a'_{1\sigma(1)}\cdot\dots\cdot a'_{n-1\sigma(n-1)} \\
&= (-1)^{i+j}\cdot \mathrm{Det}\big(S_{ij}(A)\big).
\end{aligned}$$

Dies ist die Behauptung des Lemmas. □

Laplace-Entwicklung von Determinanten. — Aus Satz IV.4.14 ergibt sich eine Berechnungsmethode für Determinanten, die einfacher anzuwenden ist als die Leibnizformel.

Satz IV.4.17. *Es sei* $A = (a_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}} \in \mathrm{Mat}(n;K)$ *eine* $(n\times n)$*-Matrix.*

i) *Es sei* $i\in\{1,\dots,n\}$ *der Index einer Zeile. Es gilt*

$$\mathrm{Det}(A) = \sum_{j=1}^{n} (-1)^{i+j}\cdot a_{ij}\cdot \mathrm{Det}\big(S_{ij}(A)\big).$$

Man spricht von ***Laplace***[12]***-Entwicklung nach der*** i***-ten Zeile****.*

ii) *Es sei* $j\in\{1,\dots,n\}$ *ein Spaltenindex. Es gilt*

$$\mathrm{Det}(A) = \sum_{i=1}^{n} (-1)^{i+j}\cdot a_{ij}\cdot \mathrm{Det}\big(S_{ij}(A)\big).$$

Dies ist die ***Laplace-Entwicklung nach der*** j***-ten Spalte****.*

Beweis. i) Es seien $\widetilde{A} = (\widetilde{a}_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}}$ die zu A komplementäre Matrix und $A\cdot\widetilde{A} = (a'_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}}$. Dann gilt nach Satz IV.4.14, (IV.1) und Lemma IV.4.16:

$$\mathrm{Det}(A) = a'_{ii} = \sum_{j=1}^{n} a_{ij}\cdot\widetilde{a}_{ji} = \sum_{j=1}^{n} a_{ij}\cdot\mathrm{Det}(A_{ij}) = \sum_{j=1}^{n}(-1)^{i+j}\cdot a_{ij}\cdot\mathrm{Det}\big(S_{ij}(A)\big).$$

Teil ii) wird genauso bewiesen. □

[12] Pierre-Simon (Marquis de) Laplace (1749 - 1827), französischer Mathematiker, Physiker und Astronom.

Bemerkung IV.4.18. Die Vorzeichen für die Entwicklung kann man sich als „Schachbrett“ merken, z.B. für eine (5×5)-Matrix:

$$\left(\begin{array}{c|c|c|c|c} + & - & + & - & + \\ \hline - & + & - & + & - \\ \hline + & - & + & - & + \\ \hline - & + & - & + & - \\ \hline + & - & + & - & + \end{array}\right).$$

Beispiele IV.4.19. i) Wir bestimmen die Determinante der Matrix

$$A := \begin{pmatrix} 2 & -2 & 3 \\ 4 & 1 & 2 \\ 1 & 5 & 1 \end{pmatrix} \in \mathrm{Mat}(3;\mathbb{Q})$$

durch Entwicklung nach der zweiten Zeile:

$$\begin{aligned} \mathrm{Det}(A) &= -4 \cdot \mathrm{Det}\begin{pmatrix} -2 & 3 \\ 5 & 1 \end{pmatrix} + \mathrm{Det}\begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} - 2 \cdot \mathrm{Det}\begin{pmatrix} 2 & -2 \\ 1 & 5 \end{pmatrix} \\ &= -4 \cdot (-2-15) + (2-3) - 2 \cdot (10+2) = 68 - 1 - 24 = 43. \end{aligned}$$

Wir überprüfen das Ergebnis mit der Regel von Sarrus (Beispiel IV.4.7, ii). Dazu bilden wir

$$\begin{array}{ccccc} 2 & -2 & 3 & 2 & -2 \\ 4 & 1 & 2 & 4 & 1 \\ 1 & 5 & 1 & 1 & 5 \end{array}$$

und erhalten

$$\mathrm{Det}(A) = 2 - 4 + 60 - 3 - 20 + 8 = 43.$$

ii) Wir möchten die Determinante der Matrix

$$B := \begin{pmatrix} 4 & 7 & -1 & 2 \\ 2 & 0 & 3 & -1 \\ -1 & 0 & 5 & 2 \\ 0 & 1 & -4 & 1 \end{pmatrix} \in \mathrm{Mat}(4;\mathbb{Q})$$

berechnen. Dazu entwickeln wir nach der zweiten Spalte:

$$\mathrm{Det}(B) = -7 \cdot \mathrm{Det}\begin{pmatrix} 2 & 3 & -1 \\ -1 & 5 & 2 \\ 0 & -4 & 1 \end{pmatrix} + \mathrm{Det}\begin{pmatrix} 4 & -1 & 2 \\ 2 & 3 & -1 \\ -1 & 5 & 2 \end{pmatrix}.$$

Die auftretenden Determinanten von (3×3)-Matrizen ermitteln wir wiederum mit der Regel von Sarrus: Mit

$$\begin{array}{ccccc} 2 & 3 & -1 & 2 & 3 \\ -1 & 5 & 2 & -1 & 5 \\ 0 & -4 & 1 & 0 & -4 \end{array}$$

finden wir

$$\operatorname{Det}\begin{pmatrix} 2 & 3 & -1 \\ -1 & 5 & 2 \\ 0 & -4 & 1 \end{pmatrix} = 10 - 4 + 16 + 3 = 25$$

und mit

$$\begin{array}{ccccc} 4 & -1 & 2 & 4 & -1 \\ 2 & 3 & -1 & 2 & 3 \\ -1 & 5 & 2 & -1 & 5 \end{array}$$

$$\operatorname{Det}\begin{pmatrix} 4 & -1 & 2 \\ 2 & 3 & -1 \\ -1 & 5 & 2 \end{pmatrix} = 24 - 1 + 20 + 6 + 20 + 4 = 73.$$

Zusammengenommen ergibt sich

$$\operatorname{Det}(B) = -7 \cdot 25 + 73 = -102.$$

Bemerkungen IV.4.20. i) Mit Hilfe der Laplace-Entwicklung kann man die Determinante auch rekursiv einführen. Man beginnt mit

$$\begin{array}{rcl} \operatorname{Det}\colon \operatorname{Mat}(1;K) & \longrightarrow & K \\ (a) & \longmapsto & a \end{array}$$

und definiert die Determinante einer $((n+1)\times(n+1))$-Matrix mittels Laplace-Entwicklung nach der letzten Zeile durch Determinanten von $(n\times n)$-Matrizen, $n \geq 1$. Man kann sich durch einen einfachen Induktionsbeweis vergewissern, dass man auch auf diese Weise die Leibnizformel erhält. Bei der Laplace-Entwicklung wird im Gegensatz zur Leibnizformel geklammert. Die Leserin bzw. der Leser möge sich das am Beispiel von (3×3)-Matrizen vor Augen führen.

ii) Den Aufwand für ein Rechenverfahren misst man mit der Anzahl der durchgeführten Gleitkommaoperationen. Dies sind Additionen oder Multiplikationen. Die Multiplikation mit -1 wird nicht gezählt. Bei der Leibnizformel müssen $n! - 1$ Additionen und $n! \cdot (n-1)$ Multiplikationen ausgeführt werden, insgesamt also $n \cdot n! - 1$ Gleitkommaoperationen. Durch die Klammerung bei der Laplace-Entwicklung reduziert sich der Rechnenaufwand etwas auf $2 \cdot n! - 1$ Gleitkommaoperationen. Für den Gauß-Algorithmus (s. Aufgabe IV.4.30) werden allerdings nur (n^3+Terme niederer Ordnung) Gleitkommaoperationen benötigt. Das ist somit die effizienteste Berechnungsmethode. Für eine ausführlichere Diskussion verweisen wir auf [5].

Die Cramerschen Regeln. — Satz IV.4.14 beinhaltet zwei bekannte Regeln für die Inversion einer Matrix und die Lösung gewisser Gleichungssysteme. Diese Regeln sind rechenaufwändig aber für gewisse theoretische Erwägungen wertvoll.

Lemma IV.4.21 (Cramersche[13] Regel für die Matrixinversion)**.** *Gegeben sei eine invertierbare* $(n \times n)$*-Matrix* $A \in \mathrm{GL}_n(K)$*. Dann gilt in den Bezeichnungen von Definition IV.4.13 und IV.4.15:*

$$A^{-1} = \left(\frac{\mathrm{Det}(A_{ji})}{\mathrm{Det}(A)}\right)_{\substack{i=1,\dots,n\\ j=1,\dots,n}} = \left(\frac{(-1)^{i+j}\cdot \mathrm{Det}\big(S_{ji}(A)\big)}{\mathrm{Det}(A)}\right)_{\substack{i=1,\dots,n\\ j=1,\dots,n}}.$$

Beweis. Das ergibt sich unmittelbar aus Satz IV.4.14. □

Beispiel IV.4.22. Wir berechnen die Inverse der Matrix

$$A := \begin{pmatrix} 1 & -2 & 0 \\ 1 & 0 & -3 \\ -1 & 2 & 5 \end{pmatrix} \in \mathrm{Mat}(n;\mathbb{Q})$$

mit der Cramerschen Regel. Dazu ermitteln wir zunächst die Koeffizienten der komplementären Matrix $\widetilde{A} = (\widetilde{a}_{ij})_{\substack{i=1,\dots,3\\ j=1,\dots,3}}$:

$$\begin{aligned}
\widetilde{a}_{11} &= \mathrm{Det}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -3 \\ 0 & 2 & 5 \end{pmatrix} = \mathrm{Det}\begin{pmatrix} 0 & -3 \\ 2 & 5 \end{pmatrix} = 6,\\
\widetilde{a}_{12} &= \mathrm{Det}\begin{pmatrix} 0 & -2 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 5 \end{pmatrix} = -\mathrm{Det}\begin{pmatrix} -2 & 0 \\ 2 & 5 \end{pmatrix} = 10,\\
\widetilde{a}_{13} &= \mathrm{Det}\begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & -3 \\ 1 & 0 & 0 \end{pmatrix} = \mathrm{Det}\begin{pmatrix} -2 & 0 \\ 0 & -3 \end{pmatrix} = 6,\\
\widetilde{a}_{21} &= \mathrm{Det}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & -3 \\ -1 & 0 & 5 \end{pmatrix} = -\mathrm{Det}\begin{pmatrix} 1 & -3 \\ -1 & 5 \end{pmatrix} = -2,\\
\widetilde{a}_{22} &= \mathrm{Det}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 5 \end{pmatrix} = \mathrm{Det}\begin{pmatrix} 1 & 0 \\ -1 & 5 \end{pmatrix} = 5,\\
\widetilde{a}_{23} &= \mathrm{Det}\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & -3 \\ 0 & 1 & 0 \end{pmatrix} = -\mathrm{Det}\begin{pmatrix} 1 & 0 \\ 1 & -3 \end{pmatrix} = 3,\\
\widetilde{a}_{31} &= \mathrm{Det}\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ -1 & 2 & 0 \end{pmatrix} = \mathrm{Det}\begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} = 2,
\end{aligned}$$

[13] Gabriel Cramer (1704 - 1752), Genfer Mathematiker.

$$\begin{aligned}
\widetilde{a}_{32} &= \operatorname{Det}\begin{pmatrix} 1 & -2 & 0 \\ 0 & 0 & 1 \\ -1 & 2 & 0 \end{pmatrix} = -\operatorname{Det}\begin{pmatrix} 1 & -2 \\ -1 & 2 \end{pmatrix} = 0, \\
\widetilde{a}_{33} &= \operatorname{Det}\begin{pmatrix} 1 & -2 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \operatorname{Det}\begin{pmatrix} 1 & -2 \\ 1 & 0 \end{pmatrix} = 2.
\end{aligned}$$

Ferner gilt $\operatorname{Det}(A) = 10$, so dass

$$A^{-1} = \frac{1}{10} \cdot \begin{pmatrix} 6 & 10 & 6 \\ -2 & 5 & 3 \\ 2 & 0 & 2 \end{pmatrix}.$$

Wenden wir uns nun der Cramerschen Regel für Gleichungssysteme zu. Es seien $A \in \mathrm{GL}_n(K)$ eine **invertierbare** $(n \times n)$-Matrix und $b \in K^n$. Das Gleichungssystem

$$A \cdot \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} = b$$

hat genau eine Lösung, und zwar

$$\begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} = A^{-1} \cdot b. \tag{IV.5}$$

Definition IV.4.23. Es seien $A = (a_{ij}) \in \mathrm{GL}_n(K)$ eine invertierbare $(n \times n)$-Matrix und $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in K^n$ ein Vektor. Für $i = 1, \ldots, n$ setzen wir

$$B_i = \begin{pmatrix} a_{11} & \cdots & a_{1i-1} & b_1 & a_{1i+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{ni-1} & b_n & a_{ni+1} & \cdots & a_{nn} \end{pmatrix},$$

d.h. wir ersetzen die i-te Spalte von A durch den Vektor b.

Lemma IV.4.24 (Cramersche Regel für lineare Gleichungssysteme)**.** *In Gleichung* (IV.5) *gilt*

$$s_i = \frac{\operatorname{Det}(B_i)}{\operatorname{Det}(A)}, \quad i = 1, \ldots, n.$$

Beweis. Mit der Cramerschen Regel für die Berechnung der inversen Matrix (Lemma IV.4.21) und der Formel für das Matrixprodukt (IV.1) folgt aus (IV.5)

$$s_i = \sum_{j=1}^{n} \frac{\mathrm{Det}(A_{ji})}{\mathrm{Det}(A)} \cdot b_j, \quad i = 1, ..., n.$$

Mit Lemma IV.4.16 wird daraus

$$s_i = \sum_{j=1}^{n} \frac{(-1)^{i+j} \cdot \mathrm{Det}\big(S_{ji}(A)\big)}{\mathrm{Det}(A)} \cdot b_j, \quad i = 1, ..., n.$$

Auf der anderen Seite ergibt die Laplace-Entwicklung der Determinante von B_i nach der i-ten Spalte (Satz IV.4.17, ii)

$$\mathrm{Det}(B_i) = \sum_{j=1}^{n} (-1)^{i+j} \cdot b_j \cdot \mathrm{Det}\big(S_{ji}(A)\big), \quad i = 1, ..., n.$$

Vergleicht man diese Formeln, so gelangt man zum gewünschten Ergebnis. □

Beispiel IV.4.25. Gegeben seien

$$A := \begin{pmatrix} -1 & 2-i & 0 \\ i & 0 & -i \\ 1+3i & i & 0 \end{pmatrix} \in \mathrm{Mat}(3;\mathbb{C}), \quad b := \begin{pmatrix} 3-2i \\ 0 \\ 1+5i \end{pmatrix} \in \mathbb{C}^3,$$

und gesucht ist die Lösungsmenge $\mathbb{L}(A, b)$. Zunächst wird die Determinante von A mit der Regel von Sarrus (Beispiel IV.4.7, ii) bestimmt:

$$\begin{matrix} -1 & 2-i & 0 & -1 & 2-i \\ i & 0 & -i & i & 0 \\ 1+3i & i & 0 & 1+3i & i \end{matrix},$$

also

$$\mathrm{Det}(A) = (2-i)(-i)(1+3i) - (-1)(-i)i = (2-i)(3-i) + 1 = 6 - 1 - 2i - 3i + 1 = 6 - 5i.$$

Jetzt stellen wir die in der Cramerschen Regel benötigten „Hilfsmatrizen“ auf und berechnen deren Determinanten ebenfalls mit der Sarrus-Regel:

$$B_1 = \begin{pmatrix} 3-2i & 2-i & 0 \\ 0 & 0 & -i \\ 1+5i & i & 0 \end{pmatrix}, \qquad \begin{matrix} 3-2i & 2-i & 0 & 3-2i & 2-i \\ 0 & 0 & -i & 0 & 0 \\ 1+5i & i & 0 & 1+5i & i \end{matrix},$$

$$\begin{aligned} \mathrm{Det}(B_1) &= (2-i)(-i)(1+5i) - i(-i)(3-2i) = (2-i)(5-i) - (3-2i) \\ &= 10 - 1 - 2i - 5i - 3 + 2i = 6 - 5i; \end{aligned}$$

$$B_2 := \begin{pmatrix} -1 & 3-2i & 0 \\ i & 0 & -i \\ 1+3i & 1+5i & 0 \end{pmatrix}, \qquad \begin{matrix} -1 & 3-2i & 0 & -1 & 3-2i \\ i & 0 & -i & i & 0 \\ 1+3i & 1+5i & 0 & 1+3i & 1+5i \end{matrix},$$

$$\begin{aligned} \mathrm{Det}(B_2) &= (3-2i)(-i)(1+3i) - (-1)(-i)(1+5i) = (3-2i)(3-i) + 5 - i \\ &= 9 - 2 - 3i - 6i + 5 - i = 12 - 10i; \end{aligned}$$

$$B_3 = \begin{pmatrix} -1 & 2-i & 3-2i \\ i & 0 & 0 \\ 1+3i & i & 1+5i \end{pmatrix}, \qquad \begin{matrix} -1 & 2-i & 3-2i & -1 & 2-i \\ i & 0 & 0 & i & 0 \\ 1+3i & i & 1+5i & 1+3i & i \end{matrix},$$

$$\begin{aligned} \mathrm{Det}(B_3) &= (3-2i)i^2 - (2-i)i(1+5i) \\ &= 2i - 3 - (2-i)(-5+i) = 2i - 3 + 10 - 1 - 2i - 5i = 6 - 5i. \end{aligned}$$

Schließlich ergibt sich

$$\mathbb{L}(A, b) = \left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\}.$$

Aufgabe IV.4.26.
Berechnen Sie die Determinanten der Matrizen

$$\begin{pmatrix} -1 & 7 \\ 0,25 & 5 \end{pmatrix}, \quad \begin{pmatrix} -1 & 2,5 & 4 \\ 0,5 & 5 & -2 \\ 4 & 0 & 12 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} \overline{3} & \overline{14} & \overline{2} \\ \overline{0} & \overline{7} & \overline{10} \\ \overline{8} & \overline{2} & \overline{4} \end{pmatrix}.$$

Dabei seien die ersten beiden Matrizen über $\mathbb{R}$ und die dritte über $\mathbb{F}_{17}$ definiert. Im letzten Fall ist das Ergebnis wieder in der Form $\overline{a}$ mit $0 \leq a < 17$ anzugeben.

Aufgabe IV.4.27.
Berechnen Sie die Determinanten der folgenden über $\mathbb{R}$ definierten Matrizen:

$$\text{a)} \begin{pmatrix} 1 & -2 & 0 & 5 \\ -3 & 7 & 2 & -11 \\ 2 & 2 & 0 & -6 \\ -1 & 14 & 2 & 3 \end{pmatrix} \quad \text{und} \quad \text{b)} \begin{pmatrix} 0 & 5 & 1 & -6 \\ -1 & 12 & 1 & -12 \\ 2 & -21 & 0 & 19 \\ -11 & 1 & -1 & 11 \end{pmatrix}.$$

Aufgabe IV.4.28.
a) Eine $(n \times n)$-Matrix $A = (a_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}$ mit $a_{ij} = 0$ für $i > j$ nennt man *obere Dreiecksmatrix*. Eine obere Dreiecksmatrix hat also die Form

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}.$$

Beweisen Sie

$$\mathrm{Det}(A) = a_{11} \cdot \cdots \cdot a_{nn}.$$

b) Es seien $\lambda_1, ..., \lambda_n \in K$. Beweisen Sie die Gleichung

$$\mathrm{Det}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_n^2 \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{pmatrix} = \prod_{i<j}(\lambda_j - \lambda_i).$$

Diese Determinante wird als *Vandermonde* [14]*-Determinante* bezeichnet.

Aufgabe IV.4.29.

a) Bestimmen Sie die inverse Matrix der reellen Matrix

$$\begin{pmatrix} 0 & 1 & -4 \\ 1 & 2 & -1 \\ 1 & 1 & 2 \end{pmatrix}$$

mit der Cramerschen Regel.

b) Lösen Sie folgende Gleichungssysteme mit der Cramerschen Regel.

i) ($K = \mathbb{C}$).

$$\begin{array}{rcrcl} (2-\sqrt{2}i)x_1 & + & 3x_2 & = & 2i \\ -ix_1 & + & (2+\sqrt{2}i)x_2 & = & -5 \end{array}.$$

ii) ($K = \mathbb{R}$).

$$\begin{array}{rcrcrcl} -x_1 & + & 3x_2 & - & 2x_3 & = & -44 \\ & & 2x_2 & - & x_3 & = & -23 \\ 4x_1 & & & + & 7x_3 & = & 113 \end{array}.$$

Aufgabe IV.4.30.

Entwickeln Sie mit dem Gauß-Algorithmus eine Berechnungsmethode für Determinanten, und bestimmen Sie die Determinante der reellen Matrix

$$\begin{pmatrix} 5 & 8 & 9 & 8 & 5 \\ -5 & 4 & -3 & 2 & -1 \\ 10 & 12 & 9 & 4 & 1 \\ 10 & 8 & -3 & -2 & 3 \\ 35 & 8 & 18 & 10 & 3 \end{pmatrix}$$

mit diesem Verfahren.

[14] Alexandre-Théophile Vandermonde (1735 - 1796), französischer Musiker, Mathematiker und Chemiker.

V

Eigenwerte und -vektoren

In diesem Kapitel werden wir zunächst die Begriffe des Eigenwertes und des Eigenvektors einführen. Diese Begriffe spielen auch in den Anwendungen eine große Rolle. An dieser Stelle sei etwa die Quantenmechanik erwähnt, in der man grob gesprochen mit Eigenwerten und -vektoren eines linearen Operators auf einem Hilbertraum quantenmechanische Systeme beschreibt. Hierbei wird mit unendlichdimensionalen Vektorräumen gearbeitet. In der Vorlesung wollen wir Eigenwerte und -vektoren dazu benutzen, die Struktur eines Endomorphismus f eines endlichdimensionalen K-Vektorraums V zu analysieren. Die Ergebnisse der Analyse werden durch die Jordansche Normalform zusammengefasst.
Generalvoraussetzung. Alle Vektorräume, die wir betrachten, haben mindestens die Dimension 1.

V.1 Äquivalenzrelationen auf Matrizen

Definition V.1.1. Es seien $m \geq 1$ und $n \geq 1$ natürliche Zahlen. Zwei Matrizen $A, B \in \mathrm{Mat}(m, n; K)$ heißen *äquivalent*, wenn es Matrizen $S \in \mathrm{GL}_m(K)$ und $T \in \mathrm{GL}_n(K)$ gibt, so dass

$$B = S \cdot A \cdot T^{-1}.$$

Schreibweise. $A \sim B$.

Bemerkung V.1.2. Bei der Relation „$\sim$“ handelt es sich in der Tat um eine Äquivalenzrelation. Für $A \in \mathrm{Mat}(m, n; K)$ gilt offenkundig $A = \mathbb{E}_m \cdot A \cdot \mathbb{E}_n = \mathbb{E}_m \cdot A \cdot \mathbb{E}_n^{-1}$. Da $\mathbb{E}_m \in \mathrm{GL}_m(K)$ und $\mathbb{E}_n \in \mathrm{GL}_n(K)$, folgt $A \sim A$. Es seien $A, B \in \mathrm{Mat}(m, n; K)$ und

$A \sim B$. Man wähle Matrizen $S \in GL_m(K)$ und $T \in GL_n(K)$ mit $B = S \cdot A \cdot T^{-1}$. Wir erhalten $A = (S^{-1}) \cdot B \cdot (T^{-1})^{-1}$. Es gilt $S^{-1} \in GL_m(K)$ und $T^{-1} \in GL_n(K)$, so dass wir $B \sim A$ schließen. Zu guter Letzt seien $A, B, C \in \mathrm{Mat}(m, n; K)$ mit $A \sim B$ und $B \sim C$ gegeben. Es gibt also Matrizen $S, U \in GL_m(K)$ und $T, V \in GL_n(K)$ mit $B = S \cdot A \cdot T^{-1}$ und $C = U \cdot B \cdot V^{-1}$. Zusammengenommen ergibt sich

$$C = U \cdot (S \cdot A \cdot T^{-1}) \cdot V^{-1} = (U \cdot S) \cdot A \cdot (T^{-1} \cdot V^{-1}) = (U \cdot S) \cdot A \cdot (V \cdot T)^{-1}.$$

Nun gilt $U \cdot S \in GL_m(K)$ und $V \cdot T \in GL_n(K)$, so dass $A \sim C$ folgt.

Der Hintergrund für diese Begriffsbildung ist folgender: Gegeben seien endlichdimensionale K-Vektorräume V und W sowie eine lineare Abbildung $f\colon V \longrightarrow W$. Gesucht sind durchnummerierte Basen B und C für V bzw. W, so dass die darstellende Matrix

$$M_C^B(f)$$

eine möglichst einfache Gestalt hat. Dieses Problem wurde in Folgerung IV.1.30 gelöst. Man beachte, dass für durchnummerierte Basen B und B' von V bzw. C und C' von W der Zusammenhang

$$M_{C'}^{B'}(f) = M_{C'}^{C}(\mathrm{Id}_W) \cdot M_C^B(f) \cdot \left(M_{B'}^{B}(\mathrm{Id}_V)\right)^{-1} \text{gilt.}$$

Dabei hat man $M_{C'}^{C}(\mathrm{Id}_W) \in GL_m(K)$ und $M_{B'}^{B}(\mathrm{Id}_V) \in GL_n(K)$. Damit ist die natürliche Motivation für den Äquivalenzbegriff von Matrizen gegeben, und Folgerung IV.1.30 kann umformuliert werden in:

Satz V.1.3. *Eine $(m \times n)$-Matrix A vom Rang r ist äquivalent zu der Matrix*

$$\left(\begin{array}{c|c} \mathbb{E}_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

Insbesondere sind zwei Matrizen A und B genau dann äquivalent, wenn sie denselben Rang haben.

Definition V.1.4. Zwei Matrizen $A, B \in \mathrm{Mat}(n; K)$ nennt man *ähnlich*, wenn es eine Matrix $S \in GL_n(K)$ mit

$$B = S \cdot A \cdot S^{-1}$$

gibt.
Schreibweise. $A \approx B$.

Bemerkung V.1.5. Wie in Bemerkung V.1.2 sieht man, dass es sich bei der Relation „$\approx$“ wirklich um eine Äquivalenzrelation handelt.

Der Hintergrund für dieses Problem lässt sich wie folgt verstehen: Wir betrachten einen endlichdimensionalen K-Vektorraum V und einen Endomorphismus $f\colon V \longrightarrow V$.

Problem V.1.6. Diesmal suchen wir eine durchnummerierte Basis B von V, für die die Matrix

$$M_B(f) := M_B^B(f)$$

möglichst einfach wird.

Für zwei durchnummerierte Basen B und B' von V besteht der Zusammenhang

$$M_{B'}(f) = M_{B'}^B(\mathrm{id}_V) \cdot M_B(f) \cdot \left(M_{B'}^B(\mathrm{id}_V)\right)^{-1}.$$

Es gilt dabei $M_{B'}^B(\mathrm{id}_V) \in \mathrm{GL}_n(K)$. Damit ist die Herkunft des Begriffs „Ähnlichkeit“ geklärt. Es wird sich herausstellen, dass die Lösung des gerade gestellten Problems wesentlich komplizierter ist als die des Äquivalenzproblems von Matrizen. Um es zu lösen, muss man eine detaillierte Analyse der Struktur des Endomorphismus f durchführen. Dies führt uns endlich einmal ein bißchen weg von linearen Gleichungssystemen und dem Gauß-Algorithmus (auch wenn wir ihn wieder einsetzen) und zeigt, dass die Lineare Algebra keine reine Hilfswissenschaft ist, sondern auch interessante Fragestellungen in sich birgt.

Definitionen V.1.7. i) Es seien V ein K-Vektorraum und $f\colon V \longrightarrow V$ ein Endomorphismus. Eine Zahl $\lambda \in K$ wird *Eigenwert von* f genannt, wenn es einen Vektor $v \in V \setminus \{0\}$ gibt, so dass

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

Der Vektor v heißt dann *Eigenvektor von* f (*zum Eigenwert* λ).

ii) Es seien $n \geq 1$ eine natürliche Zahl und A eine $(n \times n)$-Matrix. Man nennt eine Zahl $\lambda \in K$ einen *Eigenwert der Matrix* A, wenn es einen Vektor $v \in K^n \setminus \{0\}$ mit

$$A \cdot v = \lambda \cdot v$$

gibt. Der Vektor v heißt *Eigenvektor von* A (*zum Eigenwert* λ).

iii) In der Situation von i) bzw. ii) setzen wir

$$\begin{aligned} \mathrm{Eig}(f,\lambda) &:= \left\{ v \in V \,|\, f(v) = \lambda \cdot v \right\} = \mathrm{Ker}(f - \lambda \cdot \mathrm{Id}_V) \\ \text{bzw. } \mathrm{Eig}(A,\lambda) &:= \left\{ v \in K^n \,|\, A \cdot v = \lambda \cdot v \right\} = \mathbb{L}(A - \lambda \cdot \mathbb{E}_n, 0). \end{aligned}$$

Man spricht dabei von dem *Eigenraum von* f bzw. A (*zum Eigenwert* λ).

Bemerkungen V.1.8. i) Die Eigenräume $\mathrm{Eig}(f,\lambda)$ bzw. $\mathrm{Eig}(A,\lambda)$ sind lineare Teilräume von V bzw. K^n.

ii) Man muss sich den Eigenraum $E := \mathrm{Eig}(f,\lambda)$ des Endomorphismus f als Teilraum von V vorstellen, der dadurch ausgezeichnet ist, dass E invariant ist, d.h. $f(E) \subseteq E$, und dass f auf E besonders einfach operiert, nämlich durch Multiplikation mit dem Skalar λ. Das anschließende Lemma gibt bereits präzisere Hinweise darauf, wie Eigenvektoren bei der Lösung von Problem V.1.6 hilfreich sein können.

iii) Die Zahl 0 kommt offenbar genau dann als Eigenwert von f vor, wenn f nicht injektiv ist. Der Eigenraum zum Eigenwert 0 ist nichts anderes als der Kern der Abbildung f.

Lemma V.1.9. *Es seien V ein endlichdimensionaler K-Vektorraum und $f\colon V \longrightarrow V$ ein Endomorphismus. Dann sind folgende Bedingungen äquivalent:*

a) *Es gibt eine durchnummerierte Basis B von V, so dass*

$$M_B(f) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

b) *Es gibt eine durchnummerierte Basis B von V, die nur aus Eigenvektoren von V besteht.*

Beweis. **a) $\Longrightarrow$ b).** Es sei $B = \{b_1, ..., b_n\}$. Aus der Definition der darstellenden Matrix folgt $f(b_i) = \lambda_i \cdot b_i$, $i = 1, ..., n$.

b) $\Longrightarrow$ a). Es seien $B = \{b_1, ..., b_n\}$ und $\lambda_1, ..., \lambda_n \in K$, so dass b_i ein Eigenvektor zum Eigenwert λ_i ist, $i = 1, ..., n$. Die Darstellungsmatrix bezüglich dieser Basis hat offensichtlich die in a) angegebene Gestalt. □

Definitionen V.1.10. i) Eine Matrix $A = (a_{ij})_{\substack{i=1,...,n \\ j=1,...,n}}$ heißt *Diagonalmatrix*, wenn $a_{ij} = 0$ für alle $1 \leq i \neq j \leq n$ gilt.
Schreibweise. $A := \mathrm{Diag}(a_{11}, ..., a_{nn})$.

ii) Ein Endomorphismus $f\colon V \longrightarrow V$ des endlichdimensionalen K-Vektorraums V heißt *diagonalisierbar* oder *halbeinfach*, wenn es eine durchnummerierte Basis B von V gibt, für die $M_B(f)$ eine Diagonalmatrix ist.

iii) Eine Matrix $A \in \mathrm{Mat}(n; K)$ heißt *diagonalisierbar*, wenn sie einer Diagonalmatrix ähnlich ist.

Beispiel V.1.11. Es sei $A := \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix} \in \mathrm{Mat}(2; \mathbb{R})$. Dann ist $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ein Eigenvektor zum Eigenwert 2 und $\begin{pmatrix} -3 \\ 1 \end{pmatrix}$ einer zum Eigenwert 1. Damit ist A diagonalisierbar. Es gilt

$$\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

Man beachte, dass

$$\begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}^{-1}.$$

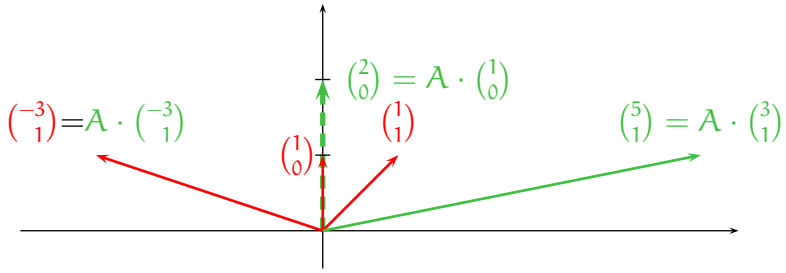


Abbildung V.1: Die Abbildung f_A.

Die Vektoren $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} -3 \\ 1 \end{pmatrix}$ definieren die ausgezeichneten Richtungen in $\mathbb{R}^2$, auf denen Multiplikation mit A als Multiplikation mit 2 bzw. 1 wirkt. Die anderen Richtungen sind nicht invariant, d.h. für

$$v \notin \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle \cup \left\langle \begin{pmatrix} -3 \\ 1 \end{pmatrix} \right\rangle$$

gilt $A \cdot v \notin \langle v \rangle$ (vgl. Abbildung 5.1).

Aufgabe V.1.12.
Beweisen Sie durch vollständige Induktion über n: Gegeben seien ein Vektorraum V und ein Endomorphismus $f\colon V \longrightarrow V$ sowie n verschiedene Eigenwerte $\lambda_1,\ldots,\lambda_n \in K$ von f mit Eigenvektoren $v_i \in V\setminus\{0\}$, i.e. $f(v_i) = \lambda_i \cdot v_i$, $i = 1, \ldots, n$. Dann ist $\{v_1, \ldots, v_n\}$ eine linear unabhängige Teilmenge von V.

V.2 Polynome

Wir erinnern zunächst an den Ring aus Aufgabe II.3.7, ii). Es sei

$$R := \mathrm{Abb}'(\mathbb{N}, K) := \left\{ p\colon \mathbb{N} \longrightarrow K \,\middle|\, p(n) \neq 0 \text{ nur für endlich viele } n \in \mathbb{N} \right\}.$$

Auf R definieren die folgenden Operationen die Struktur eines K-Vektorraums:

$$\begin{aligned} \lambda \cdot p\colon \mathbb{N} &\longrightarrow K, \quad n \longmapsto \lambda \cdot p(n),\ \lambda \in K,\ p \in R, \\ p + q\colon \mathbb{N} &\longrightarrow K, \quad n \longmapsto p(n) + q(n),\ p, q \in R. \end{aligned}$$

Die Multiplikation, mit der R zum Ring wird, ist wie folgt erklärt:

$$p \cdot q\colon \mathbb{N} \longrightarrow K, \quad n \longmapsto \sum_{\substack{k,l\in\mathbb{N} \\ k+l=n}} p(k) \cdot q(l),\ p, q \in R.$$

Die Elemente

$$\begin{array}{rcl} e_j\colon \mathbb{N} & \longrightarrow & K \\ n & \longmapsto & \begin{cases} 1, & n = j \\ 0, & \text{sonst} \end{cases} \end{array},$$

$j \in \mathbb{N}$, bilden eine Basis für den K-Vektorraum R. Offenbar haben wir die Regel

$$e_i \cdot e_j = e_{i+j}, \quad i, j \in \mathbb{N}.$$

Wir setzen $t := e_1$, $t^0 := 1 = e_0$. Dann gilt $t^j = e_j$ für alle $j \in \mathbb{N}$.
Schreibweise. Wir können $p \in R$ mit $p(j) = 0$ für alle $j > n$ also auf eindeutige Weise in der Form

$$p(t) := p = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0$$

mit $a_j \in K$, $j = 0, ..., n$, schreiben. Wir nennen $p(t)$ ein *Polynom in der Unbestimmten* t. Die Schreibweise $p(t)$ hebt die Unbestimmte t hervor und ist nicht mit dem Wert von p an einer natürlichen Zahl zu verwechseln. Ferner setzen wir $K[t] := R$ und sprechen von dem *Polynomring in der Unbestimmten* t.

Beispiel V.2.1. Wir berechnen eine Summe und ein Produkt von Polynomen:

$$(t^5 + 4t^2 + 3t) + (2t^7 - 5t^4 + 3t^2 + t - 1) = 2t^7 + t^5 - 5t^4 + 7t^2 + 4t - 1;$$

$$(t^2 + 3t + 2) \cdot (2t + 1) = 2t^3 + 7t^2 + 7t + 2.$$

Definition V.2.2. Es sei $p = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0$. Der *Grad von* p ist definiert als

$$\operatorname{Grad}(p) := \begin{cases} -\infty, & a_0 = \cdots = a_n = 0 \\ \max\{\, j \mid a_j \neq 0 \,\}, & \text{sonst} \end{cases}.$$

Lemma V.2.3. *Für zwei Polynome* $p, q \in K[t]$ *gilt:*[1]

a) $\operatorname{Grad}(p + q) \leq \max\{\operatorname{Grad}(p), \operatorname{Grad}(q)\}$.

b) $\operatorname{Grad}(p \cdot q) = \operatorname{Grad}(p) + \operatorname{Grad}(q)$.

[1] Dabei gelten folgende Rechenregeln für $-\infty$:

- $\forall n \in \mathbb{N}$: $-\infty < n$,
- $\forall n \in \mathbb{N}$: $-\infty + n = -\infty$,
- $-\infty + (-\infty) = -\infty$.

Beweis. Es seien

$$\begin{aligned} p &= a_n t^n + \cdots + a_1 t + a_0 \\ q &= b_m t^m + \cdots + b_1 t + b_0 \end{aligned}$$

mit $a_n \neq 0$ und $b_m \neq 0$, so dass $\mathrm{Grad}(p) = n$ und $\mathrm{Grad}(q) = m$. Wir können sicherlich $m \leq n$ voraussetzen.

Zu a). Wir erhalten $p + q = a_n t^n + \cdots + a_{m+1} t^{m+1} + (a_m + b_m) t^m + \cdots + (a_1 + b_1) t + (a_0 + b_0)$. Damit gilt $\mathrm{Grad}(p + q) \leq n$. Wir weisen darauf hin, dass „$<$“ genau dann auftritt, wenn $m = n$ und $a_m + b_m = 0$ gilt.

Zu b). Es gilt $p \cdot q = a_n b_m t^{n+m} + \cdots + (a_0 b_1 + a_1 b_0) t + a_0 b_0$. Da $a_n b_m \neq 0$, gilt $\mathrm{Grad}(p \cdot q) = m + n$. □

Satz V.2.4 (Division mit Rest)**.** *Es seien* $p, q \in K[t]$ *Polynome. Dann gibt es eindeutig bestimmte Polynome* s *und* r *mit* $\mathrm{Grad}(r) < \mathrm{Grad}(p)$*, so dass*

$$q = p \cdot s + r.$$

Definition V.2.5. Gilt in der Situation des Satzes $r = 0$, dann sagen wir, dass p das Polynom q teilt und schreiben $p|q$.

Beweis von Satz V.2.4. Für $\mathrm{Grad}(p) > \mathrm{Grad}(q)$ gilt $s = 0$ und $r = q$. Anstatt den Beweis auszuführen, geben wir ein Beispiel an:

$$\begin{array}{rcl} (2t^5 - 5t^4 + 7t^3 + 2t^2 + 8t - 5) : (t^2 - 3t + 5) & = & 2t^3 + t^2 + (-3) \\ -(2t^5 - 6t^4 + 10t^3) \downarrow & & \text{Rest } (-t + 10) \\ \hline t^4 - 3t^3 + 2t^2 \quad \Big\downarrow \text{ Übertrag} & & \\ -(t^4 - 3t^3 + 5t^2) & & \\ \hline -3t^2 + 8t - 5 & & \\ -(-3t^2 + 9t - 15) & & \\ \hline -t + 10. & & \end{array}$$

Der obige Satz wird z.B. in [9], Satz A.1.5, bewiesen. □

Definitionen V.2.6. i) Es sei $p = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0 \in K[t]$. Wir definieren

$$\begin{aligned} f_p \colon K &\longrightarrow K \\ \lambda &\longmapsto a_n \lambda^n + a_{n-1} \lambda^{n-1} + \cdots + a_1 \lambda + a_0. \end{aligned}$$

Die Abbildung f_p heißt die *zu* p *gehörige Polynomabbildung*.
Schreibweise. $p(\lambda) := f_p(\lambda)$.[2]

[2]Die Definition der Polynome über Abbildungen von $\mathbb{N}$ nach K erfolgte nur aus formalen Gründen und kann für das Weitere vergessen werden. In diesem Sinne führt die mißbräuchliche Schreibweise $p(\lambda)$ hoffentlich nicht zur Verwirrung.

ii) Eine Abbildung $f\colon K \longrightarrow K$ nennt man eine *Polynomabbildung*, wenn es ein Polynom $p \in K[t]$ mit $f = f_p$ gibt.

Satz V.2.7. i) *Die Zuordnung* $F\colon K[t] \longrightarrow \mathrm{Abb}(K,K)$, $p \longmapsto f_p$, *ist* K*-linear.*
ii) *Für* $\lambda \in K$ *und Polynome* $p, q \in K[t]$ *gilt* $f_{p\cdot q}(\lambda) = f_p(\lambda)\cdot f_q(\lambda)$.
iii) *Wenn* K *unendlich viele Elemente enthält, dann ist* F *injektiv.*

Beweis. Teil i) und ii) sind durch einfache Rechnungen leicht nachzuprüfen. Auf Punkt iii) kommen wir im Anschluss an den Beweis von Satz V.2.12 zu sprechen. □

Bemerkung V.2.8. Es sei $K = \{\lambda_1, ..., \lambda_s\}$ ein Körper mit nur endlich vielen Elementen. Das Polynom $p := (t-\lambda_1)\cdot\dots\cdot(t-\lambda_s)$ ist nicht das Nullpolynom, aber es gilt offensichtlich $f_p(\lambda) = 0$ für alle $\lambda \in K$. Damit ist F, wie im letzten Satz definiert, nicht injektiv.

Für $K = \mathbb{F}_2$ erhält man z.B. $p = t\cdot(t-1) = t\cdot(t+1) = t^2 + t$. Dabei ist $p(0) = 0+0 = 0$ und $p(1) = 1+1 = 0$.

Definition V.2.9. Es sei $p \in K[t]$ ein Polynom. Eine Zahl $\lambda \in K$ wird *Nullstelle von* p genannt, wenn $p(\lambda) = 0$.

Lemma V.2.10. *Wenn* λ *eine Nullstelle des Polynoms* $p \in K[t]$ *ist, dann gilt* $(t-\lambda)|p$. *Es gibt also ein Polynom* s *vom Grad* $\mathrm{Grad}(p)-1$, *so dass* $p = (t-\lambda)\cdot s$.

Sprechweise. Man spricht in diesem Zusammenhang von **der Abspaltung des Linearfaktors** $(t-\lambda)$**.**

Beweis von Lemma V.2.10. Wir führen die Division mit Rest durch: $p = (t-\lambda)\cdot s + r$. Da $\mathrm{Grad}(r) < \mathrm{Grad}(t-\lambda) = 1$, muss $r \in K$ gelten, und aus $0 = p(\lambda) = (\lambda-\lambda)\cdot s(\lambda) + r = r$ folgt $r = 0$. □

Definition V.2.11. Es seien $p \in K[t]$ ein Polynom mit $\mathrm{Grad}(p) \geq 0$ und $\lambda \in K$ eine Nullstelle von p. Dann heißt

$$\mu(p;\lambda) := \max\Big\{\, l \in \mathbb{N} \;\Big|\; (t-\lambda)^l | p \,\Big\} \geq 1$$

die *Vielfachheit* oder *Multiplizität der Nullstelle* λ.

Satz V.2.12. *Es sei* $p \in K[t]$ *ein Polynom mit* $\mathrm{Grad}(p) =: n \geq 0$ *und den (verschiedenen) Nullstellen* $\lambda_1, ..., \lambda_m$, $m \geq 0$. *Dann gibt es ein Polynom* $q \in K[t]$ *ohne Nullstellen, so dass*

$$p = (t-\lambda_1)^{l_1}\cdot\dots\cdot(t-\lambda_m)^{l_m}\cdot q. \tag{V.1}$$

Dabei haben wir $l_i := \mu(p,\lambda_i)$, $i = 1, ..., m$, *gesetzt. Insbesondere hat* p *höchtens* n *mit Vielfachheiten gezählte Nullstellen.*

Beweis. Wir führen Induktion über n. Für $n = 0$ gilt $p = a_0 \in K^*$, und es ist nichts zu zeigen.
Induktionsschritt $n \rightsquigarrow n+1$. Falls p keine Nullstelle hat, können wir $p = q$ wählen. Ansonsten sei $\lambda \in K$ so, dass $p(\lambda) = 0$. Gemäß Lemma V.2.10 schreiben wir $p = (t-\lambda) \cdot s$ für ein Polynom s vom Grad n. Nach Induktionsvoraussetzung gibt es also ein Polynom $q \in K[t]$ ohne Nullstellen, so dass

$$s = (t-\lambda_1)^{l'_1} \cdot \dots \cdot (t-\lambda_{m'})^{l'_{m'}} \cdot q.$$

Wenn $\lambda \notin \{\lambda_1, ..., \lambda_{m'}\}$, dann sind $\lambda_1, ..., \lambda_{m'}, \lambda_{m'+1} := \lambda$ die Nullstellen von p mit den Vielfachheiten $l'_1, ..., l'_{m'}, 1$. Falls (nach Umnummerierung) $\lambda = \lambda_1$, dann sind $\lambda_1, ..., \lambda_{m'}$ die Nullstellen von p mit den Vielfachheiten $l'_1 + 1, l'_2, ..., l'_{m'}$. In beiden Fällen besteht die Faktorisierung (V.1) wie behauptet. □

Beweis von Satz V.2.7, iii). Wenn f_p die Nullabbildung ist, dann hat p unendlich viele Nullstellen. Nach Satz V.2.12 kann p nur das Nullpolynom sein. □

Definition V.2.13. Ein Körper heißt *algebraisch abgeschlossen*, wenn jedes Polynom $p \in K[t]$ vom Grad $n \geq 1$ mindestens eine Nullstelle hat.

Beispiel V.2.14. Der Körper $\mathbb{R}$ der reellen Zahlen ist nicht algebraisch abgeschlossen. Das Polynom $p := t^2 + 1$ vom Grad zwei hat bekanntlich keine reelle Nullstelle.

Satz V.2.15. *Es seien K ein algebraisch abgeschlossener Körper und $p \in K[t]$ ein Polynom vom Grad $n \geq 1$ mit den verschiedenen Nullstellen $\lambda_1, ..., \lambda_m$. Dann existiert eine Zahl $a \in K^*$, so dass*

$$p = a \cdot (t-\lambda_1)^{l_1} \cdot \dots \cdot (t-\lambda_m)^{l_m}, \quad l_i := \mu(p, \lambda_i),\ i = 1, ..., m.$$

Sprechweise. Man sagt, das Polynom p zerfällt vollständig in Linearfaktoren.

Beweis von Satz V.2.15. Nach Satz V.2.12 können wir

$$p = (t-\lambda_1)^{l_1} \cdot \dots \cdot (t-\lambda_m)^{l_m} \cdot q$$

schreiben. Dabei ist q ein Polynom ohne Nullstellen. Auf Grund der Definition eines algebraisch abgeschlossenen Körpers muss $\mathrm{Grad}(q) = 0$ gelten, d.h. $q = a \in K^*$. □

Satz V.2.16 (Fundamentalsatz der Algebra)**.** *Der Körper $\mathbb{C}$ der komplexen Zahlen ist algebraisch abgeschlossen.*

Beweis. Elegante Beweise lernt man zumeist in der Funktionentheorie etwa im vierten Semester kennen. Die dort verwendeten Methoden sind völlig unabhängig von den im Folgenden entwickelten, so dass kein so genannter „Zirkelschluß“ zu befürchten ist. Man konsultiere etwa [11], Satz IV.5.11. □

Beispiel V.2.17 (Polynome vom Grad zwei). Es sei $p = at^2 + bt + c \in \mathbb{C}[t]$ ein Polynom vom Grad zwei, i.e. $a \neq 0$. Die Nullstellen λ_1 und $\lambda_2 \in \mathbb{C}$ werden nach der Formel

$$\lambda_{1/2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{V.2}$$

berechnet. Dabei überlegt man sich leicht, dass jede komplexe Zahl eine Wurzel besitzt (s. Aufgabe II.3.6).

Für $a, b, c \in \mathbb{R}$ beachte man, dass gilt:

$$\lambda_1 \in \mathbb{R} \iff \lambda_2 \in \mathbb{R} \iff b^2 - 4ac \geq 0.$$

Dabei ist $b^2 - 4ac = 0$ gleichbedeutend mit der Tatsache, dass p eine Nullstelle der Vielfachheit zwei besitzt (die automatisch reell ist). In jedem Fall hat p zwei Nullstellen (mit Vielfachheiten gezählt) im Körper $\mathbb{C}$.

Das allgemeine Verfahren. — Um für ein beliebiges Polynom $p \in K[t]$ vom Grad $n \geq 1$ eine mögliche Zerlegung in Linearfaktoren zu finden, gehe man wie folgt vor:

- Wenn p keine Nullstelle hat, dann gibt es keine solche Zerlegung, und wir sind fertig. Ansonsten bestimme man eine Nullstelle λ von p. (Dies ist natürlich ein schwieriges Problem. Für die Vorlesung, die Übungsaufgaben und die Klausur beschränken wir uns hierzu auf das „Eins-Zwei-Drei-Verfahren“, d.h. wir überprüfen, ob ± 1, ± 2 oder ± 3 vielleicht Nullstellen von p sind.) Wir bilden $p = (t - \lambda) \cdot s$.

- Nun stelle man fest, ob s eine Nullstelle besitzt. Falls nicht, brechen wir ab. Ansonsten bestimmen wir eine Nullstelle λ' und bilden $s = (t - \lambda') \cdot s'$, d.h. $p = (t - \lambda) \cdot (t - \lambda') \cdot s'$.

- Dieses Verfahren wird so lange iteriert, bis man sieht, dass keine weiteren Nullstellen vorkommen oder man ein Polynom vom Grad eins hat.
  Man beachte, dass man über dem Körper der komplexen Zahlen $\mathbb{C}$ für ein Polynom vom Grad 2 die Nullstellen mit (V.2) bestimmen kann.

Beispiel V.2.18. Es sei $p = t^4 - 5t^3 + 7t^2 + 3t - 10 \in \mathbb{C}[t]$. Es gilt $p(-1) = 0$. Daher führen wir die Division von p durch $t + 1$ aus:

$$
\begin{array}{r}
(t^4 - 5t^3 + 7t^2 + 3t - 10) : (t+1) = t^3 - 6t^2 + 13t - 10 =: s \\
\underline{-(t^4 + t^3)} \\
\quad -6t^3 + 7t^2 \\
\quad \underline{-(-6t^3 - 6t^2)} \\
\qquad\quad 13t^2 + 3t \\
\qquad \underline{-(13t^2 + 13t)} \\
\qquad\quad -10t - 10 \\
\qquad\quad \underline{-(-10t - 10)} \\
\qquad\qquad\quad 0.
\end{array}
$$

Als Nächstes stellt man fest, dass $s(2) = 0$ gilt. Die entsprechende Polynomdivision sieht wie folgt aus:

$$
\begin{array}{r}
(t^3 - 6t^2 + 13t - 10) : (t-2) = t^2 - 4t + 5 =: u \\
\underline{-(t^3 - 2t^2)} \\
\quad -4t^2 + 13t \\
\quad \underline{-(-4t^2 + 8t)} \\
\qquad\quad 5t - 10 \\
\qquad \underline{-(5t - 10)} \\
\qquad\qquad 0.
\end{array}
$$

Die Nullstellen λ_1 und λ_2 von u sind durch die Formel

$$\lambda_{1/2} = \frac{4 \pm \sqrt{16 - 20}}{2} = 2 \pm i$$

gegeben. Alles in allem bekommen wir die Faktorisierung

$$p = (t+1) \cdot (t-2) \cdot \big(t - (2-i)\big) \cdot \big(t - (2+i)\big).$$

Aufgabe V.2.19.
a) Führen Sie in $\mathbb{C}[t]$ die Polynomdivision mit Rest aus für die Polynome

$$p := 4\frac{1}{9}t^4 + (5+5i)t^3 + (4+9i)t^2 + (3+3i)t + 1 + i \quad \text{und} \quad q := \big(2 + \frac{1}{3}i\big)t^2 + 3t + 1.$$

b) Es seien $p \in \mathbb{R}[t]$ und $\lambda \in \mathbb{C}$ eine Nullstelle von p. Zeigen Sie, dass auch $\overline{\lambda}$ eine Nullstelle von p ist. Schließen Sie, dass

$$p = (t - \lambda_1)^{l_1} \cdot \cdots \cdot (t - \lambda_m)^{l_m} \cdot f_1 \cdot \cdots \cdot f_u$$

für gewisse Polynome $f_1, \ldots, f_u \in \mathbb{R}[t]$ vom Grad zwei. Hierbei seien $\lambda_1, \ldots, \lambda_m$ die reellen Nullstellen von p und $l_1, \ldots, l_m$ ihre Vielfachheiten.

Aufgabe V.2.20.
Kann ein endlicher Körper algebraisch abgeschlossen sein? Begründen Sie Ihre Antwort.

V.3 Charakteristische Polynome

Zunächst bemerken wir, dass wir Matrizen auch in einem allgemeineren Rahmen betrachten können. Dazu sei R ein kommutativer Ring (Kapitel II, Definition II.3.1).

Definitionen V.3.1. i) Für $m, n \in \mathbb{N}_{>0}$ ist die *Menge der* $(m \times n)$*-Matrizen mit Einträgen in* R definiert als

$$\mathrm{Mat}(m,n;R) := \left\{ A = (a_{ij})_{\substack{i=1,\dots,m\\ j=1,\dots,n}} \;\middle|\; a_{ij} \in R, i = 1,\dots,m, j = 1,\dots,n \right\}.$$

ii) Wir definieren die Abbildungen

- (*Addition*).

$$\begin{array}{rcl} +\colon \mathrm{Mat}(m,n;R) \times \mathrm{Mat}(m,n;R) & \longrightarrow & \mathrm{Mat}(m,n;R) \\ \left((a_{ij})_{\substack{i=1,\dots,m\\ j=1,\dots,n}}, (b_{ij})_{\substack{i=1,\dots,m\\ j=1,\dots,n}}\right) & \longmapsto & (a_{ij}+b_{ij})_{\substack{i=1,\dots,m\\ j=1,\dots,n}}. \end{array}$$

- (*Skalarmultiplikation*).

$$\begin{array}{rcl} \cdot_s\colon R \times \mathrm{Mat}(m,n;R) & \longrightarrow & \mathrm{Mat}(m,n;R) \\ \left(r, (a_{ij})_{\substack{i=1,\dots,m\\ j=1,\dots,n}}\right) & \longmapsto & (r \cdot a_{ij})_{\substack{i=1,\dots,m\\ j=1,\dots,n}}. \end{array}$$

- (*Matrixmultiplikation*).

$$\begin{array}{rcl} \cdot_m\colon \mathrm{Mat}(l,m;R) \times \mathrm{Mat}(m,n;R) & \longrightarrow & \mathrm{Mat}(l,n;R) \\ \left((a_{ij})_{\substack{i=1,\dots,l\\ j=1,\dots,m}}, (b_{ij})_{\substack{i=1,\dots,m\\ j=1,\dots,n}}\right) & \longmapsto & \left(\sum_{k=1}^{m} a_{ik} \cdot b_{kj}\right)_{\substack{i=1,\dots,l\\ j=1,\dots,n}}. \end{array}$$

- (*Determinantenabbildung*).

$$\begin{array}{rcl} \mathrm{Det}\colon \mathrm{Mat}(n;R) := \mathrm{Mat}(n,n;R) & \longrightarrow & R \\ A = (a_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}} & \longmapsto & \displaystyle\sum_{\sigma \in S_n} \mathrm{Sign}(\sigma) \cdot a_{1\sigma(1)} \cdot \dots \cdot a_{n\sigma(n)}. \end{array}$$

iii) Wie für einen Körper erklärt man so genannte *Elementarmatrizen* $S_i(a)$, $Q_{ij}(a)$, P_{ij}, $a \in R$, $1 \leq i \neq j \leq n$ (vgl. Definition IV.1.9).

Eigenschaften V.3.2. i) *Für* $(\mathrm{Mat}(m,n;R), +, \cdot_s)$ *sind die Vektorraumaxiome (Definition III.1.1) sinngemäß erfüllt.*[3]

ii) $(\mathrm{Mat}(n;R), +, \cdot_m)$ *ist ein Ring.*

iii) *Alle Eigenschaften der Determinante, die mit Hilfe der Leibniz-Formel bewiesen werden, bleiben gültig. Insbesondere gilt:*

[3]Man spricht in diesem Zusammenhang von einem **R-Modul**.

a) *Die Eigenschaften aus Definition IV.4.1, i), sind erfüllt.*

b) $\mathrm{Det}(A) = \mathrm{Det}(A^t)$ *(s. Satz IV.4.8).*

c) *Für ein Produkt* $S \in \mathrm{Mat}(n; R)$ *von Elementarmatrizen und eine beliebige Matrix* $A \in \mathrm{Mat}(n; R)$ *hat man*

$$\mathrm{Det}(S \cdot A) = \mathrm{Det}(S) \cdot \mathrm{Det}(A) \overset{R \text{ kommutativ}}{=} \mathrm{Det}(A) \cdot \mathrm{Det}(S) = \mathrm{Det}(A \cdot S). \quad \text{(V.3)}$$

d) *Es gibt zu jeder Matrix* $A \in \mathrm{Mat}(n, R)$ *die* komplementäre Matrix $\widetilde{A}$ *(vgl. Definition IV.4.13) mit* $\widetilde{A} \cdot A = A \cdot \widetilde{A} = \mathrm{Det}(A) \cdot \mathbb{E}_n$. *Insbesondere bleibt die Laplace-Entwicklung für Determinanten erlaubt.*

Bemerkung V.3.3. Bei allem, was mit multiplikativen Inversen und Invertieren zu tun hat, muss man vorsichtig sein. Z.B. gilt für die Elementarmatrix $S_1(2) \in \mathrm{Mat}(n; \mathbb{Z})$ zwar $\mathrm{Det}(S_1(2)) = 2 \neq 0$, dennoch ist $S_1(2)$ nicht invertierbar, d.h. es gibt keine Matrix $T \in \mathrm{Mat}(n; \mathbb{Z})$ mit $T \cdot S_1(2) = \mathbb{E}_n$. Dies hängt natürlich mit der Tatsache zusammen, dass 2 in $\mathbb{Z}$ nicht invertierbar ist (vgl. Aufgabe V.3.20).

Definitionen V.3.4. Es sei $A \in \mathrm{Mat}(n; K)$ eine $(n \times n)$-Matrix mit Einträgen im Körper K.

i) Dann ist

$$t \cdot \mathbb{E}_n - A \in \mathrm{Mat}(n; K[t]),$$

und

$$\chi_A(t) := \mathrm{Det}(t \cdot \mathbb{E}_n - A) \in K[t]$$

wird das *charakteristische Polynom von* A genannt.

ii) Die *Spur der Matrix* $A = (a_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$ ist die Zahl

$$\mathrm{Spur}(A) = a_{11} + \cdots + a_{nn}.$$

iii) Ein Polynom $p = a_n t^n + \cdots + a_0 \in K[t]$ vom Grad n ist *normiert*, wenn $a_n = 1$ gilt.

Bemerkung V.3.5. Mit der Leibniz-Formel überlegt man sich, dass das charakteristische Polynom von A die Gestalt

$$\chi_A(t) = t^n - \mathrm{Spur}(A) \cdot t^{n-1} + \cdots + (-1)^n \cdot \mathrm{Det}(A)$$

hat. Insbesondere ist $\chi_A(t)$ normiert vom Grad n.

Lemma V.3.6. *Für zwei ähnliche Matrizen* $A, B \in \mathrm{Mat}(n; K)$ *gilt*

$$\chi_A(t) = \chi_B(t).$$

Beweis. Nach Definition V.1.4 gibt es eine Matrix $S \in GL_n(K)$ mit $B = S \cdot A \cdot S^{-1}$. Damit erhält man

$$t \cdot \mathbb{E}_n - B = t \cdot \mathbb{E}_n - S \cdot A \cdot S^{-1} = S \cdot (t \cdot \mathbb{E}_n) \cdot S^{-1} - S \cdot A \cdot S^{-1} = S \cdot (t \cdot \mathbb{E}_n - A) \cdot S^{-1}.$$

Da S nach Satz IV.1.17 ein Produkt von Elementarmatrizen ist, gilt nach Gleichung (V.3)

$$\begin{aligned} \chi_B(t) &= \mathrm{Det}(t \cdot \mathbb{E}_n - B) = \mathrm{Det}\Big(S \cdot (t \cdot \mathbb{E}_n - A) \cdot S^{-1}\Big) \\ &= \mathrm{Det}(S) \cdot \mathrm{Det}(t \cdot \mathbb{E}_n - A) \cdot \mathrm{Det}(S^{-1}) \\ &= \mathrm{Det}(S) \cdot \mathrm{Det}(S^{-1}) \cdot \chi_A(t) = \chi_A(t). \end{aligned}$$

Dies ist die gewünschte Gleichung. □

Bemerkungen V.3.7. i) Im obigen Satz gilt nicht die Umkehrung. Es gibt Matrizen, die nicht ähnlich sind, aber die dennoch dasselbe charakteristische Polynom haben. Die Jordansche Normalform wird uns erlauben, genauere Aussagen darüber zu machen (s. Satz V.8.1 und Beispiel V.8.6).

ii) Die zu dem charakteristischen Polynom $\chi_A(t)$ einer Matrix $A \in \mathrm{Mat}(n; K)$ gehörige Polynomabbildung ist:

$$\begin{aligned} f_{\chi_A(t)} \colon K &\longrightarrow K \\ \lambda &\longmapsto \mathrm{Det}(\lambda \cdot \mathbb{E}_n - A). \end{aligned}$$

Auf Grund der Eigenschaften der Determinante sind folgende Aussagen äquivalent:

a) $\lambda \in K$ ist eine Nullstelle des charakteristischen Polynoms $\chi_A(t)$.

b) $\mathrm{Det}(\lambda \cdot \mathbb{E}_n - A) = 0$.

c) $\lambda \cdot \mathbb{E}_n - A$ ist nicht invertierbar.

d) $\mathrm{Rg}(\lambda \cdot \mathbb{E}_n - A) < n$.

e) $\mathbb{L}(\lambda \cdot \mathbb{E}_n - A, 0) \neq \{0\}$.

f) $\mathrm{Ker}(f_{\lambda \cdot \mathbb{E}_n - A}) \neq \{0\}$.

g) $\exists v \in K^n \setminus \{0\} : 0 = (\lambda \cdot \mathbb{E}_n - A) \cdot v = \lambda \cdot (\mathbb{E}_n \cdot v) - A \cdot v = \lambda \cdot v - A \cdot v$.

Die Äquivalenz von a) und g) führt zu dem folgenden wichtigen Resultat.

Satz V.3.8. *Die Eigenwerte einer Matrix* $A \in \mathrm{Mat}(n; K)$ *sind die Nullstellen des charakteristischen Polynoms* $\chi_A(t)$.

Bemerkung V.3.9. Satz V.3.8 gibt uns einen wichtigen Ansatz zur Berechnung der Eigenwerte einer gegebenen Matrix. Aus Lemma V.3.6 und Satz V.3.8 ergibt sich z.B., dass ähnliche Matrizen dieselben Eigenwerte haben. (Dies kann man sich allerdings auch mit elementaren Methoden überlegen (vgl. Bemerkung V.3.17).)

Das folgende Ergebnis beantwortet die Frage, welche Polynome als charakteristische Polynome auftreten können.

Satz V.3.10. *Es sei* $p(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0 \in K[t]$ *ein normiertes Polynom vom Grad* $n \geq 1$. *Dann existiert eine Matrix* $A \in \mathrm{Mat}(n; K)$ *mit*

$$\chi_A(t) = p(t).$$

Beweis. Wir beweisen folgende Aussage durch vollständige Induktion über n:

Es seien $n \geq 1$ *eine natürliche Zahl,* $a_0, \ldots, a_{n-1} \in K$ *Elemente des Körpers und*

$$A := (b_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}} := \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & -a_{n-2} \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix} \in \mathrm{Mat}(n; K),$$

d.h. $b_{i+1,i} = 1$, $i = 1, \ldots, n-1$, $b_{in} = -a_{i-1}$, $i = 1, \ldots, n$, *und* $b_{ij} = 0$ *in allen anderen Fällen. Dann gilt*

$$\chi_A(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0.$$

Für $n = 1$ gilt offenbar $\chi_A(t) = t + a_0$. Für den **Induktionsschritt „$n \rightsquigarrow n+1$“** entwickeln wir die Determinante der Matrix

$$\begin{pmatrix} t & 0 & \cdots & 0 & a_0 \\ -1 & t & \cdots & 0 & a_1 \\ 0 & -1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & t & a_{n-1} \\ 0 & \cdots & 0 & -1 & t + a_n \end{pmatrix} \in \mathrm{Mat}\big(n+1; K[t]\big)$$

nach der ersten Zeile, so dass

$$\begin{aligned} \chi_A(t) &= t \cdot \mathrm{Det} \begin{pmatrix} t & 0 & \cdots & 0 & a_1 \\ -1 & t & \cdots & 0 & a_2 \\ 0 & -1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & t & a_{n-1} \\ 0 & \cdots & 0 & -1 & t + a_n \end{pmatrix} + (-1)^{n+2} a_0 \underbrace{\mathrm{Det} \begin{pmatrix} -1 & t & 0 & \cdots & 0 \\ 0 & -1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -1 & t \\ 0 & \cdots & \cdots & 0 & -1 \end{pmatrix}}_{=(-1)^n} \\ &\overset{\text{IV}}{=} t \cdot (t^n + a_n t^{n-1} + \cdots + a_2 t + a_1) + a_0 \\ &= t^{n+1} + a_n t^n + \cdots + a_2 t^2 + a_1 t + a_0. \end{aligned}$$

Die Abkürzung „IV“ steht dabei für „Induktionsvoraussetzung“. □

Bemerkung V.3.11. Wir weisen nochmals drauf hin, dass die Matrix A bei weitem nicht die einzige Matrix mit dem angegebenen charakteristischen Polynom ist.

Für jedes Polynom $p = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0$ mit $a_n \neq 0$ ist $(1/a_n) \cdot p$ ein normiertes Polynom mit denselben Nullstellen wie p. Angesichts der Definition eines algebraisch abgeschlossenen Körpers V.2.13 schließen wir mit dem letzten Satz:

Folgerung V.3.12. *Für einen Körper K sind folgende Eigenschaften äquivalent:*

a) *K ist algebraisch abgeschlossen.*

b) *Jede Matrix $A \in \mathrm{Mat}(n; K)$ mit $n \geq 1$ besitzt einen Eigenwert.*

Dies ist eine wichtige Beobachtung: Die Analyse eines Endomorphismus $f\colon V \longrightarrow V$ eines endlichdimensionalen K-Vektorraums bzw. einer Matrix $A \in \mathrm{Mat}(n; K)$ wollen wir auf Eigenwerte und zugehörige Eigenvektoren stützen. Dazu muss es natürlich genügend Eigenwerte und -vektoren geben. Die obige Folgerung sagt uns, dass wir dies im Allgemeinen nur erwarten können, wenn der Körper K algebraisch abgeschlossen ist. Zum ersten Mal gehen also Eigenschaften des Grundkörpers in unsere Theorie ein.

Beispiel V.3.13. Es sei $K = \mathbb{R}$. Wie wir wissen, ist $\mathbb{R}$ nicht algebraisch abgeschlossen, so dass es Matrizen ohne Eigenwert geben muss. Ein intuitives Beispiel, das dieses Phänomen erklärt, ist das der Matrix[4]

$$A := \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}, \quad 0 \leq \alpha < 2\pi.$$

Wir erinnern an Beispiel III.5.19. Die zu A gehörige lineare Abbildung $f_A\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ ist die Drehung um den Winkel α. Ein Vektor $v \in \mathbb{R}^2 \setminus \{0\}$ ist genau dann ein Eigenvektor, wenn er durch f_A auf ein Vielfaches von sich selbst abgebildet wird oder, anders gesagt, die von v erzeugte Gerade durch f_A auf sich selbst abgebildet wird. Für eine Drehung kann dies nur für $\alpha = 0$ (identische Abbildung; $A = \mathbb{E}_2$) oder $\alpha = \pi$ (Drehung um 180 Grad; $A = -\mathbb{E}_2$) passieren. In den letzten beiden Fällen haben wir nur den Eigenwert $\lambda = 1$ bzw. $\lambda = -1$ und $\mathrm{Eig}(A, \lambda) = \mathbb{R}^2$. Für $\alpha \in [0, 2\pi) \setminus \{0, \pi\}$ erwarten wir keinen Eigenwert.

[4] Für die verwendeten Eigenschaften der trigonometrischen Funktionen verweisen wir auf [8].

Rechnerisch ergibt sich

$$\begin{aligned}\chi_A(t) &= \operatorname{Det}\begin{pmatrix} t-\cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & t-\cos(\alpha)\end{pmatrix} \\ &= (t-\cos(\alpha))^2+\sin^2(\alpha) \\ &= t^2-2\cos(\alpha)\cdot t+\underbrace{\cos^2(\alpha)+\sin^2(\alpha)}_{=1} = t^2-2\cos(\alpha)\cdot t+1.\end{aligned}$$

Die Nullstellen von $\chi_A(t)$ sind nach (V.2) durch $\cos(\alpha)\pm\sqrt{\cos^2(\alpha)-1}$ gegeben. Da $\cos^2(\alpha)\in[0,1]$, existiert eine reelle Nullstelle nur für $\cos(\alpha)=\pm 1$. Da wir $\alpha\in[0,2\pi)$ voraussetzen, gilt dies für $\alpha=0$ und $\alpha=\pi$ in Übereinstimmung mit unseren geometrischen Überlegungen.

Schließlich seien V ein endlichdimensionaler K-Vektorraum und $f\colon V\longrightarrow V$ ein Endomorphismus. Wie wir bereits des Öfteren gesehen haben, ist die Darstellungsmatrix von f bis auf Ähnlichkeit bestimmt, d.h. sind zwei durchnummerierte Basen B und B' von V gegeben, dann folgt $M_B(f)\approx M_{B'}(f)$. Die folgende Definition ist also wegen Lemma V.3.6 sinnvoll.

Definition V.3.14. Für V und f wie oben setzt man

$$\chi_f(t) := \chi_{M_B(f)}(t).$$

Dabei sei B eine durchnummerierte Basis von V. Man nennt $\chi_f(t)$ das *charakteristische Polynom von* f.

Eigenschaften V.3.15. *Wie für das charakteristische Polynom einer Matrix hat man:*

$$\forall\lambda\in K:\chi_f(\lambda)=0\Longleftrightarrow\lambda \textit{ ist Eigenwert von } f.$$

Definitionen V.3.16. Es seien V ein endlichdimensionaler K-Vektorraum und $f\colon V\longrightarrow V$ ein Endomorphismus bzw. $A\in\operatorname{Mat}(n;K)$. Ferner sei $\lambda\in K$ ein Eigenwert von f bzw. A.

i) Die Zahl

$$\begin{aligned}\alpha(f,\lambda) &:= \mu(\chi_f(t);\lambda) \text{ (s. Definition V.2.11)} \\ \text{bzw. } \alpha(A,\lambda) &:= \mu(\chi_A(t);\lambda)\end{aligned}$$

heißt die *algebraische Vielfachheit des Eigenwerts* λ.

ii) Die Zahl

$$\begin{aligned}\gamma(f,\lambda) &:= \operatorname{Dim}_K(\operatorname{Eig}(f,\lambda)) \text{ (s. Definition V.1.7)} \\ \text{bzw. } \gamma(A,\lambda) &:= \operatorname{Dim}_K(\operatorname{Eig}(A,\lambda))\end{aligned}$$

heißt die *geometrische Vielfachheit des Eigenwerts* λ.

Bemerkung V.3.17. Es seien A und $B \in \mathrm{Mat}(n;K)$ ähnliche Matrizen. Nach Lemma V.3.6 gilt $\chi_A(t) = \chi_B(t)$. Damit haben A und B dieselben Eigenwerte, und für jeden Eigenwert $\lambda \in K$ gilt $\alpha(A,\lambda) = \alpha(B,\lambda)$. Es gilt aber auch $\gamma(A,\lambda) = \gamma(B,\lambda)$. Denn für $S \in \mathrm{GL}_n(K)$ mit $B = S \cdot A \cdot S^{-1}$ und $v \in \mathrm{Eig}(A,\lambda)$ gilt

$$B \cdot (S \cdot v) = S \cdot A \cdot S^{-1} \cdot S \cdot v = S \cdot A \cdot v = S \cdot (\lambda \cdot v) = \lambda \cdot (S \cdot v),$$

i.e. $S \cdot v \in \mathrm{Eig}(B,\lambda)$ und $S \cdot \mathrm{Eig}(A,\lambda) \subseteq \mathrm{Eig}(B,\lambda)$. Ebenso gilt $S^{-1} \cdot \mathrm{Eig}(B,\lambda) \subseteq \mathrm{Eig}(A,\lambda)$. Damit induziert die Multiplikation mit S von links einen Isomorphismus zwischen $\mathrm{Eig}(A,\lambda)$ und $\mathrm{Eig}(B,\lambda)$, so dass die Aussage über die Gleichheit der geometrischen Vielfachheiten folgt. Auch erhält man einen elementaren Beweis dafür, dass A und B dieselben Eigenwerte haben.

Die geometrische Vielfachheit gibt die maximale Anzahl linear unabhängiger Eigenvektoren zum Eigenwert λ an und wird durch die algebraische Vielfachheit abgeschätzt:

Satz V.3.18. *In der obigen Situation gilt immer*

$$\gamma(f,\lambda) \leq \alpha(f,\lambda) \quad \textit{bzw.} \quad \gamma(A,\lambda) \leq \alpha(A,\lambda).$$

Beweis. Wir führen den Beweis für einen Endomorphismus f. Dazu seien $r := \gamma(f,\lambda)$ und $b_1, ..., b_r$ eine Basis für $\mathrm{Eig}(f,\lambda)$. Man ergänze zu einer durchnummerierten Basis $B = \{ b_1, ..., b_r, b_{r+1}, ..., b_n \}$ von V. Für diese Basis gilt

$$M_B(f) = \left(\begin{array}{c|c} \lambda \cdot \mathbb{E}_r & * \\ \hline 0 & A' \end{array}\right).$$

Man berechnet $\chi_f(t) = (t-\lambda)^r \cdot \chi_{A'}(t)$. Also teilt $(t-\lambda)^r$ das charakteristische Polynom, so dass

$$\alpha(f,\lambda) = \max\left\{ l \in \mathbb{N} \,\middle|\, (t-\lambda)^l \text{ teilt } \chi_f(t) \right\} \geq r$$

wie behauptet folgt. □

Beispiel V.3.19. Wie man sich leicht überlegt, gilt $\alpha(A,\lambda) = \gamma(A,\lambda)$ für jede diagonalisierbare Matrix A und jeden Eigenwert λ von A (s. auch Satz V.4.1). Für die Matrix $A := \begin{pmatrix} 2 & -1 \\ 0 & 2 \end{pmatrix} \in \mathrm{Mat}(2;\mathbb{Q})$ berechnet man $\chi_A(t) = (t-2)^2$. Damit ist $\lambda = 2$ der einzige Eigenwert, und wir haben $\alpha(A,2) = 2$. Wegen $2 \cdot \mathbb{E}_2 - A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ gilt

$$\mathrm{Eig}(A,2) = \mathbb{L}(2 \cdot \mathbb{E}_2 - A, 0) = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle,$$

also $\gamma(A,2) = 1 < \alpha(A,2)$. Somit ist A insbesondere nicht diagonalisierbar.

Aufgabe V.3.20 (Matrizen über $\mathbb{Z}$).
Es sei $n \geq 1$ eine natürliche Zahl.
a) Zeigen Sie, dass für $A, B \in \mathrm{Mat}(n, \mathbb{Z})$

$$\mathrm{Det}(A \cdot B) = \mathrm{Det}(A) \cdot \mathrm{Det}(B)$$

gilt.
b) Schließen Sie, dass eine Matrix $A \in \mathrm{Mat}(n; \mathbb{Z})$ genau dann invertierbar ist,[5] wenn

$$\mathrm{Det}(A) \in \{\pm 1\}.$$

c) Zeigen Sie, dass eine invertierbare Matrix $A \in \mathrm{Mat}(n; \mathbb{Z})$ als Produkt von Elementarmatrizen aus $\mathrm{Mat}(n; \mathbb{Z})$ geschrieben werden kann.

Aufgabe V.3.21.
Zeigen Sie unter Verwendung der Formel für das Matrixprodukt, dass für alle $A, B \in \mathrm{Mat}(n; K)$ gilt

$$\mathrm{Spur}(A \cdot B) = \mathrm{Spur}(B \cdot A).$$

Aufgabe V.3.22.
Berechnen Sie das charakteristische Polynom der komplexen Matrix

$$A := \begin{pmatrix} 1 & 0 & 0 & -25 \\ 0 & 0 & 0 & 26 \\ -3 & 1 & 0 & 70 \\ 0 & 0 & 1 & 5 \end{pmatrix}.$$

V.4 Diagonalisierbare Endomorphismen

Wir können mit den zuletzt eingeführten Begriffen das elementare Diagonalisierungskriterium aus Lemma V.1.9 präzisieren.

Satz V.4.1. *Es seien V ein n-dimensionaler K-Vektorraum und $f\colon V \longrightarrow V$ ein Endomorphismus bzw. $A \in \mathrm{Mat}(n; K)$. Dann sind die folgenden Bedingungen äquivalent:*

a) *Der Endomorphismus f bzw. die Matrix A ist diagonalisierbar.*

b) A) *Das charakteristische Polynom $\chi_f(t)$ bzw. $\chi_A(t)$ zerfällt vollständig in Linearfaktoren und* B) *für jeden Eigenwert λ von f bzw. A gilt $\alpha(f, \lambda) = \gamma(f, \lambda)$ bzw. $\alpha(A, \lambda) = \gamma(A, \lambda)$.*

[5] d.h. es gibt eine Matrix $B \in \mathrm{Mat}(n, \mathbb{Z})$ mit $A \cdot B = B \cdot A = \mathbb{E}_n$

c) *Es gilt*

$$V = \mathrm{Eig}(f, \lambda_1) \oplus \cdots \oplus \mathrm{Eig}(f, \lambda_m)$$
$$\textit{bzw.} \qquad K^n = \mathrm{Eig}(A, \lambda_1) \oplus \cdots \oplus \mathrm{Eig}(A, \lambda_m).$$

Dabei seien $\lambda_1, ..., \lambda_m \in K$ *die verschiedenen Eigenwerte von* f *bzw.* A*.*

Beweis. Wir beweisen den Satz für einen Endomorphismus f und fangen mit der Implikation **„a)⟹b)“** an. Nach Lemma V.1.9 können wir eine durchnummerierte Basis $B = \{ b_1, ..., b_n \}$ von V wählen mit

$$M_B(f) = \begin{pmatrix} \lambda_1 \cdot \mathbb{E}_{l_1} & & & \\ & \lambda_2 \cdot \mathbb{E}_{l_2} & & 0 \\ 0 & & \ddots & \\ & & & \lambda_m \cdot \mathbb{E}_{l_m} \end{pmatrix}. \tag{V.4}$$

Dabei seien $\lambda_1, ..., \lambda_m \in K$ die verschiedenen Eigenwerte von f. Offenbar gilt $\gamma(f, \lambda_i) \geq l_i$, $i = 1, ..., m$. Aus (V.4) ergibt sich das charakteristische Polynom

$$\chi_f(t) = (t - \lambda_1)^{l_1} \cdot \cdots \cdot (t - \lambda_m)^{l_m}$$

und somit A). Es folgt

$$l_i = \alpha(f, \lambda_i) \geq \gamma(f, \lambda_i) \geq l_i, \quad i = 1, ..., m,$$

so dass tatsächlich $\alpha(f, \lambda) = \gamma(f, \lambda)$ für jeden Eigenwert λ von f wie in B) behauptet gilt.

Zu „b)⟹c)“. Wir setzen zunächst

$$W := \mathrm{Eig}(f, \lambda_1) + \cdots + \mathrm{Eig}(f, \lambda_m).$$

Nun seien $v_i \in \mathrm{Eig}(f, \lambda_i)$, $i = 1, ..., m$, Eigenvektoren mit $v_1 + \cdots + v_m = 0$. Nach Aufgabe V.1.12 muss also $v_1 = \cdots = v_m = 0$ sein, so dass

$$W = \mathrm{Eig}(f, \lambda_1) \oplus \cdots \oplus \mathrm{Eig}(f, \lambda_m) \quad \text{(Definition III.2.1, ii)}.$$

Aus der Dimensionsformel (Satz III.5.40) folgt leicht

$$\begin{aligned} \mathrm{Dim}_K(W) &= \mathrm{Dim}_K(\mathrm{Eig}(f, \lambda_1)) + \cdots + \mathrm{Dim}_K(\mathrm{Eig}(f, \lambda_m)) \\ &= \gamma(f, \lambda_1) + \cdots + \gamma(f, \lambda_m) \\ &\overset{B)}{=} \alpha(f, \lambda_1) + \cdots + \alpha(f, \lambda_m) \overset{A)}{=} n. \end{aligned}$$

Aus $\mathrm{Dim}_K(W) = \mathrm{Dim}_K(V)$ folgt sofort $W = V$.

Zu „c)⟹a)“. Man wähle Basen $v_1^j, ..., v_{l_j}^j$ für $\mathrm{Eig}(f, \lambda_j)$, $j = 1, ..., m$. Dann ist

$$B := \{ b_1, ..., b_n \} := \{ v_1^1, ..., v_{l_1}^1, v_1^2, ..., v_{l_2}^2, ..., v_1^m, ..., v_{l_m}^m \}$$

eine durchnummerierte Basis von V, bzgl. derer die Darstellungsmatrix durch (V.4) gegeben ist. □

Folgerung V.4.2. *Ein Endomorphismus f bzw. eine Matrix A wie in Satz V.4.1 mit einem charakteristischen Polynom $\chi_f(t)$ bzw. $\chi_A(t)$, das vollständig in Linearfaktoren zerfällt und nur Nullstellen der Vielfachheit eins (so genannte **einfache Nullstellen**) besitzt, ist diagonalisierbar.*

Bemerkungen V.4.3. i) Falls K algebraisch abgeschlossen ist, dann ist Bedingung A) in Satz V.4.1, b), immer erfüllt (s. Satz V.2.15).

ii) Wir haben nun eine Klasse von Endomorphismen bzw. Matrizen intrinsisch ausgezeichnet, für die unser anfangs gestelltes Problem V.1.6 eine einfache Lösung besitzt. Für die durch diesen Satz nicht abgedeckten Endomorphismen bzw. Matrizen werden wir eine nicht unerhebliche Zusatzarbeit zu leisten haben.

iii) Zwei Diagonalmatrizen $A = \mathrm{Diag}(a_1, ..., a_n)$ und $B = \mathrm{Diag}(b_1, ..., b_n)$ sind genau dann ähnlich, wenn es eine Permutation $\sigma \in S_n$ mit $a_i = b_{\sigma(i)}$, $i = 1, ..., n$, gibt.

iv) Wir bemerken, dass ein „allgemeines" Polynom $p \in \mathbb{C}[t]$ nur einfache Nullstellen hat und dass das Vorkommen mehrfacher Nullstellen eine spezielle Situation darstellt. Dies können wir uns für Polynome $p = at^2 + bt + c \in \mathbb{C}[t]$ mit $a \neq 0$ klarmachen. Nach der Lösungsformel (V.2) tritt eine doppelte Nullstelle genau dann auf, wenn $b^2 - 4ac = 0$ gilt. Für allgemeine Wahlen von a, b und c wird diese Beziehung nicht gelten, und p wird zwei einfache Nullstellen haben. Auch für Polynome höheren Grades wird das Vorkommen einer mehrfachen Nullstelle durch das Verschwinden einer so genannten **Diskriminante** charakterisiert. Man konsultiere etwa:
B.L. van der Waerden, *Algebra* I, Siebte Auflage, Heidelberger Taschenbücher **12**, Springer-Verlag, Berlin-New York, 1966, xi+271 Seiten.
In diesem Sinne hat auch eine „allgemeine" Matrix $A \in \mathrm{Mat}(n; \mathbb{C})$ genau n verschiedene Eigenwerte und ist damit diagonalisierbar.

Das Verfahren zur Diagonalisierung.

Schritt 1. — Man stelle fest, ob das charakteristische Polynom vollständig in Linearfaktoren zerfällt, und bestimme ggfs. die Zerlegung in Linearfaktoren. Falls es nicht zerfällt, dann ist f bzw. A nicht diagonalisierbar.

Z.B.

$$\begin{array}{ccc} f\colon \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} & \longmapsto & \begin{pmatrix} -s_2 + s_3 \\ -3s_1 - 2s_2 + 3s_3 \\ -2s_1 - 2s_2 + 3s_3 \end{pmatrix}. \end{array}$$

Bzgl. der Standardbasis $E = \{e_1, e_2, e_3\}$ von $\mathbb{R}^3$ gilt:

$$M_E(f) = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix}.$$

Mit der Regel von Sarrus

$$\begin{array}{rrrrr} t & 1 & -1 & t & 1 \\ 3 & t+2 & -3 & 3 & t+2 \\ 2 & 2 & t-3 & 2 & 2 \end{array}$$

folgt

$$\chi_f(t) = t(t+2)(t-3) - 6 - 6 + 2t + 4 + 6t - 3t + 9 = t(t^2 - t - 6) + 5t + 1 = t^3 - t^2 - t + 1.$$

Offenbar ist $\lambda = 1$ eine Nullstelle.

$$\begin{array}{r} (t^3 - t^2 - t + 1) : (t-1) = t^2 - 1 = (t+1)\cdot(t-1) \\ \underline{-(t^3 - t^2)} \\ -t + 1 \\ \underline{-(-t+1)} \\ 0, \end{array}$$

also

$$\chi_f(t) = (t-1)^2 \cdot (t+1).$$

Schritt 2. — Man stelle durch Berechnung der Eigenräume fest, ob die algebraischen Multiplizitäten mit den geometrischen übereinstimmen. Falls nicht, dann ist f bzw. A nicht diagonalisierbar.

$$\mathrm{Eig}(f, 1) = \mathbb{L}(\mathbb{E}_3 - M_E(f), 0) = \mathbb{L}\left(\begin{pmatrix} 1 & 1 & -1 \\ 3 & 3 & -3 \\ 2 & 2 & -2 \end{pmatrix}, 0\right).$$

Damit ist leicht zu sehen, dass

$$\mathrm{Eig}(f, 1) = \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\rangle.$$

Also

$$\alpha(f, 1) = 2 = \gamma(f, 1).$$

Offenbar gilt automatisch

$$\alpha(f,-1) = 1 = \gamma(f,-1).$$

Damit ist f diagonalisierbar. Wir berechnen noch $\mathrm{Eig}(f,-1)$. Es gilt

$$\mathrm{Eig}(f,-1) = \mathbb{L}(-\mathbb{E}_3 - M_E(f), 0) = \mathbb{L}\left(\left(\begin{array}{ccc} -1 & 1 & -1 \\ 3 & 1 & -3 \\ 2 & 2 & -4 \end{array}\right), 0\right).$$

Der Gauß-Algorithmus liefert

$$\left(\begin{array}{ccc} -1 & 1 & -1 \\ 3 & 1 & -3 \\ 2 & 2 & -4 \end{array}\right) \overset{II+3\cdot I; III+2\cdot I}{\rightsquigarrow} \left(\begin{array}{ccc} -1 & 1 & -1 \\ 0 & 4 & -6 \\ 0 & 4 & -6 \end{array}\right) \overset{III-II}{\rightsquigarrow} \left(\begin{array}{ccc} -1 & 1 & -1 \\ 0 & 4 & -6 \\ 0 & 0 & 0 \end{array}\right).$$

Wir wählen $s_3 = 2s$. Es folgt $s_2 = 3s$ und $s_1 = s$, also

$$\mathrm{Eig}(f,-1) = \left\langle \left(\begin{array}{c} 1 \\ 3 \\ 2 \end{array}\right) \right\rangle.$$

Schritt 3. — Aus Schritt 2 erhält man eine Basis $B = \{b_1, ..., b_n\}$ aus Eigenvektoren zu den Eigenwerten $\lambda_1, ..., \lambda_m$. Dann ist $M_B(f)$ eine Diagonalmatrix. Für $V = K^n$ und $f = f_A$ sei $S^{-1} := T := (b_1|b_2|\cdots|b_n)$ die Matrix mit den Spalten $b_1, ..., b_n$. Dann ist

$$S \cdot A \cdot S^{-1} = T^{-1} \cdot A \cdot T = \left(\begin{array}{ccc} \lambda_1 \cdot \mathbb{E}_{l_1} & & \\ & \lambda_2 \cdot \mathbb{E}_{l_2} & \\ 0 & & \ddots \\ & & \lambda_m \cdot \mathbb{E}_{l_m} \end{array}\right).$$

In unserem Beispiel

$$T = \left(\begin{array}{ccc} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{array}\right).$$

Wir bestimmen T^{-1} mit dem Gauß-Verfahren:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{array}\right) \overset{III-I}{\rightsquigarrow} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & 0 & 1 & 0 \\ 0 & 1 & 1 & -1 & 0 & 1 \end{array}\right) \overset{III-II}{\rightsquigarrow} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & 0 & 1 & 0 \\ 0 & 0 & -2 & -1 & -1 & 1 \end{array}\right)$$

$$\overset{I+\frac{1}{2}\cdot III; II+\frac{3}{2}\cdot III; -\frac{1}{2}\cdot III}{\rightsquigarrow} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & -\frac{3}{2} & -\frac{1}{2} & \frac{3}{2} \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{array}\right),$$

also

$$T^{-1} = \frac{1}{2} \cdot \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ 1 & 1 & -1 \end{pmatrix}.$$

Kontrolle.

$$\begin{aligned} & \frac{1}{2} \cdot \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ 1 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix} \\ = & \frac{1}{2} \cdot \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ 1 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -3 \\ 1 & 1 & -2 \end{pmatrix} \\ = & \frac{1}{2} \cdot \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} . \surd \end{aligned}$$

Aufgabe V.4.4.
Gegeben sei die Matrix

$$A := \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix} \in \mathrm{Mat}(2;\mathbb{R}), \quad 0 \leq \alpha < 2\pi.$$

a) Berechnen Sie das charakteristische Polynom von A und zeigen Sie, dass A für jeden Wert $\alpha \in [0, 2\pi)$ diagonalisierbar ist.
b) Berechnen Sie die Eigenräume von A.
c) Interpretieren Sie Ihre Ergebnisse geometrisch in der Ebene $\mathbb{R}^2$.

Aufgabe V.4.5.
Es sei

$$A := \frac{1}{4} \cdot \begin{pmatrix} 8 & 0 & 0 & 0 \\ 0 & -3 & -6 & -7 \\ 0 & 0 & 8 & 0 \\ 0 & 11 & 6 & 15 \end{pmatrix} \in \mathrm{Mat}(4, \mathbb{R}).$$

a) Berechnen Sie das charakteristische Polynom von A und zerlegen Sie es in Linearfaktoren.
b) Zeigen Sie, dass A über $\mathbb{R}$ diagonalisierbar ist.
c) Geben Sie eine Matrix S an, so dass $S \cdot A \cdot S^{-1}$ eine Diagonalmatrix ist.

V.5 Der Satz von Cayley-Hamilton

Lemma V.5.1. i) *Es sei V ein endlichdimensionaler K-Vektorraum. Zu jedem Endomorphismus $f \in \mathrm{End}_K(V)$ gibt es genau eine Abbildung*

$$\mathrm{ev}_f \colon K[t] \longrightarrow \mathrm{End}_K(V),$$

so dass gilt:

a) ev_f *ist* K*-linear.*

b) $\mathrm{ev}_f(1) = \mathrm{Id}_V$.

c) $\mathrm{ev}_f(t) = f$.

d) *Für alle* $p, q \in K[t]$ *gilt* $\mathrm{ev}_f(p \cdot q) = \mathrm{ev}_f(p) \circ \mathrm{ev}_f(q)$.

ii) *Zu jeder Matrix* $A \in \mathrm{Mat}(n; K)$ *gibt es genau eine Abbildung*

$$\mathrm{ev}_A \colon K[t] \longrightarrow \mathrm{Mat}(n; K)$$

mit den folgenden Eigenschaften:

a) ev_A *ist* K*-linear.*

b) $\mathrm{ev}_A(1) = \mathbb{E}_n$.

c) $\mathrm{ev}_A(t) = A$.

d) *Für alle* $p, q \in K[t]$ *gilt* $\mathrm{ev}_A(p \cdot q) = \mathrm{ev}_A(p) \cdot \mathrm{ev}_A(q)$.

Schreibweise. $p(A) := \mathrm{ev}_A(p)$; $p(f) := \mathrm{ev}_f(p)$.

Definitionen V.5.2. i) Es seien V ein endlichdimensionaler K-Vektorraum und $f \colon V \longrightarrow V$ ein Endomorphismus. Die *Potenzen* von f werden rekursiv durch

- $f^0 := \mathrm{Id}_V$,
- $f^{n+1} := f \circ f^n$, $n \in \mathbb{N}$,

erklärt.

ii) Es seien $n \geq 1$ und A eine $(n \times n)$-Matrix. Die *Potenzen* von A werden rekursiv vermöge

- $A^0 := \mathbb{E}_n$,
- $A^{n+1} := A \cdot A^n$, $n \in \mathbb{N}$,

definiert.

Beweis von Lemma V.5.1. Wir beweisen ii). Wir zeigen zunächst die Eindeutigkeit der Abbildung ev_A. Es sei also $\mathrm{ev}_A \colon K[t] \longrightarrow \mathrm{Mat}(n; K)$ eine Abbildung mit

den Eigenschaften a) - d). Damit gilt für ein Polynom $p = a_m t^m + a_{m-1} t^{m-1} + \cdots + a_1 t + a_0$:

$$\begin{aligned} \mathrm{ev}_A(p) &\overset{\text{a)}}{=} a_m \cdot \mathrm{ev}_A(t^m) + a_{m-1} \cdot \mathrm{ev}_A(t^{m-1}) + \cdots + a_1 \cdot \mathrm{ev}_A(t) + a_0 \cdot \mathrm{ev}_A(1) \\ &\overset{\text{d)}}{=} a_m \cdot \mathrm{ev}_A(t)^m + a_{m-1} \cdot \mathrm{ev}_A(t)^{m-1} + \cdots + a_1 \cdot \mathrm{ev}_A(t) + a_0 \cdot \mathrm{ev}_A(1) \\ &\overset{\text{b)\&c)}}{=} a_m \cdot A^m + a_{m-1} \cdot A^{m-1} + \cdots + a_1 \cdot A + a_0 \cdot \mathbb{E}_n. \end{aligned}$$

Damit ist der Wert ev_A in jedem Polynom $p \in K[t]$ festgelegt, und die Eindeutigkeit folgt.

Umgekehrt können wir die Abbildung $\mathrm{ev}_A \colon K[t] \longrightarrow \mathrm{Mat}(n; K)$ durch die Formel

$$\mathrm{ev}_A(a_m t^m + a_{m-1} t^{m-1} + \cdots + a_1 t + a_0) := a_m \cdot A^m + a_{m-1} \cdot A^{m-1} + \cdots + a_1 \cdot A + a_0 \cdot \mathbb{E}_n$$

definieren. Dann sind die Eigenschaften b) und c) offenkundig erfüllt, und a) und d) kann man leicht nachrechnen. Für Eigenschaft d) z.B. seien $p := a_m t^m + a_{m-1} t^{m-1} + \cdots + a_1 t + a_0$ und $q = b_l t^l + b_{l-1} t^{l-1} + \cdots + b_1 t + b_0$ zwei Polynome. Dann erhält man

$$\begin{aligned} \mathrm{ev}_A(p) \cdot \mathrm{ev}_A(q) &= \Big(a_m \cdot A^m + a_{m-1} \cdot A^{m-1} + \cdots + a_1 \cdot A + a_0 \cdot \mathbb{E}_n\Big) \cdot \\ &\quad \cdot \Big(b_l \cdot A^l + b_{l-1} \cdot A^{l-1} + \cdots + b_1 \cdot A + b_0 \cdot \mathbb{E}_n\Big) \\ &\overset{\text{Distributivgesetz}}{=} \sum_{i=0}^{m} \sum_{j=0}^{l} a_i b_j \cdot A^{i+j} \\ &= \sum_{k=0}^{m+l} \Big(\sum_{i+j=k} a_i b_j \cdot A^{i+j} \Big) \\ &= \sum_{k=0}^{m+l} \Big(\sum_{i+j=k} a_i b_j \Big) \cdot A^k = \mathrm{ev}_A(p \cdot q). \end{aligned}$$

Dies ist die behauptete Eigenschaft. □

Bemerkung V.5.3. Die Mengen $K[A] := \mathrm{Bild}(\mathrm{ev}_A)$ bzw. $K[f] := \mathrm{Bild}(\mathrm{ev}_f)$ sind kommutative Unterringe von $(\mathrm{Mat}(n; K), +, \cdot)$ bzw. $(\mathrm{End}_K(f), +, \circ)$. Das bedeutet a) $\mathbb{E}_n \in K[A]$ bzw. $\mathrm{Id}_V \in K[f]$, b) $K[A]$ ist eine Untergruppe von $(\mathrm{Mat}(n; K), +)$ bzw. $K[f]$ ist eine Untergruppe von $(\mathrm{End}_K(V), +)$, c) $B, C \in K[A]$ impliziert $B \cdot C \in K[A]$ bzw. $f, g \in K[f]$ beinhaltet $f \circ g \in K[f]$ und d) $B \cdot C = C \cdot B$ für alle $B, C \in K[A]$ bzw. $f \circ g = g \circ f$ für alle $f, g \in K[f]$. Damit sind $(K[A], +, \cdot)$ bzw. $(K[f], +, \circ)$ insbesondere selber kommutative Ringe.

Satz V.5.4 (Cayley[6]–Hamilton[7])**.** i) *Für jeden Endomorphismus* $f \in \mathrm{End}_K(V)$ *gilt*

$$\chi_f(f) = 0.$$

ii) *Für jede Matrix* $A \in \mathrm{Mat}(n; K)$ *gilt*

$$\chi_A(A) = 0.$$

Sprechweise. Man sagt, A bzw. f erfüllt seine eigene charakteristische Gleichung.

Beweis von Satz V.5.4. Wir weisen ii) nach. Da $K[A]$ ein kommutativer Ring ist, haben wir über ihm die Theorie der Matrizen und Determinanten zur Verfügung (s. Abschnitt V.3). Wir definieren die Matrix

$$C := \begin{pmatrix} A - a_{11} \cdot \mathbb{E}_n & -a_{12} \cdot \mathbb{E}_n & \cdots\cdots\cdots\cdots & -a_{1n} \cdot \mathbb{E}_n \\ -a_{21} \cdot \mathbb{E}_n & A - a_{22} \cdot \mathbb{E}_n & -a_{23} \cdot \mathbb{E}_n \cdots & -a_{2n} \cdot \mathbb{E}_n \\ \vdots & \ddots & \ddots & \vdots \\ -a_{n1} \cdot \mathbb{E}_n & \cdots\cdots\cdots\cdots & -a_{nn-1} \cdot \mathbb{E}_n & A - a_{nn} \cdot \mathbb{E}_n \end{pmatrix} \in \mathrm{Mat}\big(n; K[A]\big),$$

d.h. für $C = (c_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$ gilt $c_{ii} = A - a_{ii} \cdot \mathbb{E}_n$, $i = 1, \dots, n$, und $c_{ij} = -a_{ij} \cdot \mathbb{E}_n$, $1 \le i \ne j \le n$. Anschaulich gesprochen haben wir in der Matrix $t \cdot \mathbb{E}_n - A \in \mathrm{Mat}(n; K[t])$ die Unbestimmte t durch A ersetzt. Mit der Leibnizformel und den Eigenschaften der Abbildung ev_A (Lemma V.5.1) erhält man sofort

$$\mathrm{Det}(C) = \chi_A(A) \quad \text{in } K[A].$$

Für $j = 1, \dots, n$ gilt

$$\sum_{i=1}^{n} c_{ij} \cdot e_i = A \cdot e_j - \sum_{i=1}^{n} a_{ij} \cdot e_i = j\text{-te Spalte von } A - \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} = 0. \tag{V.5}$$

Behauptung. *Für* $k = 1, \dots, n$ *gilt* $\mathrm{Det}(C) \cdot e_k = 0$.

Da $\mathrm{Det}(C) \cdot e_k$ die k-te Spalte der Matrix $\mathrm{Det}(C)$ ist, impliziert diese Behauptung die gesuchte Eigenschaft $\mathrm{Det}(C) = 0$.

Beweis der Behauptung. Es sei $\widetilde{C} = (\widetilde{c}_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$ die zu C komplementäre Matrix (s. Eigenschaft V.3.2, iii), d), und Definition IV.4.13, ii). Für die Matrix $C \cdot \widetilde{C} =: (c'_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$ gilt nach Eigenschaft V.3.2, iii), d):

$$c'_{ij} = \begin{cases} \mathrm{Det}(C), & i = j \\ 0, & i \ne j \end{cases}.$$

[6]Arthur Cayley (1821 - 1895), englischer Mathematiker.

[7]William Rowan Hamilton (1805 - 1865), irischer Mathematiker und Physiker.

Somit erhalten wir für $k = 1, ..., n$

$$\begin{aligned}
\operatorname{Det}(C) \cdot e_k &= \sum_{i=1}^{n} c'_{ik} \cdot e_i \\
&= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} c_{ij} \cdot \widetilde{c}_{jk} \right) \cdot e_i \\
&\overset{K[A] \text{ kommutativ}}{=} \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \widetilde{c}_{jk} \cdot c_{ij} \right) \cdot e_i \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \widetilde{c}_{jk} \cdot \left(c_{ij} \cdot e_i \right) \\
&= \sum_{j=1}^{n} \widetilde{c}_{jk} \cdot \underbrace{\left(\sum_{i=1}^{n} c_{ij} \cdot e_i \right)}_{\overset{(V.5)}{=} 0} = 0.
\end{aligned}$$

Somit ist der Satz bewiesen. □

Das Minimalpolynom. — Für einen Endomorphismus $f\colon V \longrightarrow V$ des endlichdimensionalen Vektorraums V bzw. eine Matrix $A \in \operatorname{Mat}(n; K)$ liefert der Satz von Cayley-Hamilton die Existenz eines Polynoms $p(t) \in K[t]$ mit $p(f) = 0$ bzw. $p(A) = 0$. Wir wollen nun unter allen Polynomen, die diese Eigenschaft haben, das „kleinste“ auszeichnen.

Satz V.5.5. *Für $f\colon V \longrightarrow V$ wie oben bzw. $A \in \operatorname{Mat}(n; K)$ gibt es genau ein normiertes Polynom $p \in K[t]$ von positivem Grad mit den folgenden beiden Eigenschaften:*

a) $p(f) = 0$ *bzw.* $p(A) = 0$.

b) p *teilt jedes Polynom* $q \in K[t]$ *mit* $q(f) = 0$ *bzw.* $q(A) = 0$.

Definition V.5.6. Das Polynom aus Satz V.5.5 heißt das *Minimalpolynom von* f bzw. A.
Schreibweise. $\mu_f(t)$ bzw. $\mu_A(t)$.

Beweis von Satz V.5.5. Wir zeigen die Aussage für einen Endomorphismus f. Wir wählen ein Polynom $\widetilde{p} \in K[t] \setminus \{0\}$ mit $\widetilde{p}(f) = 0$, das unter allen Polynomen q positiven Grades mit $q(f) = 0$ den minimalen Grad hat. Ein solches Polynom existiert nach dem Satz von Cayley-Hamilton. Es sei etwa $\widetilde{p} = b_m t^m + \cdots + b_1 t + b_0$ mit $b_m \neq 0$. Wir erhalten das normierte Polynom $p := (1/b_m) \cdot \widetilde{p}$. Nach Konstruktion gilt a) $p(f) = 0$. Gegeben sei ein Polynom $q \in K[t]$ mit $q(f) = 0$.

Division mit Rest liefert $q = s \cdot p + r$ mit $s, r \in K[t]$ und $\mathrm{Grad}(r) < \mathrm{Grad}(p)$. Dann gilt nach Lemma V.5.1

$$0 = q(f) = s(f)p(f) + r(f) = r(f).$$

Nach Wahl von p folgt aus $\mathrm{Grad}(r) < \mathrm{Grad}(p)$ und $r(f) = 0$ bereits $r = 0$, d.h. $p|q$. Damit ist auch Eigenschaft b) nachgewiesen. Für ein anderes normiertes Polynom p' von positivem Grad mit Eigenschaft a) und b) folgt zunächst $p|p'$ und $p'|p$. Damit müssen beide Polynome denselben Grad m haben, und es gibt eine Zahl $a \in K^*$ mit $p' = a \cdot p$. Da aber sowohl p' als auch p normiert sind, muss $a = 1$ und somit $p' = p$ gelten. Dies liefert die Eindeutigkeitsaussage im Satz. □

Folgerung V.5.7. *Für f bzw. A wie oben gilt*

$$\mu_f | \chi_f \quad \textit{bzw.} \quad \mu_A | \chi_A.$$

Inbesondere hat man $\mathrm{Grad}(\mu_{f/A}) \leq \mathrm{Grad}(\chi_{f/A}) = n$.

Bemerkung V.5.8. Um das Minimalpolynom zu definieren, benötigt man lediglich die Existenz eines Polynoms p vom Grad $m > 0$ mit $p(f) = 0$ bzw. $p(A) = 0$. Diese Aussage lässt sich recht elementar ohne den Satz von Cayley-Hamilton beweisen (s. Aufgabe V.5.11). Allerdings folgt so nicht die wichtige Aussage, dass das Minimalpolynom das charakteristische Polynom teilt.

Satz V.5.9. *Die Nullstellen des Minimalpolynoms sind genau die Eigenwerte von f bzw. A. D.h. es gilt*

$$\lambda \textit{ ist Nullstelle von } \mu_f(t) \textit{ bzw. } \mu_A(t) \iff \lambda \textit{ ist Nullstelle von } \chi_f(t) \textit{ bzw. } \chi_A(t).$$

Beweis. Wir zeigen die Aussage für $A \in \mathrm{Mat}(n; K)$. Aus $\mu_A(t) | \chi_A(t)$ folgt sofort $\mu_A(\lambda) = 0 \implies \chi_A(\lambda) = 0$ für alle $\lambda \in K$. Dies ist die Richtung „$\Longrightarrow$“. Für die Umkehrung „$\Longleftarrow$“ sei λ eine Nullstelle von $\chi_A(t)$, d.h. ein Eigenwert von A. Wir wählen einen Eigenvektor $v \in K^n \setminus \{0\}$ zum Eigenwert λ. Es gilt

$$\begin{aligned} 0 &= \mu_A(A)(v) = \left(A^m + a_{m-1} \cdot A^{m-1} + \cdots + a_1 \cdot A + a_0 \cdot \mathbb{E}_n\right)(v) \\ &= \lambda^m \cdot v + a_{m-1}\lambda^{m-1} \cdot v + \cdots + a_1 \lambda \cdot v + a_0 \cdot v = \mu_A(\lambda) \cdot v. \end{aligned}$$

Da $v \neq 0$, muss $\mu_A(\lambda) = 0$ gelten, d.h. λ ist eine Nullstelle von $\mu_A(t)$. □

Beispiel V.5.10. Wir betrachten die Matrix

$$A := \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in \mathrm{Mat}(4; \mathbb{Q}).$$

Das charakteristische Polynom von A ist gegeben durch

$$\chi_A(t) = (t-2)^2 \cdot (t-1)^2.$$

Wir wissen bereits, dass das Minimalpolynom von der Form $\mu_A(t) = (t-2)^i \cdot (t-1)^j$ mit $1 \leq i, j \leq 2$ ist. Man überprüft $(2 \cdot \mathbb{E}_4 - A) \cdot (\mathbb{E}_4 - A)^2 \neq 0$ und $(2 \cdot \mathbb{E}_4 - A)^2 \cdot (\mathbb{E}_4 - A) = 0$. Damit folgt

$$\mu_A(t) = (t-2)^2 \cdot (t-1).$$

Aufgabe V.5.11.
Es seien V ein n-dimensionaler K-Vektorraum und $f\colon V \longrightarrow V$ ein Endomorphismus bzw. $A \in \mathrm{Mat}(n; K)$, $n \geq 1$. Zeigen Sie, dass ein Polynom $p \in K[t]$ mit $n^2 \geq \mathrm{Grad}(f) > 0$ existiert, so dass $p(f) = 0$ bzw. $p(A) = 0$. Dabei sollen Sie nur die Konzepte der linearen (Un-)Abhängigkeit und der Dimension eines Vektorraums verwenden (und natürlich nicht den Satz von Cayley-Hamilton).

V.6 Trigonalisierbare Endomorphismen

Definitionen V.6.1. i) Es seien $n \geq 1$ eine natürliche Zahl und R ein Ring. Eine Matrix $A = (a_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}} \in \mathrm{Mat}(n; R)$ wird als *obere Dreiecksmatrix* bezeichnet, wenn

$$\forall n \geq i > j \geq 1 : \quad a_{ij} = 0$$

gilt, d.h. wenn alle Einträge unterhalb der Diagonalen verschwinden.

ii) Es sei V ein endlichdimensionaler K-Vektorraum. Ein Endomorphismus $f\colon V \longrightarrow V$ wird *trigonalisierbar* genannt, wenn es eine durchnummerierte Basis B von V gibt, für die die Darstellungsmatrix $M_B(f)$ eine obere Dreiecksmatrix ist.

iii) Eine Matrix $A \in \mathrm{Mat}(n; K)$ heißt *trigonalisierbar*, wenn sie einer oberen Dreiecksmatrix ähnlich ist.

Satz V.6.2. *Für f bzw. A wie oben sind die folgenden Aussagen äquivalent:*

a) *f bzw. A ist trigonalisierbar.*

b) *Das charakteristische Polynom von f bzw. A zerfällt vollständig in Linearfaktoren.*

Beweis. Wir zeigen den Satz für Matrizen und beginnen mit **„a)⟹b)“**. Es sei $S \in \mathrm{GL}_n(K)$ so, dass

$$S \cdot A \cdot S^{-1} = \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}.$$

Dann gilt

$$\chi_A(t) \overset{\text{Lemma V.3.6}}{=} \chi_{S\cdot A\cdot S^{-1}}(t) \overset{\text{Aufgabe IV.4.28, a)}}{=} (t-a_{11})\cdot\dots\cdot(t-a_{nn}).$$

Die Richtung **„b)⟹a)“** beweisen wir durch Induktion über n. Für $n = 1$ ist offenbar nichts zu beweisen. Für den Schritt „$n \rightsquigarrow n+1$“ wählen wir zunächst eine Nullstelle λ von $\chi_A(t)$, die existiert, weil $\chi_A(t)$ nach Voraussetzung in Linearfaktoren zerfällt. Es sei $v_1 \in K^{n+1}\setminus\{0\}$ ein Eigenvektor von A zum Eigenwert λ, den man durch Vektoren $v_2,\dots,v_{n+1}$ zu einer durchnummerierten Basis von K^{n+1} ergänze. Man bilde die Matrix $T := (v_1|v_2|\cdots|v_{n+1}) \in GL_{n+1}(K)$ mit den Spalten $v_1,\dots,v_{n+1}$. Dann gilt

$$B := T^{-1}\cdot A\cdot T = \left(\begin{array}{c|ccc} \lambda & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{array}\right)$$

und

$$\chi_A(t) = (t-\lambda)\cdot\chi_{A'}(t).$$

Das Polynom $\chi_{A'}(t)$ zerfällt offensichtlich ebenfalls in Linearfaktoren, so dass es nach Induktionsvoraussetzung eine Matrix $S' \in GL_n(K)$ mit

$$S'\cdot A'\cdot S'^{-1} = \begin{pmatrix} a'_{11} & & * \\ & \ddots & \\ 0 & & a'_{nn} \end{pmatrix}$$

gibt. Es ergibt sich

$$\underbrace{\left(\begin{array}{c|ccc} 1 & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & S' & \\ 0 & & & \end{array}\right)}_{=:S''}\cdot B\cdot \underbrace{\left(\begin{array}{c|ccc} 1 & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & S'^{-1} & \\ 0 & & & \end{array}\right)}_{=S''^{-1}} = \left(\begin{array}{c|ccc} \lambda & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & S'\cdot A'\cdot S'^{-1} & \\ 0 & & & \end{array}\right)$$

$$= \left(\begin{array}{c|ccc} \lambda & * & \cdots & * \\ \hline 0 & a'_{11} & & * \\ \vdots & & \ddots & \\ 0 & 0 & & a'_{nn} \end{array}\right).$$

Für $S := S''\cdot T^{-1} \in GL_{n+1}(K)$ ist $S\cdot A\cdot S^{-1}$ daher eine obere Dreiecksmatrix. □

Bemerkung V.6.3. Für einen algebraisch abgeschlossenen Körper K, z.B. $K = \mathbb{C}$, ist Bedingung b) in Satz V.6.2 immer erfüllt, so dass jeder Endomorphismus bzw. jede Matrix über K trigonalisierbar ist. Dies bringt uns der Lösung des Normalformproblems V.1.6 wieder einen Schritt näher. Das Ergebnis aus Satz V.6.2 ist noch nicht völlig zufrieden stellend, da z.B. eine Matrix vielen oberen Dreiecksmatrizen ähnlich sein kann und wir auch nicht ohne weiteres entscheiden können, wann zwei obere Dreiecksmatrizen ähnlich sind.

Aufgabe V.6.4.
Gegeben sei

$$A := \begin{pmatrix} 21 & -24 & 5 \\ 14 & -13 & 4 \\ -37 & 42 & -9 \end{pmatrix} \in \mathrm{Mat}(3;\mathbb{R}).$$

a) Berechnen Sie das charakteristische Polynom von A.
b) Ermitteln Sie alle Nullstellen von χ_A in $\mathbb{R}$ und in $\mathbb{C}$.
c) Ist die Matrix A über $\mathbb{R}$ diagonalisierbar oder trigonalisierbar? Ist sie über $\mathbb{C}$ diagonalisierbar oder trigonalisierbar?

V.7 Nilpotente Endomorphismen

In diesem Abschnitt werden wir mit der Einführung einer verfeinerten oberen Dreiecksgestalt beginnen, die insbesondere die in Bemerkung V.6.3 angeschnittenen Probleme beseitigt.

Definitionen V.7.1. i) Es sei V ein endlichdimensionaler K-Vektorraum. Ein Endomorphismus $f\colon V \longrightarrow V$ ist *nilpotent*, wenn es eine natürliche Zahl $p \geq 1$ mit $f^p = 0$ gibt.

ii) Eine *nilpotente Matrix* ist eine Matrix $A \in \mathrm{Mat}(n;K)$, zu der es eine natürliche Zahl $p \geq 1$ mit $A^p = 0$ gibt.[8]

Bemerkung V.7.2. Es sei z.B. A eine nilpotente Matrix. Wie der folgende Satz zeigt, ist A einer oberen Dreiecksmatrix mit Nullen auf der Diagonalen ähnlich. Insbesondere gilt $\chi_A(t) = t^n$. Ist umgekehrt $A \in \mathrm{Mat}(n;K)$ mit $\chi_A(t) = t^n$ gegeben, dann folgt nach dem Satz von Cayley-Hamilton $A^n = 0$, so dass A nilpotent ist. Nilpotente Matrizen bzw. Endomorphismen sind also genau diejenigen trigonalisierbaren Endomorphismen bzw. Matrizen, deren einziger Eigenwert $\lambda = 0$ ist. Sie sollten daher relativ einfach zu verstehen sein, so dass es sinnvoll erscheint, die Untersuchungen mit ihnen zu beginnen.

[8] Die auftretenden Potenzen wurden in Definition V.5.2 eingeführt.

Satz V.7.3. *Es sei* $f\colon V \longrightarrow V$ *ein nilpotenter Endomorphismus des* n*-dimensionalen* K*-Vektorraums* V *bzw.* $A \in \mathrm{Mat}(n;K)$ *eine nilpotente Matrix. Dann gibt es eine durchnummerierte Basis* $B = \{b_1, ..., b_n\}$ *von* V *bzw.* K^n*, so dass*

$$\begin{array}{rll} & f(b_i) = b_{i-1} & \textit{bzw.} \quad A \cdot b_i = b_{i-1} \\ \textit{oder} & f(b_i) = 0 & \textit{bzw.} \quad A \cdot b_i = 0 \end{array}$$

für $i = 1, ..., n$ *gilt.*

Beweis. Wir beweisen den Satz für einen Endomorphismus $f\colon V \longrightarrow V$. Zunächst sei $p \geq 1$ minimal mit der Eigenschaft $f^p = 0$. Dann setzen wir $W_i := \mathrm{Ker}(f^i)$, $i = 1, ..., p$. Mit $W_0 := \{0\}$ erhalten wir

$$W_0 \subseteq W_1 \subseteq W_2 \subseteq \cdots \subseteq W_{p-1} \subseteq W_p = V.$$

Offenbar gilt

$$f(W_i) \subseteq W_{i-1}, \quad i = 1, ..., p. \tag{V.6}$$

Diese einfache Beobachtung gibt den entscheidenden Hinweis darauf, dass eine Basis wie im Satz behauptet existieren könnte. Wir müssen diese Basis jetzt sorgfältig konstruieren. Dazu verwenden wir das folgende Resultat.

Behauptung. *Es gibt lineare Teilräume* $U_i \subseteq W_i \subseteq V$, $i = 1, ..., p$, *so dass*

a) $W_i = W_{i-1} \oplus U_i$, $i = 1, ..., p$.

b) $f(U_i) \subseteq U_{i-1}$, $i = 2, ..., p$.

c) $V = U_1 \oplus \cdots \oplus U_p$.

Wir widmen uns jetzt dem Beweis dieser Behauptung. Die Unterräume U_1,..., U_p werden rekursiv eingeführt, beginnend mit U_p. Für den Raum U_p wählen wir ein beliebiges lineares Komplement zu W_{p-1} in $W_p = V$. Wenn die Unterräume U_j mit $j > i$ bereits konstruiert worden sind, dann erhalten wir U_i auf folgende Weise. Es gilt

$$f(U_{i+1}) \cap W_{i-1} = \{0\}. \tag{V.7}$$

In der Tat folgt für ein Element $v \in U_{i+1}$ mit $w := f(v) \in W_{i-1}$

$$f^i(v) = f^{i-1}\Big(f(v)\Big) = 0,$$

d.h. $v \in W_i$. Da aber v auch in U_{i+1} liegt und $U_{i+1} \cap W_i = \{0\}$ gilt, erhalten wir $v = 0$ und damit $w = 0$ wie gewünscht. Da weiter $U_{i+1} \subseteq W_{i+1}$, beinhaltet (V.6), dass $f(U_{i+1}) \subseteq W_i$. Wegen (V.7) können wir daher ein lineares Komplement $U_i \subseteq W_i$ zu W_{i-1} finden, welches $f(U_{i+1})$ enthält. Insgesamt erhalten wir Unterräume U_1,...,U_p, für die a) und b) gilt. Aus a) folgt sofort $W_1 = U_1$. Damit gilt $W_2 =$

$W_1 \oplus U_2 = U_1 \oplus U_2$. So fortfahrend sehen wir, dass $W_j = U_1 \oplus \cdots \oplus U_j$ für $j = 1, ..., p$ und damit insbesondere c) gilt. Schließlich ist die Abbildung

$$\begin{array}{rcl} f_{|U_i}\colon U_i & \longrightarrow & U_{i-1} \\ v & \longmapsto & f(v) \end{array}$$

für $i = 2, ..., p$ injektiv. Der Kern dieser Abbildung ist nämlich

$$U_i \cap \mathrm{Ker}(f) = U_i \cap W_1 \overset{i\geq 2}{\subseteq} U_i \cap W_{i-1} = \{0\}.$$

Diese letzte Beobachtung und die Behauptung implizieren, dass die Vektoren in der nachstehenden Tabelle zusammen eine Basis von V bilden.

$$\begin{array}{ccccc} u_1^p, ..., u_{\pi_p}^p & & & & \\ f(u_1^p), ..., f(u_{\pi_p}^p), & & u_1^{p-1}, ..., u_{\pi_{p-1}}^{p-1} & & \\ \vdots & & \vdots & & \vdots \\ f^{p-1}(u_1^p), ..., f^{p-1}(u_{\pi_p}^p), & f^{p-2}(u_1^{p-1}), ..., f^{p-2}(u_{\pi_{p-1}}^{p-1}) & , ..., & f(u_1^2), ..., f(u_{\pi_2}^2), & u_1^1, ..., u_{\pi_1}^1. \end{array} \tag{V.8}$$

Dabei steht in der ersten Zeile eine Basis von U_p, in der zweiten eine Basis von U_{p-1} und in der letzten eine Basis von U_1. Wir erhalten eine durchnummerierte Basis $B = \{b_1, ..., b_n\}$, indem wir zunächst die Vektoren der ersten Spalte von unten nach oben für b_1,...,b_p erklären, dann die der zweiten von unten nach oben für b_{p+1},...,b_{2p} usw.:

$$\begin{array}{l} b_1 := f^{p-1}(u_1^p), ..., b_p := u_1^p, \\ b_{p+1} := f^{p-1}(u_2^p), ..., b_{2p} := u_2^p, \\ \quad \vdots \\ b_{\pi_p \cdot p+1} := f^{p-2}(u_1^{p-1}), ..., b_{(\pi_p+1)p-1} := u_1^{p-1}, \\ \quad \vdots \\ b_{n-\pi_1+1} := u_1^1, ..., b_n := u_{\pi_1}^1. \end{array}$$

Für die so konstruierte Basis ist die Behauptung des Satzes klar. □

Bemerkungen V.7.4. i) Man beachte, dass der Beweis des Satzes einen expliziten Algorithmus zum Auffinden der Basis B zur Verfügung stellt.

ii) Die Basis B ist natürlich nicht eindeutig bestimmt, denn für die Komplemente U_1,..., U_p und ihre Basen sind viele Wahlen möglich und erlaubt.

iii) Die Konstruktion der Basis B beschreiben wir noch einmal in Worten: Da $f^p = 0$ aber $f^{p-1} \neq 0$, wählen wir zunächst eine maximale Anzahl linear unabhängiger Vektoren v mit $f^{p-1}(v) \neq 0$, nämlich u_1^p,...,$u_{\pi_p}^p$. Diese Vektoren zusammen mit ihren Bildern unter den Potenzen f,...,f^{p-1} sind auf Grund der

Behauptung linear unabhängig. Danach suchen wir eine maximale Menge linear unabhängiger Vektoren v mit $f^{p-1}(v) = 0$ aber $f^{p-2}(v) \neq 0$, die sich nicht als Linearkombinationen von $f(u_1^p)$, ...,$f(u_{\pi_p}^p)$ schreiben lassen. Dies sind die Vektoren $u_1^{p-1}, ..., u_{\pi_{p-1}}^{p-1}$. So fahren wir fort, bis die Basis B konstruiert ist.

Definitionen V.7.5. i) Es seien $l \in \mathbb{N}$, $l \geq 1$, und $\lambda \in K$. Der *Jordan-Block*[9] *der Größe* l *zum Eigenwert* λ ist die Matrix

$$J(l,\lambda) := \begin{pmatrix} \lambda & 1 & & \\ & \lambda & 1 & \mathbf{0} \\ & & \ddots & \ddots \\ & \mathbf{0} & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} \in \mathrm{Mat}(l;K),$$

d.h. $J(l,\lambda) = (a_{ij})_{\substack{i=1,...,l \\ j=1,...,l}}$ mit $a_{ii} = \lambda$, $i = 1, ..., l$, $a_{ii+1} = 1$, $i = 1, ..., l-1$, und $a_{ij} = 0$ in den übrigen Fällen.

ii) Eine *Partition der natürlichen Zahl* n ist ein Tupel

$$\underline{\pi} = (\pi_n, ..., \pi_1), \quad \text{mit } \pi_i \in \mathbb{N},\ i = 1, ..., n,$$

so dass

$$n \cdot \pi_n + (n-1) \cdot \pi_{n-1} + \cdots + 2 \cdot \pi_2 + \pi_1 = n.$$

iii) Für eine Partition $\underline{\pi} = (\pi_n, ..., \pi_1)$ von $n \in \mathbb{N}$ und eine Zahl $\lambda \in K$ ist die *Jordan-Blockmatrix zu der Partition* $\underline{\pi}$ *und dem Eigenwert* λ gegeben durch

$$J(\underline{\pi},\lambda) := \begin{pmatrix} \overbrace{\begin{matrix} J(p,\lambda) & & \\ & \ddots & \\ & & J(p,\lambda) \end{matrix}}^{\pi_p \text{ Blöcke}} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \underbrace{\begin{matrix} J(1,\lambda) & & \\ & \ddots & \\ & & J(1,\lambda) \end{matrix}}_{\pi_1 \text{ Blöcke}} \end{pmatrix} \in \mathrm{Mat}(n;K).$$

Dabei sei $p := \max\{\, j = 1, ..., n \mid \pi_j \neq 0 \,\}$.

[9] Marie Ennemond Camille Jordan, genannt Camille Jordan (1838 - 1922 in Paris), französischer Mathematiker.

Beispiel V.7.6. Für $n = 3$ ergeben sich die folgenden Möglichkeiten:

Partition	Jordan-Blockmatrix
(1,0,0)	$\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$
(0,1,1)	$\left(\begin{array}{cc\|c} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ \hline 0 & 0 & \lambda \end{array}\right)$
(0,0,3)	$\left(\begin{array}{c\|c\|c} \lambda & 0 & 0 \\ \hline 0 & \lambda & 0 \\ \hline 0 & 0 & \lambda \end{array}\right)$

.

Satz V.7.7. *Es seien* $f\colon V \longrightarrow V$ *ein Endomorphismus des* n*-dimensionalen* K*-Vektorraums* V *bzw.* $A \in \mathrm{Mat}(n;K)$ *und* $\lambda \in K$*. Es gelte* $\chi_f(t) = (t-\lambda)^n$ *bzw.* $\chi_A(t) = (t-\lambda)^n$*.*[10]

i) *Dann existieren eine Partition* $\underline{\pi}$ *von* n *und eine durchnummerierte Basis* $B = \{b_1, ..., b_n\}$ *von* V *bzw.* K^n*, so dass*

$$M_B(f) = J(\underline{\pi}, f)$$

bzw.

$$T^{-1} \cdot A \cdot T = J(\underline{\pi}, f), \quad T := (b_1|\cdots|b_n) \in \mathrm{GL}_n(K).$$

ii) *Für die Partition* $\underline{\pi} = (\pi_n, ..., \pi_1)$ *sei*

$$\begin{aligned} \nu &:= \pi_n + \cdots + \pi_1 = \textit{Anzahl der Jordan-Blöcke} \\ p &:= \max\{\, i = 1, ..., n \,|\, \pi_i \neq 0 \,\}. \end{aligned}$$

Dann gilt

a) $\nu = \mathrm{Dim}_K\big(\mathrm{Eig}(f,\lambda)\big)$ *bzw.* $\nu = \mathrm{Dim}_K\big(\mathrm{Eig}(A,\lambda)\big)$.

b) $\mu_f(t) = (t-\lambda)^p$ *bzw.* $\mu_A(t) = (t-\lambda)^p$.

iii) *Für zwei Zahlen* $\lambda, \lambda' \in K$ *und zwei Partitionen* $\underline{\pi}, \underline{\pi}'$ *von* n *gilt:*

$$J(\underline{\pi}, \lambda) \approx J(\underline{\pi}', \lambda') \Longleftrightarrow J(\underline{\pi}, \lambda) = J(\underline{\pi}', \lambda') \Longleftrightarrow \lambda = \lambda' \wedge \underline{\pi} = \underline{\pi}'.$$

Beweis. **Zu i).** Wir zeigen die Aussage für den Endomorphismus f. Für $\lambda = 0$ ist das die Aussage von Satz V.7.3. Einen Endomorphismus wie im Satz schreiben wir in der Form

$$f = (\underbrace{f - \lambda \cdot \mathrm{Id}_V}_{=:g}) + \lambda \cdot \mathrm{Id}_V.$$

[10] Damit ist f bzw. A trigonalisierbar.

Aus den Eigenschaften von Darstellungsmatrizen folgt

$$M_B(f) = M_B(g) + M_B(\lambda \cdot \mathrm{Id}_V) = M_B(g) + \lambda \cdot \mathbb{E}_n \tag{V.9}$$

für jede durchnummerierte Basis B von V. Auf der anderen Seite gilt nach dem Satz von Cayley-Hamilton

$$g^n = \chi_f(f) = 0,$$

i.e. g ist nilpotent. Die Behauptung folgt somit durch Anwendung von Satz V.7.3 auf g und Gleichung (V.9).

Zu ii). Für einen nilpotenten Endomorphismus zeigt der Beweis von Satz V.7.3, dass die Anzahl der Jordan-Blöcke durch $\mathrm{Dim}_K(U_1) = \mathrm{Dim}_K(\mathrm{Ker}(g))$ gegeben ist.[11] In unserem Fall ist $\mathrm{Dim}_K(\mathrm{Ker}(g)) = \mathrm{Dim}_K(\mathrm{Eig}(f,\lambda))$ wie behauptet. Dies zeigt a). Da das Minimalpolynom $\mu_f(t)$ das charakteristische Polynom $\chi_f(t) = (t-\lambda)^n$ teilt, ist es von der Form $(t-\lambda)^j$ mit

$$\begin{aligned} j &= \min\{ i \in \mathbb{N} \mid g^i = (f - \lambda \cdot \mathrm{Id}_V)^i = 0 \} \\ &= \min\{ i \in \mathbb{N} \mid M_B(g)^i = 0 \}. \end{aligned}$$

Wir wählen die durchnummerierte Basis B so, dass $M_B(f) = J(\underline{\pi}, \lambda)$ und daher $M_B(g) = J(\underline{\pi}, 0)$. Es gilt

$$M_B(g)^i = \begin{pmatrix} J(p,0)^i & & & 0 \\ & J(p,0)^i & & \\ & & \ddots & \\ 0 & & & J(1,0)^i \end{pmatrix}$$

[11] Die Anzahl der Jordan-Blöcke ist durch die Anzahl der Vektoren in der untersten Zeile von Tabelle (V.8) gegeben. Diese ist zum einen offenbar $\mathrm{Dim}_K(U_1)$, zum anderen aber $\pi_n + \cdots + \pi_1$, wie man (V.8) entnimmt.

Nun überzeugt man sich leicht, dass

$$J(j,0)^i = \begin{pmatrix} 0 & \cdots & 0 & \overset{a_{1i+1}}{\downarrow} 1 & & O \\ & & & & \ddots & \\ & & & & & 1 \leftarrow a_{(j-i)j} \\ & O & & & & 0 \\ & & & & & \vdots \\ & & & & & 0 \end{pmatrix},$$

so dass genau dann $J(j,0)^i = 0$ gilt, wenn $i \geq j$. Damit erhält man die Aussage zum Minimalpolynom.

Zu iii). Die Äquivalenz „$J(\underline{\pi},\lambda) = J(\underline{\pi}',\lambda') \Longleftrightarrow \underline{\pi} = \underline{\pi}'$ und $\lambda = \lambda'$“ ist offensichtlich, ebenso wie die Implikation „$J(\underline{\pi},\lambda) = J(\underline{\pi}',\lambda') \Longrightarrow J(\underline{\pi},\lambda) \approx J(\underline{\pi}',\lambda')$“. Zeigen wir nun „$J(\underline{\pi},\lambda) \approx J(\underline{\pi}',\lambda') \Longrightarrow J(\underline{\pi},\lambda) = J(\underline{\pi}',\lambda')$“. Da λ der einzige Eigenwert von $A := J(\underline{\pi},\lambda)$ und λ' der einzige Eigenwert von $B := J(\underline{\pi}',\lambda')$ ist und ähnliche Matrizen dieselben Eigenwerte haben, folgt bereits $\lambda = \lambda'$. Um die Gleichheit der Partitionen einzusehen, erinnern wir an die Beweisstrategie von Satz V.7.3. Dort hatten wir die Unterräume

$$W_i := \left\{ v \in K^n \,|\, (A - \lambda \cdot \mathbb{E}_n)^i \cdot v = 0 \right\},\ i \in \mathbb{N},$$

von K^n eingeführt. Mit $p := \min\{\, i \in \mathbb{N} \,|\, (A - \lambda \cdot \mathbb{E}_n)^i = 0 \,\}$ erhalten wir die „aufsteigende Kette“[12]

$$W_0 = \{0\} \subseteq W_1 \subseteq W_2 \subseteq \cdots \subseteq W_{p-1} \subsetneq W_p = W.$$

Des Weiteren haben wir $W_i = U_i \oplus W_{i-1}$, $i = 1, ..., p$, geschrieben. Dabei sei

$$(A - \lambda \cdot \mathbb{E}_n) \cdot U_i \subseteq U_{i-1},\ i = 2, ..., p.$$

Die Partition $\underline{\pi} = (\pi_n, ..., \pi_1)$ ist nun gegeben durch[13]

$$\pi_i = \begin{cases} 0, & i > p \\ \mathrm{Dim}_K(U_p), & i = p \\ \mathrm{Dim}_K(U_i) - \mathrm{Dim}_K(U_{i+1}), & i = 1, ..., p-1 \end{cases}.$$

Nach Konstruktion gilt $\mathrm{Dim}_K(U_i) = \mathrm{Dim}_K(W_i) - \mathrm{Dim}_K(W_{i-1})$, so dass

$$\pi_i = -\mathrm{Dim}_K(W_{i+1}) + 2 \cdot \mathrm{Dim}_K(W_i) - \mathrm{Dim}_K(W_{i-1}) \quad \text{für } i = 1, ..., n$$

[12] Die Inklusion $W_{p-1} \subseteq W_p$ ist auf Grund der Wahl von p strikt.

[13] Wegen $W_{p-1} \subsetneq W_p$ gilt $U_p \neq \{0\}$, i.e. $\mathrm{Dim}_K(U_p) \geq 1$.

folgt. Es sei $S \in GL_n(K)$ so, dass $B = SAS^{-1}$. Man berechnet auch $S \cdot (A - \lambda \cdot \mathbb{E}_n) \cdot S^{-1} = B - \lambda \cdot \mathbb{E}_n$, $B^i = S \cdot A^i \cdot S^{-1}$ und $S \cdot (A - \lambda \cdot \mathbb{E}_n)^i \cdot S^{-1} = (B - \lambda \cdot \mathbb{E}_n)^i$, $i \in \mathbb{N}$. Setzt man

$$W'_i := \left\{ v \in K^n \mid (B - \lambda \cdot \mathbb{E}_n)^i \cdot v = 0 \right\},\ i \in \mathbb{N},$$

dann überprüft man wie in Bemerkung V.3.17, dass $W'_i = S \cdot W_i$ für alle $i \in \mathbb{N}$ erfüllt ist. Insbesondere hat man $\mathrm{Dim}_K(W'_i) = \mathrm{Dim}_K(W_i)$ für alle $i \in \mathbb{N}$, so dass auch

$$\begin{aligned} \pi'_i &= -\mathrm{Dim}_K(W'_{i+1}) + 2 \cdot \mathrm{Dim}_K(W'_i) - \mathrm{Dim}_K(W'_{i-1}) \\ &= -\mathrm{Dim}_K(W_{i+1}) + 2 \cdot \mathrm{Dim}_K(W_i) - \mathrm{Dim}_K(W_{i-1}) = \pi_i \end{aligned}$$

für $i = 1, ..., n$ gilt. □

Beispiel V.7.8. Wir betrachten die Matrix

$$A := \begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix} \in \mathrm{Mat}(3; \mathbb{Q}).$$

Mit der Regel von Sarrus bestimmen wir das charakteristische Polynom:

$$\begin{array}{ccccc} t-3 & -4 & -3 & t-3 & -4 \\ 1 & t & 1 & 1 & t \\ -1 & -2 & t-3 & -1 & -2 \end{array},$$

also

$$\begin{aligned} \chi_A(t) &= t(t-3)^2 + 4 + 6 - 3t + 2t - 6 + 4t - 12 \\ &= t(t^2 - 6t + 9) + 3t - 8 = t^3 - 6t^2 + 12t - 8 \\ &= t^3 + 3(-2t^2) + 3((-2)^2 t) + (-2)^3 = (t-2)^3. \end{aligned}$$

Wir müssen

$$C := A - 2 \cdot \mathbb{E}_3 = \begin{pmatrix} 1 & 4 & 3 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix}$$

untersuchen. Es gilt

$$C^2 = \begin{pmatrix} 1 & 4 & 3 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 4 & 3 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 2 & 2 \\ 0 & -2 & -2 \\ 0 & 2 & 2 \end{pmatrix} \neq 0.$$

(Damit ist bereits klar, dass die Jordansche Normalform von A durch

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

gegeben ist.) Wir wissen, dass U_3, U_2 und U_1 jeweils eindimensional sind. Wir müssen nur einen erzeugenden Vektor v für einen geeigneten Unterraum U_3 ermitteln. Dann ist $B = \{ C^2 \cdot v, C \cdot v, v \}$ die gesuchte Basis. Offensichtlich gilt

$$\mathbb{L}(C^2, 0) = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\rangle.$$

Damit können wir

$$v := \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{und} \quad U_3 = \langle v \rangle$$

wählen. (Es sind viele andere Wahlen möglich.) Es ergibt sich

$$C^2 \cdot v = \begin{pmatrix} 2 \\ -2 \\ 2 \end{pmatrix}, \quad C \cdot v = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}, \quad v = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

und somit

$$T = \begin{pmatrix} 2 & 3 & 0 \\ -2 & -1 & 0 \\ 2 & 1 & 1 \end{pmatrix}.$$

Die Inverse von T wird mit dem Gauß-Algorithmus bestimmt:

$$\left(\begin{array}{ccc|ccc} 2 & 3 & 0 & 1 & 0 & 0 \\ -2 & -1 & 0 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \overset{\text{II+I; III−I}}{\rightsquigarrow} \left(\begin{array}{ccc|ccc} 2 & 3 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 1 & 0 \\ 0 & -2 & 1 & -1 & 0 & 1 \end{array}\right)$$

$$\overset{\text{III+II; } \frac{1}{2}\cdot\text{II}}{\rightsquigarrow} \left(\begin{array}{ccc|ccc} 2 & 3 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{array}\right) \overset{\text{I}-3\cdot\text{II}}{\rightsquigarrow} \left(\begin{array}{ccc|ccc} 2 & 0 & 0 & -\frac{1}{2} & -\frac{3}{2} & 0 \\ 0 & 1 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{array}\right)$$

$$\overset{\frac{1}{2}\cdot\text{I}}{\rightsquigarrow} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{4} & -\frac{3}{4} & 0 \\ 0 & 1 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{array}\right).$$

Damit gilt

$$T^{-1} = \begin{pmatrix} -\frac{1}{4} & -\frac{3}{4} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Probe.

$$\begin{pmatrix} -\frac{1}{4} & -\frac{3}{4} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2 & 3 & 0 \\ -2 & -1 & 0 \\ 2 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \cdot \surd$$

V.8 Die Jordansche Normalform

Das folgende Resultat liefert die vollständige Lösung von Problem V.1.6 über einem algebraisch abgeschlossenen Körper.

Satz V.8.1 (Die Jordansche Normalform)**.** *Es sei* $f\colon V \longrightarrow V$ *ein Endomorphismus des* n*-dimensionalen* K*-Vektorraums* V *bzw.* $A \in \mathrm{Mat}(n;K)$*. Falls* f *bzw.* A *trigonalisierbar ist (z.B., wenn* K *algebraisch abgeschlossen ist), d.h. falls*

$$\begin{aligned} & \chi_f(t) = (t-\lambda_1)^{l_1} \cdot \cdots \cdot (t-\lambda_m)^{l_m} \\ \textit{bzw.} \quad & \chi_A(t) = (t-\lambda_1)^{l_1} \cdot \cdots \cdot (t-\lambda_m)^{l_m} \end{aligned}$$

mit $\lambda_i \neq \lambda_j \in K$ *für* $1 \leq i \neq j \leq m$*, dann gilt:*

i) *Es gibt Partitionen* $\underline{\pi}^j = (\pi^j_{l_j}, ..., \pi^j_1)$ *von* l_j*,* $j = 1, ..., m$*, und eine durchnummerierte Basis* $B = \{b_1, ..., b_n\}$ *von* V *bzw.* K^n*, so dass*

$$M_B(f) = J \quad \textit{bzw.} \quad T^{-1} \cdot A \cdot T = J$$

mit $T := (b_1|\cdots|b_n) \in GL_n(K)$ *und*

$$J := \begin{pmatrix} J(\underline{\pi}^1, \lambda_1) & & & 0 \\ & J(\underline{\pi}^2, \lambda_2) & & \\ & & \ddots & \\ 0 & & & J(\underline{\pi}^m, \lambda_m) \end{pmatrix}.$$

ii) *Für* $j = 1, ..., m$ *setze man*

$$\begin{aligned} \nu_j &:= \pi^j_{l_j} + \cdots + \pi^j_1 \quad \textit{(=Anzahl der Jordan-Blöcke zum Eigenwert } \lambda_j\textit{)} \\ p_j &:= \max\{\, i = 1, ..., l_j \,|\, \pi^j_i \neq 0 \,\}. \end{aligned}$$

Dann gilt

a) $\nu_j = \mathrm{Dim}_K(\mathrm{Eig}(f, \lambda_j))$ *bzw.* $\nu_j = \mathrm{Dim}_K(\mathrm{Eig}(A, \lambda_j))$*. (Die Anzahl der Jordan-Blöcke zum Eigenwert* λ *ist gleich der Dimension des Eigenraums zum Eigenwert* λ*.)*

b) $\mu_f(t) = (t-\lambda_1)^{p_1} \cdot \cdots \cdot (t-\lambda_s)^{p_s}$ *bzw.* $\mu_A(t) = (t-\lambda_1)^{p_1} \cdot \cdots \cdot (t-\lambda_s)^{p_s}$*. (Die Größe des größten Jordan-Blocks zum Eigenwert* λ *ist gleich der Vielfachheit der Nullstelle* λ *des Minimalpolynoms.)*

iii) *Zwei Matrizen sind genau dann ähnlich, wenn sie dieselben Eigenwerte* λ_1*,...,*λ_m *haben und ihre Jordanschen Normalformen übereinstimmen.*

Definition V.8.2. Die Matrix J in Teil i) des Satzes heißt die *Jordansche Normalform* von f bzw. A.

Bemerkung V.8.3. Die einzige willkürliche Wahl, die in die Jordansche Normalform eingeht, ist die Nummerierung der Nullstellen (vgl. Bemerkung V.4.3, iii). Anders als $\mathbb{R}$ und $\mathbb{Q}$ besitzt ein algebraisch abgeschlossener Körper keine „natürliche“ Anordnung. Daher muss man entweder eine künstliche Anordnung vorgeben oder sich mit der kleinen Mehrdeutigkeit abfinden.

Die Zerlegung in verallgemeinerte Eigenräume. — Wir bereiten den Beweis von Satz V.8.1 vor. Die natürliche Idee, den Satz über die Jordansche Normalform für einen Endomorphismus f zu beweisen, ist, den Raum V in eine direkte Summe

$$V = V_1 \oplus \cdots \oplus V_m$$

zu zerlegen, so dass es Endomorphismen $f_i\colon V_i \longrightarrow V_i$, $i = 1, ..., m$, mit folgenden Eigenschaften gibt:

a) Für jedes $v \in V$ und die eindeutig bestimmten Elemente $v_i \in V_i$, $i = 1, ..., m$, mit $v = v_1 + \cdots + v_m$ gilt $f(v) = f(v_1) + \cdots + f(v_m)$.

b) Es gilt $\mathrm{Dim}_K(V_i) = l_i$ und $\chi_{f_i}(t) = (t - \lambda_i)^{l_i}$, $i = 1, ..., m$.

Hat man eine solche Zerlegung gegeben, dann kann man Satz V.7.7 auf die Summanden anwenden und erhält so die gewünschte Normalform. Darauf gehen wir im Anschluss an den Beweis von Satz V.8.5 ausführlich ein.

Definition V.8.4. Es sei f bzw. A wie in Satz V.8.1. Für einen Eigenwert λ von f bzw. A heißt

$$\begin{aligned} V(f,\lambda) &:= \left\{ v \in V \mid \exists i \in \mathbb{N} : (f - \lambda \cdot \mathrm{Id}_V)^i(v) = 0 \right\} = \bigcup_{i\in\mathbb{N}} \mathrm{Ker}\big((f - \lambda \cdot \mathrm{Id}_V)^i\big) \\ \text{bzw.}\quad V(A,\lambda) &:= \left\{ v \in K^n \mid \exists i \in \mathbb{N} : (A - \lambda \cdot \mathbb{E}_n)^i \cdot v = 0 \right\} = \bigcup_{i\in\mathbb{N}} \mathbb{L}\big((A - \lambda \cdot \mathbb{E}_n)^i, 0\big) \end{aligned}$$

der *verallgemeinerte Eigenraum zum Eigenwert* λ.

Satz V.8.5. *Unter den Voraussetzungen von Satz V.8.1 setze man*

$$V_i := V(f,\lambda_i) \quad \textit{bzw.} \quad V_i := V(A,\lambda_i), \quad i = 1, ..., m.$$

Dann gilt:

i) $V = V_1 \oplus \cdots \oplus V_m$ *bzw.* $K^n = V_1 \oplus \cdots \oplus V_m$.

ii) $f(V_i) \subseteq V_i$ *bzw.* $A \cdot V_i \subseteq V_i$, $i = 1, ..., m$.

iii) $\mathrm{Dim}_K(V_i) = l_i$, $i = 1, ..., m$.

Beweis. Wir beweisen die Version für Endomorphismen. **Zu ii).** Es seien λ ein Eigenwert von f und $v \in V(f,\lambda)$ ein Element des zugehörigen verallgemeinerten Eigenraums. Wir wählen eine natürliche Zahl $i \geq 1$, für die $(f-\lambda\cdot \mathrm{Id}_V)^i(v) = 0$ gilt, und berechnen

$$\begin{aligned} 0 &= (f-\lambda\cdot\mathrm{Id}_V)^{i+1}(v) = (f-\lambda\cdot\mathrm{Id}_V)^i(f(v)-\lambda\cdot v) \\ &= (f-\lambda\cdot\mathrm{Id}_V)^i\big(f(v)\big) - \lambda\cdot(f-\lambda\cdot\mathrm{Id}_V)^i(v) \\ &= (f-\lambda\cdot\mathrm{Id}_V)^i\big(f(v)\big). \end{aligned}$$

Dies zeigt $f(v) \in V(f,\lambda)$.

Zu i) und iii). Wir beweisen die Aussagen durch Induktion über m. Für $m = 1$ ist nichts zu zeigen. Im **Induktionsschritt** $m \rightsquigarrow m+1$ setzen wir $g := f-\lambda_1\cdot\mathrm{Id}_V$. Mit den Unterräumen $W_i(\lambda_1) := \mathrm{Ker}(g^i)$, $i \in \mathbb{N}$, erhalten wir die Filtrierung

$$\{0\} = W_0(\lambda_1) \subseteq W_1(\lambda_1) \subseteq W_2(\lambda_1) \subseteq \cdots \subseteq V(f,\lambda_1).$$

Es sei

$$r := \min\big\{\, i \in \mathbb{N} \,|\, W_i(\lambda_1) = V(f,\lambda_1) \,\big\}.$$

(Diese Zahl existiert, weil $V(f,\lambda_1)$ endlichdimensional ist.) Weiter sei

$$U := \mathrm{Bild}(g^r).$$

Behauptung. *Es gilt:*
α) $V = V(f,\lambda_1) \oplus U$,
β) $f\big(V(f,\lambda_1)\big) \subset V(f,\lambda_1)$, $f(U) \subset U$.

Den Beweis von **Teil α)** beginnen wir mit der Beobachtung, dass die Dimensionsformel (Satz III.5.40) für lineare Abbildungen

$$\mathrm{Dim}_K(V) = \mathrm{Dim}_K\big(\mathrm{Ker}(g^r)\big) + \mathrm{Dim}_K\big(\mathrm{Bild}(g^r)\big) = \mathrm{Dim}_K\big(V(f,\lambda_1)\big) + \mathrm{Dim}_K(U)$$

ergibt. Weiter zeigt die Dimensionsformel für Unterräume (Aufgabe III.5.49)

$$\begin{aligned} \mathrm{Dim}_K(V(f,\lambda_1)+U) &= \mathrm{Dim}_K\big(V(f,\lambda_1)\big) + \mathrm{Dim}_K(U) - \mathrm{Dim}_K\big(V(f,\lambda_1)\cap U\big) \\ &= \mathrm{Dim}_K(V) - \mathrm{Dim}_K\big(V(f,\lambda_1)\cap U\big). \end{aligned}$$

Es reicht folglich,

$$V(f,\lambda_1) \cap U = \{0\} \tag{V.10}$$

nachzuweisen. Seien also $w \in V(f,\lambda_1)\cap U$ und $v \in V$ mit $g^r(v) = w$. Wegen

$$0 = g^r(w) = g^{2r}(v)$$

hat man $v \in V(f,\lambda_1)$. Dann gilt aber $w = g^r(v) = 0$ nach Wahl von r.

Zu Teil β). Die Inklusion $f\big(V(f,\lambda_1)\big) \subset V(f,\lambda_1)$ haben wir bereits überprüft, Offenbar gilt auch $g(U) \subset U$, so dass

$$\forall v \in U: \quad f(v) = g(v) + \lambda_1 \cdot v \in U$$

folgt. ✓

Es seien jetzt

$$\begin{array}{rccl} f_{|V(f,\lambda_1)}\colon V(f,\lambda_1) & \longrightarrow & V(f,\lambda_1), & \qquad f_{|U}\colon U \longrightarrow U, \\ v & \longmapsto & f(v) & \qquad\quad v \longmapsto f(v) \end{array}$$

$B' = \{b_1, ..., b_l\}$ eine durchnummerierte Basis von $V(f,\lambda_1)$, $B'' = \{b_{l+1}, ..., b_n\}$ eine durchnummerierte Basis von U und $B = \{b_1, ..., b_l, b_{l+1}, ..., b_n\}$. Dies ist eine durchnummerierte Basis von V. Die entsprechende Darstellungsmatrix nimmt die Gestalt

$$M_B(f) = \left(\begin{array}{c|c} M_{B'}(f_{|V(f,\lambda_1)}) & 0 \\ \hline 0 & M_{B''}(f_{|U}) \end{array}\right) \tag{V.11}$$

an. Aus

$$(f_{|V(f,\lambda_1)} - \lambda_1 \cdot \mathrm{Id}_{V(f,\lambda_1)})^r = 0$$

folgt

$$\chi_{f_{|V(f,\lambda_1)}}(t) = (t-\lambda_1)^{\mathrm{Dim}_K(V(f,\lambda_1))}. \tag{V.12}$$

Schließlich haben wir wegen (V.11)

$$\chi_f(t) = \chi_{f_{|V(f,\lambda_1)}}(t) \cdot \chi_{f_{|U}}(t). \tag{V.13}$$

Die bereits gezeigte Aussage (V.10) beinhaltet, dass λ_1 kein Eigenwert von $f_{|U}$ sein kann. Aus (V.12) und (V.13) folgern wir

$$\mathrm{Dim}_K\big(V(f,\lambda_1)\big) = l_1.$$

Dieselbe Argumentation funktioniert auch für $\lambda_2, ..., \lambda_m$, so dass wir **Teil iii)**, nämlich

$$\mathrm{Dim}_K\big(V(f,\lambda_i)\big) = l_i, \quad i = 1, ..., m,$$

einsehen. Außerdem folgt

$$\chi_{f_{|U}}(t) = (t-\lambda_2)^{l_2} \cdot \cdots \cdot (t-\lambda_m)^{l_m}.$$

Die Induktionsvoraussetzung für $f_{|U}$ besagt

$$U = V(f_{|U}, \lambda_2) \oplus \cdots \oplus V(f_{|U}, \lambda_m)$$

und

$$\mathrm{Dim}_K\big(V(f_{|U}, \lambda_i)\big) = l_i, \quad i = 2, ..., m.$$

Für $i = 2, ..., m$ gilt

$$V(f_{|U}, \lambda_i) \subseteq V(f, \lambda_i).$$

Da beide Vektorräume die Dimension l_i haben, sind sie somit gleich, $i = 2, ..., m$. Damit haben wir endgültig

$$V = V(f, \lambda_1) \oplus U = V(f, \lambda_1) \oplus V(f, \lambda_2) \oplus \cdots \oplus V(f, \lambda_m)$$

nachgewiesen. □

Beweis von Satz V.8.1. Man beachte, dass man auf Grund von Satz V.8.5, ii), die Endomorphismen

$$\begin{aligned} f_i := f_{|V_i} \colon V_i &\longrightarrow V_i \\ v &\longmapsto f(v_i) \end{aligned}$$

definieren kann, $i = 1, ..., m$. Ferner kann man nach Satz V.7.7 eine Basis $B^i = \{ b_1^i, ..., b_{l_i}^i \}$ für V_i wählen, so dass $M_{B^i}(f_i) = J(\underline{\pi}^i, \lambda_i)$ für die entsprechende Partition $\underline{\pi}^i$ von l_i, $i = 1, ..., m$. Für die Basis $B = \{ b_1^1, ..., b_{l_1}^1, ..., b_1^m, ..., b_{l_m}^m \}$ gilt dann

$$M_B(f) = \begin{pmatrix} M_{B^1}(f_1) & & & 0 \\ & M_{B^2}(f_2) & & \\ & & \ddots & \\ 0 & & & M_{B^m}(f_m) \end{pmatrix} = J,$$

und **Teil i)** von Satz V.8.1 ist bewiesen. **Punkt ii)** beweist man analog zu der entsprechenden Eigenschaft in Satz V.7.7. Für **Teil iii)** definieren wir wie zuvor

$$W_i(\lambda_j) := \operatorname{Ker}\big((f - \lambda_j \cdot \mathrm{Id}_V)^i\big), \quad i \in \mathbb{N},\ j = 1, ..., m,$$

und folgern aus dem Beweis von Satz V.7.7, dass

$$\pi_i^j = -\operatorname{Dim}_K\big(W_{i+1}(\lambda_j)\big) + 2 \cdot \operatorname{Dim}_K\big(W_i(\lambda_j)\big) - \operatorname{Dim}_K\big(W_{i-1}(\lambda_j)\big), \quad i = 1, ..., l_j,\ j = 1, ..., m.$$

Man leitet die gewünschte Aussage sofort daraus ab. □

Algorithmus zur Bestimmung der Jordanschen Normalform und des Minimalpolynoms. — Wie die Beweise zur Jordanschen Normalform zeigen, existiert ein Algorithmus zur Bestimmung der Jordanschen Normalform einer gegebenen Matrix A. Wir wollen ihn hier noch einmal erläutern. Eine interessante Konsequenz ist, dass damit auch ein Algorithmus existiert, mit dem man entscheiden kann, ob zwei gegebene Matrizen A und B ähnlich sind.

Schritt 1. — Man bestimme das charakteristische Polynom der Matrix A und schreibe es in der Form

$$\chi_A(t) = (t-\lambda_1)^{l_1} \cdot \cdots \cdot (t-\lambda_m)^{l_m}.$$

Z.B.:

$$A := \begin{pmatrix} 0 & 1 & 2 & 0 & -2 \\ -1 & -2 & 1 & 0 & -1 \\ 1 & 1 & -2 & 0 & 1 \\ -1 & -4 & 0 & 2 & 0 \\ 1 & 1 & -4 & 0 & 3 \end{pmatrix} \in \mathrm{Mat}(5;\mathbb{R}).$$

Man berechnet

$$\chi_A(t) = (t+1)^3 \cdot (t-2)^2.$$

Die Eigenwerte sind also -1 und 2.
Schritt 2. — Man berechne $\mathrm{Dim}_K(W_i(\lambda_j))$ für

$$W_i(\lambda_j) := \left\{ v \in K^n \,|\, (A - \lambda_j \cdot \mathbb{E}_{l_j})^i \cdot v = 0 \right\},\ i = 1,...,l_j,\ j = 1,...,m.$$

Die Partitionen $\underline{\pi}^j = (\pi^j_{l_j},...,\pi^j_1)$, $j = 1,...,m$, und damit die Jordansche Normalform sind durch die Formel

$$\pi^j_i = -\mathrm{Dim}_K\big(W_{i+1}(\lambda_j)\big) + 2\cdot \mathrm{Dim}_K\big(W_i(\lambda_j)\big) - \mathrm{Dim}_K\big(W_{i-1}(\lambda_j)\big), \quad i = 1,...,l_j,$$

gegeben. (In Beispielen kann man diese Zahlen u.U. leicht berechnen (s.u.).)
Berechnung der Dimension von Eig(A, -1) mit dem Gauß-Algorithmus:

$$\begin{pmatrix} -1 & -1 & -2 & 0 & 2 \\ 1 & 1 & -1 & 0 & 1 \\ -1 & -1 & 1 & 0 & -1 \\ 1 & 4 & 0 & -3 & 0 \\ -1 & -1 & 4 & 0 & -4 \end{pmatrix} \overset{\text{I}\rightleftharpoons\text{II}}{\rightsquigarrow} \begin{pmatrix} 1 & 1 & -1 & 0 & 1 \\ -1 & -1 & -2 & 0 & 2 \\ -1 & -1 & 1 & 0 & -1 \\ 1 & 4 & 0 & -3 & 0 \\ -1 & -1 & 4 & 0 & -4 \end{pmatrix}$$

$$\overset{\text{II+I, III+I, IV−I, V+I}}{\rightsquigarrow} \begin{pmatrix} 1 & 1 & -1 & 0 & 1 \\ 0 & 0 & -3 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 1 & -3 & -1 \\ 0 & 0 & 3 & 0 & -3 \end{pmatrix} \overset{\text{II}\rightleftharpoons\text{IV, III}\rightleftharpoons\text{V}}{\rightsquigarrow} \begin{pmatrix} 1 & 1 & -1 & 0 & 1 \\ 0 & 3 & 1 & -3 & -1 \\ 0 & 0 & 3 & 0 & -3 \\ 0 & 0 & -3 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\overset{\text{IV+III}}{\rightsquigarrow} \begin{pmatrix} 1 & 1 & -1 & 0 & 1 \\ 0 & 3 & 1 & -3 & -1 \\ 0 & 0 & 3 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Aus dieser Zeilenstufenform folgt $\mathrm{Dim}_{\mathbb{R}}(\mathrm{Eig}(A,-1)) = 2$. Die Partition zum Eigenwert $\lambda_1 = -1$ muss daher $(0,1,1)$ sein.

Berechnung der Dimension von Eig(A,2) mit dem Gauß-Algorithmus:

$$\begin{pmatrix} 2 & -1 & -2 & 0 & 2 \\ 1 & 4 & -1 & 0 & 1 \\ -1 & -1 & 4 & 0 & -1 \\ 1 & 4 & 0 & 0 & 0 \\ -1 & -1 & 4 & 0 & -1 \end{pmatrix} \overset{\mathrm{I}\rightleftharpoons\mathrm{IV}}{\rightsquigarrow} \begin{pmatrix} 1 & 4 & 0 & 0 & 0 \\ 1 & 4 & -1 & 0 & 1 \\ -1 & -1 & 4 & 0 & -1 \\ 2 & -1 & -2 & 0 & 2 \\ -1 & -1 & 4 & 0 & -1 \end{pmatrix}$$

$$\overset{\mathrm{II}-\mathrm{I},\ \mathrm{III}+\mathrm{I},\ \mathrm{IV}-2\cdot\mathrm{I},\ \mathrm{V}+\mathrm{I}}{\rightsquigarrow} \begin{pmatrix} 1 & 4 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 3 & 4 & 0 & -1 \\ 0 & -9 & -2 & 0 & 2 \\ 0 & 3 & 4 & 0 & -1 \end{pmatrix} \overset{\mathrm{IV}+3\cdot\mathrm{III},\ \mathrm{V}-\mathrm{III}}{\rightsquigarrow} \begin{pmatrix} 1 & 4 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 3 & 4 & 0 & -1 \\ 0 & 0 & 10 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\overset{\mathrm{IV}+10\cdot\mathrm{II}}{\rightsquigarrow} \begin{pmatrix} 1 & 4 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 3 & 4 & 0 & -1 \\ 0 & 0 & 0 & 0 & 9 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \overset{\mathrm{II}\rightleftharpoons\mathrm{III}}{\rightsquigarrow} \begin{pmatrix} 1 & 4 & 0 & 0 & 0 \\ 0 & 3 & 4 & 0 & -1 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 9 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Man liest $\mathrm{Dim}_{\mathbb{R}}(\mathrm{Eig}(A,2)) = 1$ ab, so dass $\underline{\pi}^2 = (1,0)$ die Partition zum Eigenwert $\lambda_2 = 2$ ist.

Die Jordansche Normalform der Matrix A ist also die Matrix

$$J := \left(\begin{array}{ccc|cc} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ \hline 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{array}\right).$$

Schritt 3. — Das Minimalpolynom kann nun abgelesen werden:

$$\chi_A(t) = \chi_J(t) = (t+1)^2 \cdot (t-2)^2.$$

Man kann weiterhin eine Matrix T mit $T^{-1} \cdot A \cdot T = J$ suchen.

Schritt 4. — Man wende den Algorithmus aus dem Satz über nilpotente Endomorphismen sinngemäß auf die verallgemeinerten Eigenräume an.

Für A gilt $V(A,-1) = \left\{ v \in \mathbb{R}^3 \mid (A + \mathbb{E}_5)^2 \cdot v = 0 \right\}$. Man hat

$$(A+\mathbb{E}_5)^2 = \begin{pmatrix} 1 & 1 & 2 & 0 & -2 \\ -1 & -1 & 1 & 0 & -1 \\ 1 & 1 & -1 & 0 & 1 \\ -1 & -4 & 0 & 3 & 0 \\ 1 & 1 & -4 & 0 & 4 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 & 9 & 0 & -9 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -9 & -6 & 9 & 6 \\ 0 & 0 & -9 & 0 & 9 \end{pmatrix}.$$

Wir müssen in dem **dreidimensionalen** Raum $V(A,-1)$ ein lineares Komplement $U_2(-1)$ zu dem **zweidimensionalen** Unterraum $\mathrm{Eig}(A,-1)$ finden. Es genügt daher, einen Vektor $v \in V(A,-1)$ zu finden, der nicht in $\mathrm{Eig}(A,-1)$ liegt, z.B. $v = e_1$, und $U_2(-1) = \langle v \rangle$ zu setzen. Unsere ersten beiden Basisvektoren sind also

$$b_1 := (A + \mathbb{E}_5) \cdot e_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \\ 1 \end{pmatrix} \quad \text{und} \quad b_2 := e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Nun müssen wir einen erzeugenden Vektor $b_3 \in \mathrm{Eig}(A,-1)$ finden, so dass $\mathrm{Eig}(A,-1) = \langle b_1 \rangle \oplus \langle b_3 \rangle$. Offenbar können wir

$$b_3 := \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

wählen. Wir haben $V(A,2) = \left\{ v \in \mathbb{R}^3 \mid (A - 2 \cdot \mathbb{E}_5)^2 \cdot v = 0 \right\}$. Dabei gilt

$$(A - 2 \cdot \mathbb{E}_5)^2 = \begin{pmatrix} -2 & 1 & 2 & 0 & -2 \\ -1 & -4 & 1 & 0 & -1 \\ 1 & 1 & -4 & 0 & 1 \\ -1 & -4 & 0 & 0 & 0 \\ 1 & 1 & -4 & 0 & 1 \end{pmatrix}^2 = \begin{pmatrix} 3 & -6 & -3 & 0 & 3 \\ 6 & 15 & -6 & 0 & 6 \\ -6 & -6 & 15 & 0 & -6 \\ 6 & 15 & -6 & 0 & 6 \\ -6 & -6 & 15 & 0 & -6 \end{pmatrix}.$$

Man sieht, dass der Vektor

$$b_5 := \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}$$

zwar in $V(A,2)$ nicht aber in $\mathrm{Eig}(A,2)$ liegt, i.e. $V(A,2) = U_2(2) \oplus \mathrm{Eig}(A,2)$ mit $U_2(2) = \langle b_5 \rangle$. Wir ergänzen unsere Basis mit dem Vektor

$$b_4 := (A - 2 \cdot \mathbb{E}_5) \cdot b_5 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix}.$$

Damit gilt

$$T := (b_1|b_2|b_3|b_4|b_5) = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 & -1 \end{pmatrix}$$

und

$$T^{-1} = \begin{pmatrix} 0 & -1 & 0 & 0 & 0 \\ 1 & 1 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{pmatrix}.$$

Man überprüft:

$$\begin{aligned} & T^{-1} \cdot A \cdot T \\ = & \begin{pmatrix} 0 & -1 & 0 & 0 & 0 \\ 1 & 1 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 2 & 0 & -2 \\ -1 & -2 & 1 & 0 & -1 \\ 1 & 1 & -2 & 0 & 1 \\ -1 & -4 & 0 & 2 & 0 \\ 1 & 1 & -4 & 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 & -1 \end{pmatrix} \\ = & \begin{pmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix} \cdot \surd \end{aligned}$$

Beispiel V.8.6. Wir haben mehrmals erwähnt, dass es nicht ähnliche Matrizen mit demselben charakteristischen Polynom gibt. Mit Hilfe der Jordanschen Normalform können wir uns einen genaueren Überblick verschaffen. Es sei $\chi(t) = (t-\lambda_1)^{l_1} \cdot \dots \cdot (t-\lambda_m)^{l_m}$. Die Jordanschen Normalformen von Matrizen A mit $\chi_A(t) = \chi(t)$ stehen in Bijektion zu Tupeln $(\underline{\pi}^1, ..., \underline{\pi}^m)$, in denen $\underline{\pi}^j = (\pi^j_{l_j}, ..., \pi^j_1)$ eine Partition von l_j ist, $j = 1, ..., m$. Damit kann man sofort eine Liste erstellen.

Wir betrachten das Beispiel

$$\chi(t) = (t+2)^3(t-5)^3.$$

Die möglichen Partitionen von 3 sind $(1,0,0)$, $(0,1,1)$, $(0,0,3)$, so dass sich insgesamt neun Möglichkeiten ergeben:

$\underline{\pi}^1$	$\underline{\pi}^2$	JNF	μ
(1,0,0)	(1,0,0)	$\left(\begin{array}{ccc\|ccc} -2 & 1 & 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{array}\right)$	$(t+2)^3 \cdot (t-5)^3$
(1,0,0)	(0,1,1)	$\left(\begin{array}{ccc\|ccc} -2 & 1 & 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{array}\right)$	$(t+2)^3 \cdot (t-5)^2$
(1,0,0)	(0,0,3)	$\left(\begin{array}{ccc\|ccc} -2 & 1 & 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{array}\right)$	$(t+2)^3 \cdot (t-5)$
(0,1,1)	(1,0,0)	$\left(\begin{array}{ccc\|ccc} -2 & 1 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{array}\right)$	$(t+2)^2 \cdot (t-5)^3$
(0,1,1)	(0,1,1)	$\left(\begin{array}{ccc\|ccc} -2 & 1 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{array}\right)$	$(t+2)^2 \cdot (t-5)^2$.
(0,1,1)	(0,0,3)	$\left(\begin{array}{ccc\|ccc} -2 & 1 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{array}\right)$	$(t+2)^2 \cdot (t-5)$
(0,0,3)	(1,0,0)	$\left(\begin{array}{ccc\|ccc} -2 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{array}\right)$	$(t+2) \cdot (t-5)^3$
(0,0,3)	(0,1,1)	$\left(\begin{array}{ccc\|ccc} -2 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{array}\right)$	$(t+2) \cdot (t-5)^2$
(0,0,3)	(0,0,3)	$\left(\begin{array}{ccc\|ccc} -2 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{array}\right)$	$(t+2) \cdot (t-5)$

Aufgabe V.8.7.
Gegeben sei

$$A := \begin{pmatrix} 3 & 5 & -16 & 6 \\ 7 & 2 & -17 & 7 \\ 1 & 0 & -4 & 1 \\ -5 & -5 & 15 & -8 \end{pmatrix} \in \mathrm{Mat}(4;\mathbb{R})$$

a) Berechnen Sie das charakteristische Polynom von A und zerlegen Sie es in Linearfaktoren.
b) Geben Sie die Jordansche Normalform und das Minimalpolynom μ_A von A an.
c) Bestimmen Sie eine Matrix $S \in \mathrm{GL}_4(\mathbb{R})$, so dass $S \cdot A \cdot S^{-1}$ die Jordansche Normalform von A ist.

Literaturhinweise

[1] T. Arens, F. Hettlich, Ch. Karpfinger, U. Kockelkorn, K. Lichtenegger, H. Stachel, *Mathematik*, 2. überarb. Aufl., Heidelberg, Spektrum Akademischer Verlag, 2012, xiv+1506 S. ISBN 978-3-642-40472-6/set.

[2] E. Brieskorn, *Lineare Algebra und analytische Geometrie.* I., mit historischen Anmerkungen von Erhard Scholz, Friedr. Vieweg & Sohn, 1983. viii+636 S. ISBN: 3-528-08561-4.

[3] G. Fischer, *Lineare Algebra. Eine Einführung für Studienanfänger*, 16. überarb. und erw. Aufl., Vieweg Studium: Grundkurs Mathematik, Wiesbaden, Friedr. Vieweg & Sohn, 2008, xxii+384 S. ISBN 978-3-658-03944-8/pbk.

[4] U. Friedrichsdorf, A. Prestel, *Mengenlehre für den Mathematiker*, Vieweg Studium **58**, Grundkurs Mathematik, Braunschweig/Wiesbaden, Friedr. Vieweg & Sohn, 1985, vi+103 S. ISBN 3-528-07258-X.

[5] St. Funken, *Numerik* I — *Einführung in die Numerische Lineare Algebra*, Vorlesungsskript,
`http://www.uni-ulm.de/fileadmin/website_uni_ulm/mawi.inst.070/ws11_12/Numerik1/numerik1a_skript.pdf`

[6] W. Klingenberg, *Lineare Algebra und Geometrie*, 3. Aufl., Springer-Lehrbuch, Berlin, Springer-Verlag, 1992, xiii+293 S. ISBN 978-3-642-77646-5.

[7] H. Kurzweil, *Endliche Körper. Verstehen, Rechnen, Anwenden*, 2. überarb. Aufl., Springer-Lehrbuch, Berlin, Springer, 2008, xi+178 S. ISBN 978-3-540-79597-1/pbk.

[8] A. Schmitt, *Analysis* I, Vorlesungsskript,
`http://userpage.fu-berlin.de/~aschmitt`.

[9] A. Schmitt, *Algebra und Zahlentheorie*, Vorlesungsskript, http://userpage.fu-berlin.de/~aschmitt.

[10] A. Schmitt, *Algebra* I, Vorlesungsskript, http://userpage.fu-berlin.de/~aschmitt.

[11] A. Schmitt, *Funktionentheorie*, Vorlesungsskript, http://userpage.fu-berlin.de/~aschmitt.

[12] H.-P. Tuschik, H. Wolter, *Mathematische Logik – kurzgefaßt*, 2. Aufl., Spektrum Hochschultaschenbuch, Heidelberg, Spektrum Akademischer Verlag, 2002, viii+214 S. ISBN 3-8274-1387-7/pbk.

Stichwortverzeichnis

Abbildung, 36
    duale —, 119
    identische —, 38
    inverse —, 41
    Polynom- —, 161
    transponierte —, 119
abelsche Gruppe, 44, 49
Abspaltung eines Linearfaktors, 162
Addition
    in einem Ring, 48
    in einem Vektorraum, 59
    von Matrizen, 166
    von Tupeln, 8
ähnliche Matrizen, 156
algebraisch abgeschlossener Körper, 163
algebraische Vielfachheit eines Eigenwerts, 171
Algorithmus für die Jordansche Normalform, 193, 199
allgemeine lineare Gruppe, 109
    eines Vektorraums, 86
alternierend, 133
alternierende Gruppe, 131
äquivalente Matrizen, 155
Äquivalenzklasse, 52
Äquivalenzrelation, 52, 155, 156
Assoziativgesetz, 14
    der Addition, 14
    der Skalarmultiplikation, 14, 59
    für Abbildungen, 41
    für Gruppen, 43
assoziierte lineare Abbildung, 110
aufzählende Form, 33
Austauschlemma, 72
Auswertungsabbildung, 92, 178
axiomatische Mengenlehre, 32

Basis, 68
    duale —, 120
    durchnummerierte —, 86
bijektiv, 39
Bild, 36, 37

Cantor, 31
Cayley, 181
    Satz von — -Hamilton, 181
charakteristisches Polynom
    einer Matrix, 167
    eines Endomorphismus, 171
    Nullstellen des —n —s, 168
Cramer, 149
Cramersche Regel
    für die Matrixinversion, 149
    für Gleichungssysteme, 150

Darstellungsmatrix, 110
De Morgan, 36
De Morganschen Gesetze, 36
Descartes, 35

Determinante, 3, 133, 143
eines Endomorphismus, 142
und Rang, 106
Vandermonde- —, 153
Determinanten von Elementarmatrizen, 136
Determinantenabbildung, 132, 166
Eindeutigkeit der —, 138
Existenz der —, 138
Diagonale, 38
diagonalisierbare Matrix, 158
diagonalisierbarer Endomorphismus, 158
Diagonalisierbarkeit
Kriterium für —, 173
Diagonalisierung
Verfahren zur —, 175
Diagonalmatrix, 158
Differenz, 34
Dimension, 7, 26, 72
Dimensionsformel
für lineare Abbildungen, 90, 98
für Unterräume, 93
direkte Summe, 63
von Unterräumen, 64
disjunkte Vereinigung, 53
Diskriminante, 175
Distributivgesetz, 14
erstes —, 14
zweites —, 14
Distributivgesetze
der Mengenlehre, 34, 35
der Skalarmultiplikation, 60
in einem Ring, 49
Division mit Rest, 48, 55
im Polynomring, 161
Drehung, 81, 170
Dreiecksmatrix
obere —, 152, 184
duale
Abbildung, 119
Basis, 120
Dualraum, 119
durchnummerierte Basis, 86
Durchschnitt, 34

Ebene, 11, 13
Eigenraum, 157
verallgemeinerter —, 196
Eigenvektor, 157
Eigenwert, 157
algebraische Vielfachheit eines —s, 171
geometrische Vielfachheit eines —s, 171
eindeutig lösbar, 8
einfache Nullstelle, 175
Einheit, 117
Einheitsmatrix, 80
Eins-Zwei-Drei-Verfahren, 164
Element einer Menge, 32
Elementarmatrix, 104, 166, 173
Elementarmatrizen und Zeilen- und Spaltenoperationen, 105
endlichdimensional, 72
endlicher Körper, 165
Endomorphismenring
eines Vektorraums, 85
Endomorphismus, 85
diagonalisierbarer —, 158
Eigenvektor eines —, 157
Eigenwert eines —, 157
halbeinfacher —, 158
nilpotenter —, 186
Potenzen eines —, 179
trigonalisierbarer —, 184
erweiterte Koeffizientenmatrix, 19
Erzeugendensystem, 65
minimales —, 70
Evaluationsabbildung, 178
Extensionalitätsaxiom, 32

Faktorgruppe, 54

Familie
  von Elementen, 38
  von Teilmengen, 38
fast alle, 63
Fehlstand, 128
frei, 67
Fundamentalsatz der Algebra, 163

ganze Zahlen, 32
Gauß, 1
Gauß-Algorithmus, 1, 153
geometrische Vielfachheit
  eines Eigenwerts, 171
  ist kleiner gleich algebraische Vielfachheit, 172
  und Ähnlichkeit, 172
geordnetes Paar, 35
Gerade, 5, 11, 13
gerade Permutation, 130
Gleichungssystem, 9
  homogenes —, 6, 9
  inhomogenes —, 9
Grad eines Polynoms, 160
Graph, 37
Gruppe, 43
  abelsche —, 44, 49
  allgemeine lineare —, 109
  alternierende —, 131
  kommutative —, 44
  symmetrische —, 123
  zyklische —, 47
Gruppenhomomorphismus, 125

halbeinfacher Endomorphismus, 158
Hamilton, 181
  Satz von Cayley- —, 181
homogenes Gleichungssystem, 6, 9
Homomorphiesatz, 96
Homomorphismus, 76
  Gruppen- —, 125
Homothetie, 77

identische Abbildung, 38
Identität, 38, 45
Indexmenge, 38
Induktionsanfang, 36
Induktionsschritt, 36
inhomogenes Gleichungssystem, 9
injektiv, 39
inverse
  Abbildung, 41, 45
  Matrix, 107
inverses Element, 14
  in einer Gruppe, 44
invertierbare
  Matrix, 107
  Matrizen über $\mathbb{Z}$, 173
isomorph, 87
Isomorphismus, 79, 85
  von Ringen, 57

Jordan, 189
Jordan-Block, 189
Jordan-Blockmatrix, 189
Jordansche Normalform, 196
  Algorithmus für die — —, 193, 199
  Satz über die — —, 195

kartesisches Produkt, 35
Kern, 78
Kirchhoff, 29
Kirchhoffsche Regeln, 29
Klein, 47
kleiner Satz von Fermat, 57
Kleinsche Vierergruppe, 47
Knotenregel, 29
Koeffizient, 9
Koeffizientenmatrix, 19
Kokern, 98
kommutative Gruppe, 44
kommutativer Ring, 49, 166
kommutatives Diagramm, 110
Kommutativgesetz, 13

in einer Gruppe, 44
Komplement, 36
komplementäre Matrix, 144, 167
Komposition von Abbildungen, 39
Komprehension, 33
kongruent, 94
Koordinaten, 2
Körper, 1
algebraisch abgeschlossener —, 163
der komplexen Zahlen, 50
endlicher —, 165
mit vier Elementen, 51
mit zwei Elementen, 50
Kriterium
für Diagonalisierbarkeit, 173
für Trigonalisierbarkeit, 184
Kronecker, 80
Kronecker-Delta, 80

Laplace, 146
Laplace-Entwicklung, 146, 167
leere Menge, 2, 32
Leibniz, 138
Leibnizformel, 138
linear
abhängig, 67
unabhängig, 67
unabhängige Teilmenge
maximale — — —, 70
lineare
Abbildung, 76
assoziierte — —, 110
Hülle, 14, 64, 66
linearer
Isomorphismus, 79
Teilraum, 11, 61
Unterraum, 11
lineares
Erzeugnis, 64
Gleichungssystem, 9
Komplement, 88
Linearfaktor
Abspaltung eines —s, 162
zerfällt in —en, 163
Linearkombination, 64
Linkstranslation, 48
Lösungsmenge, 2, 9

Maschenregel, 29
Matrix, 3, 9
diagonalisierbare —, 158
Eigenvektor einer —, 157
Eigenwert einer —, 157
inverse —, 107
invertierbare —, 107
komplementäre —, 144, 167
nilpotente —, 186
Potenzen einer —, 179
Spur einer —, 167, 173
transponierte —, 120
trigonalisierbare —, 184
Matrixmultiplikation, 166
Matrizen, 166
Addition von —, 166
ähnliche —, 156
äquivalente —, 155
Skalarmultiplikation bei —, 166
über $\mathbb{Z}$, 173
über einem Ring, 166
maximale linear unabhängige Teilmenge, 70
Menge, 31
der Äquivalenzklassen, 52
der ganzen Zahlen, 32
der natürlichen Zahlen, 32
der rationalen Zahlen, 32
der reellen Zahlen, 32
leere —, 2, 32
minimales Erzeugendensystem, 70
Minimalpolynom, 182
Nullstellen des —s, 183
teilt charakteristisches Polynom, 183

Modul über einem Ring, 166
Multiplikation
in einem Ring, 48
mit einer natürlichen Zahl, 38
Multiplizität einer Nullstelle, 162

natürliche Zahlen, 32
Neutralelement, 13
in einer Gruppe, 44
nilpotente Matrix, 186
nilpotenter Endomorphismus, 186
Normalteiler, 55
normiert, 133
normiertes Polynom, 167
nulldimensional, 7
Nullring, 49, 85
Nullstelle
eines Polynoms, 162
einfache —, 175
Multiplizität einer —, 162
Vielfachheit einer —, 162
Nullstellen
des charakteristischen Polynoms, 168
des Minimalpolynoms, 183
Nullteilerfreiheit, 49
Nulltupel, 8
Nullvektor, 8

obere Dreiecksmatrix, 152, 184
Ohm, 29
Ohmsche Gleichung, 29

parallel, 6
Parallelverschiebung, 13
Parität, 128
Partition
einer natürlichen Zahl, 189
Jordan-Blockmatrix zu einer —, 189
Peano-Axiome, 36
Permutation, 123
gerade —, 130
ungerade —, 130
Polynom, 160
charakteristisches —, 167, 171
Grad eines —s, 160
normiertes —, 167
Nullstelle eines —s, 162
Polynomabbildung, 161
Polynome vom Grad zwei, 164
Polynomring, 160
Division mit Rest im —, 161
Potenzen
einer Matrix, 179
eines Endomorphismus, 179
Potenzmenge, 35
Prinzip
der vollständigen Induktion, 36
vom kleinsten Element, 48
Produkt, 9
von Matrizen, 101, 166

Quotientenvektorraum, 94

Rang
einer linearen Abbildung, 81
einer Matrix, 91
Spalten- —, 75
Zeilen- —, 75
rationale Zahlen, 32
Rechtstranslation, 48
reelle Zahlen, 32
Reflexivität einer Relation, 52
Regel
Cramersche —, 149, 150
von Sarrus, 140
Relation, 52
Repräsentantensystem
vollständiges —, 52
Ring, 48
der formalen Potenzreihen, 52
der $(n \times n)$-Matrizen, 102
kommutativer —, 49, 166
Modul über einem —, 166

Polynom- —, 160

Sarrus, 140
Regel von —, 140
Satz
über die Jordansche Normalform, 195
von Cayley–Hamilton, 181
Schar, 8
Schnittpunkt, 6
Skalarmultiplikation, 9, 59
bei Matrizen, 166
Spaltenoperation, 28
Spaltenrang, 75
gleich Zeilenrang, 91, 122
Spaltenvektor, 8
spezielle Lösung, 11
Spur einer Matrix, 167, 173
Steinitz, 73
Steinitzscher Austauschsatz, 73
Summe, 8
von Unterräumen, 64
surjektiv, 39
Symmetrie einer Relation, 52
symmetrische
Gruppe, 123
als Untergruppe der allgemeinen linearen Gruppe, 128
Mengendifferenz, 35

Teilmenge, 32
Teilraum, 11
Transitivität einer Relation, 52
transponierte
Abbildung, 119
Matrix, 120
Transposition, 124
trigonalisierbare Matrix, 184
trigonalisierbarer Endomorphismus, 184
Trigonalisierbarkeit
Kriterium für —, 184

triviale Lösung, 6, 25
Tupel, 8
Addition, 8
Null- —, 8
Skalarmultiplikation, 9

Umkehrabbildung, 41
unabhängige Bedingungen, 7
Unbestimmte, 160
unendlichdimensional, 72
ungerade Permutation, 130
unterbestimmt, 25
Untergruppe, 47
Untermenge, 32
Unterraum, 11
Unterring, 52
Untervektorraum, 61
Urbild, 37

Vandermonde, 153
Vandermonde-Determinante, 153
Vektor
Null- —, 8
Vektorraum, 10, 59
der Matrizen, 61
der Zeilenvektoren, 75
Vektorraum-Homomorphismus, 76
verallgemeinerter Eigenraum, 196
Zerlegung in —e —e, 196
Vereinigung, 34
disjunkte —, 53
Verfahren zur Diagonalisierung, 175
Verknüpfung von Abbildungen, 39
Verknüpfungstabelle, 45
Vielfachheit
algebraische —, 171
einer Nullstelle, 162
geometrische —, 171
vollständiges Repräsentantensystem, 52
Vorzeichen, 128
einer Transposition, 130

Wohldefiniertheit, 54

Zeilenrang, 75
    gleich Spaltenrang, 91, 122
Zeilenstufenform, 16, 26
    ist nicht eindeutig, 19
Zeilenumformung, 14
zentrische Streckung, 9
zerfällt in Linearfaktoren, 163
Zerlegung in verallgemeinerte Eigenräume, 196
Zorn, 123
Zornsches Lemma, 123
Zwischenwertsatz, 39
zyklische Gruppe, 47