Ralf Gerkmann

Mathematisches Institut
Ludwig-Maximilians-Universität München

*Vorlesung im Sommersemester 2019*

# *Lineare Algebra*

*(Mathematik II für das gymnasiale Lehramt)*

(Version vom 24. Juli 2019)

## Inhaltsverzeichnis

# § 0. Einleitung

***Inhaltsübersicht***

In diesem Abschnitt soll kurz wiederholt werden, in welcher Form lineare Gleichungssysteme im Schulunterricht behandelt wurden. Dabei legen wir allerdings von Anfang an Wert auf die korrekte Verwendung mathematischer Notation, vor allem bei der Angabe von Lösungsmengen.

Anschließend geben wir einen kurzen Abriss der Vorlesungsinhalte. Außerdem erläutern wir die Bedeutung der Linearen Algebra für andere Teilgebiete der Mathematik, die zum großen Teil Gegenstand späterer Vorlesungen sind, und gehen auch kurz auf Anwendung der Linearen Algebra außerhalb der Mathematik ein.

Im Schulunterricht werden Lineare Gleichungssysteme häufig als Textaufgaben formuliert, die zunächst in formale Schreibweise übersetzt werden müssen. Eine solche Aufgabe könnte zum Beispiel lauten:

„Peter und sein Vater Wolfgang sind zusammen 56 Jahre alt. Wolfgang ist zweieinhalb mal so alt wie Peter. Wie alt sind die beiden?“

Um die Aufgabe zu lösen, führt man zunächst zwei Variablen oder „Unbekannte“ ein, eine Variable $x$ für das Alter von Peter und eine weitere $y$ für das Alter von Wolfgang. Anschließend übersetzt man die beiden umgangssprachlich formulierten Sätze in Gleichungen. Der erste Satz (Gesamtalter 56) liefert die Gleichung $x + y = 56$, der zweite (Wolfgang zweieinhalb mal so alt wie Peter) die Gleichung $y = \frac{5}{2}x$. Das zu lösende lineare Gleichungssystem lautet also

$$x + y = 56 \quad , \quad y = \tfrac{5}{2}x.$$

Häufig werden in der Schule zwei Lösungsverfahren für lineare Gleichungssysteme behandelt: das Einsetzungsverfahren und das Eliminationsverfahren. Mit dem ***Einsetzungsverfahren*** lässt sich das angegebene lineare Gleichungssystem besonders schnell lösen: Man setzt einfach $y = \frac{5}{2}x$ in die erste Gleichung ein und löst nach $x$ auf

$$x + (\tfrac{5}{2}x) = 56 \quad \Leftrightarrow \quad \tfrac{7}{2}x = 56 \quad \Leftrightarrow \quad x = \tfrac{2}{7} \cdot 56 = 16$$

und erhält den $y$-Wert anschließend durch Einsetzen in die zweite Gleichung: $y = \frac{5}{2}x = \frac{5}{2} \cdot 16 = 40$. Insgesamt ist also $x = 16$, $y = 40$ die gesuchte Lösung.

Auch Gleichungssysteme mit mehrer als zwei Variablen können mit dem Einsetzungsverfahren gelöst werden, zum Beispiel

$$\begin{array}{rcrcrcr} x & - & 2y & - & z & = & 3 \\ -3x & - & 3y & + & z & = & 1 \\ -3x & + & 2y & + & 2z & = & 5 \end{array}$$

Zunächst löst man zum Beispiel die erste Gleichung nach $x$ auf und erhält $x = 3 + 2y + z$.

Einsetzen in die zweite Gleichung ergibt

$$-3(3+2y+z)-3y+z=1 \quad\Leftrightarrow\quad -9-6y-3z-3y+z=1 \quad\Leftrightarrow\quad -9-9y-2z=1 \quad\Leftrightarrow\quad -9y-2z=10$$

und Einsetzen in die dritte Gleichung entsprechend

$$-3(3+2y+z)+2y+2z=5 \quad\Leftrightarrow\quad -9-6y-3z+2y+2z=5 \quad\Leftrightarrow\quad -9-4y-z=5 \quad\Leftrightarrow\quad -4y-z=14.$$

Wir erhalten also ein neues lineares Gleichungssystem, in dem nur noch zwei Variablen vorkommen, nämlich

$$-9y-2z=10 \quad , \quad -4y-z=14.$$

Mit diesem System wiederholen wir die Prozedur. Die zweite Gleichung des neuen Systems kann umgeformt werden zu $-z=14+4y \Leftrightarrow z=-14-4y$, und Einsetzen in die erste Gleichung liefert

$$-9y-2(-14-4y)=10 \quad\Leftrightarrow\quad -9y+28+8y=10 \quad\Leftrightarrow\quad -y=-18 \quad\Leftrightarrow\quad y=18.$$

Erneutes Einsetzen ergibt $z=-14-4\cdot 18=-14-72=-86$ und $x=3+2y+z=3+2\cdot 18+(-86)=3+36-86=-47$. Insgesamt erhalten wir also die Lösung $x=-47$, $y=18$, $z=-86$.

Kommen wir nun zum zweiten gängigen Lösungsverfahren, dem ***Eliminationsverfahren***. Hier wird das gesamte lineare Gleichungssystem solange kontinuierlich umgeformt, bis die Lösung direkt abgelesen werden kann. Beim Gleichungssystem in zwei Variablen erscheint dies zunächst aufwändiger.

$$\begin{array}{rcrcrl} x & + & y & = & 56 & \\ \frac{5}{2}x & - & y & = & 0 & |\cdot\mathbf{2} \end{array} \quad\Leftrightarrow\quad \begin{array}{rcrcrl} x & + & y & = & 56 & \\ 5x & - & 2y & = & 0 & |+(\mathbf{-5})\cdot\mathbf{I} \end{array} \quad\Leftrightarrow$$

$$\begin{array}{rcrcrl} x & + & y & = & 56 & \\ & & -7y & = & -280 & |\cdot(-\frac{1}{7}) \end{array} \quad\Leftrightarrow\quad \begin{array}{rcrcrl} x & + & y & = & 56 & |+(\mathbf{-1})\cdot\mathbf{II} \\ & & y & = & 40 & \end{array} \quad\Leftrightarrow\quad \begin{array}{rcr} x & = & 16 \\ y & = & 40 \end{array}$$

Durch die Notation hinter dem senkrechten Strich „|“ wird der jeweilige Rechenschritt angegeben. Beispielsweise wird im ersten Schritt die zweite Gleichung mit den Wert 2 multipliziert. Im zweiten Schritt wird zur zweiten Gleichung das (−5)-fache der ersten Gleichung addiert. (Wir werden später eine etwas systematischere und allgemeinere Notation für die Rechenschritte einführen.) Beim System in drei Variablen ist der Schreibaufwand im Vergleich zum Einsetzungsverfahren in etwa gleich groß, aber die gesamte Rechnung ist deutlich übersichtlicher.

$$\begin{array}{rcrcrcrl} x & - & 2y & - & z & = & 3 & \\ -3x & - & 3y & + & z & = & 1 & |+\mathbf{3}\cdot\mathbf{I} \\ -3x & + & 2y & + & 2z & = & 5 & |+\mathbf{3}\cdot\mathbf{I} \end{array} \quad\Leftrightarrow\quad \begin{array}{rcrcrcrl} x & - & 2y & - & z & = & 3 & \\ & & -4y & - & z & = & 14 & |\cdot(-\frac{1}{4}) \\ & & -9y & - & 2z & = & 10 & \end{array} \quad\Leftrightarrow$$

$$\begin{array}{rcrcrcrl} x & - & 2y & - & z & = & 3 & \\ & & y & + & \frac{1}{4}z & = & -\frac{7}{2} & \\ & & -9y & - & 2z & = & 10 & |+\mathbf{9}\cdot\mathbf{II} \end{array} \quad\Leftrightarrow\quad \begin{array}{rcrcrcrl} x & - & 2y & - & z & = & 3 & \\ & & y & + & \frac{1}{4}z & = & -\frac{7}{2} & \\ & & & & \frac{1}{4}z & = & -\frac{43}{2} & |\cdot\mathbf{4} \end{array} \quad\Leftrightarrow$$

$$\begin{array}{rcrcrcrl} x & - & 2y & - & z & = & 3 & |+\mathbf{III} \\ & & y & + & \frac{1}{4}z & = & -\frac{7}{2} & |+(-\frac{1}{4})\cdot\mathbf{III} \\ & & & & z & = & -86 & \end{array} \iff \begin{array}{rcrcrcrl} x & - & 2y & & & = & -83 & |+\mathbf{2}\cdot\mathbf{II} \\ & & y & & & = & 18 & \\ & & & & z & = & -86 & \end{array}$$

$$\iff \begin{array}{rcrcrcr} x & & & & & = & -47 \\ & & y & & & = & 18 \\ & & & & z & = & -86 \end{array}$$

Wir werden in den nächsten Kapitel sehen, dass der Schreibaufwand durch die sog. ***Matrixschreibweise*** noch erheblich reduziert werden kann. In dieser Form ist das Eliminationsverfahren, besonders bei größeren Systemen, dem Einsetzungsverfahren eindeutig überlegen. Auch aus theoretischer Sicht bietet es Vorteile. Der formale Nachweis, dass jedes lineare Gleichungssystem in endlich vielen festgelegten Schritten gelöst werden kann, und dass das Verfahren korrekt ist (also jeder einzelne Schritt die Lösungsmenge nicht verändert), ist beim Eliminationsverfahren deutlich einfacher und übersichtlicher als beim Einsetzungsverfahren.

An dieser Stelle sollten wir auch kurz wiederholen, wie die Lösungsmenge eines linearen Gleichungssystems definiert ist. Die Lösungsmenge eines Systems in zwei Unbekannten ist die Menge aller Paare $(x,y) \in \mathbb{R}^2$, die sämtliche Gleichungen des Systems erfüllen. Bei drei Variablen betrachtet man entsprechend die Menge aller Tripel $(x,y,z) \in \mathbb{R}^3$ mit dieser Eigenschaft. Die Lösungsmenge des ersten Beispielsystems ist also geben durch

$$\mathscr{L} \quad = \quad \{(16,40)\}.$$

Wichtig ist hierbei die Verwendung der Tupelschreibweise und die Einhaltung der Reihenfolge. So wäre es falsch, $\{16,40\}$ als Lösungsmenge anzugeben. Die zweielementige Lösungsmenge $\{16,40\}$ könnte höchstens dadurch zu Stande kommen, dass man an Stelle eines linearen Gleichungssystems in *zwei* Variablen eine quadratische Gleichung in *einer* Variablen betrachtet, zum Beispiel die Gleichung $x^2 - 56x + 640 = 0$. Solche Polynomgleichungen werden aber für den größten Teil der Vorlesung keine Rolle spielen. Ebenso ist die Lösungsmenge des zweiten Beispielsystems gleich der einelementigen Menge $\{(-47,18,-86)\}$, und nicht etwa die dreielementige Menge $\{-47,18,-86\}$.

Bereits aus dem Schulunterricht ist auch bekannt, dass Gleichungssysteme nicht immer eindeutig lösbar sind, die Lösungsmenge also nicht immer aus genau einem Lösungstupel besteht. Betrachten wir beispielsweise das Gleichungssystem $x + y = 2$, $x - y = 6$, $3x + y = 5$. Durch Anwendung des Eliminationsverfahrens erhält man

$$\begin{array}{rcrcrl} x & + & y & = & 2 & \\ x & - & y & = & 6 & |+(\mathbf{-1})\cdot\mathbf{I} \\ 3x & + & y & = & 5 & |+(\mathbf{-3})\cdot\mathbf{I} \end{array} \iff \begin{array}{rcrcrl} x & + & y & = & 2 & \\ & & -2y & = & 4 & \\ & & -2y & = & -1 & |+(\mathbf{-1})\cdot\mathbf{II} \end{array} \iff \begin{array}{rcrcr} x & + & y & = & 2 \\ & & y & = & -2 \\ & & 0 & = & -5 \end{array}$$

Die dritte Gleichung im letzten System könnte ausführlicher in der Form $0\cdot x + 0\cdot y = -5$ geschrieben werden. Dadurch wird noch deutlicher, dass die Gleichung, und damit auch das gesamte System, durch kein Zahlenpaar $(x,y) \in \mathbb{R}^2$ erfüllt werden kann. Die Lösungsmenge des linearen Gleichungssystem ist damit gegeben durch $\mathscr{L} = \varnothing$, die leere Menge.

Auch hier ist auf die korrekte Schreibweise zu achten. Beispielsweise wäre die Angabe $\mathscr{L} = \{\varnothing\}$ falsch, denn dies würde bedeuten, dass die leere Menge $\varnothing$ selbst ein Element der Lösungsmenge ist. Als Elemente der Lösungsmenge kommen aber nur Zahlenpaare in Frage, und nicht Mengen. Schließlich ergibt es keinen Sinn, die leere Menge $\varnothing$ in ein Gleichungssystem einzusetzen.

Als weiteres Beispiel für ein nicht eindeutig lösbares lineares Gleichungssystem betrachten wir das System bestehend aus den Gleichungen $2x - 3y + 5z = 1$, $x + y - z = 2$. Wieder wenden wir das Eliminationsverfahren an.

$$\begin{array}{rcrcrcr} 2x & - & 3y & + & 5z & = & 1 \\ x & + & y & - & z & = & 2 \end{array} \quad\Leftrightarrow\quad \begin{array}{rcrcrcrl} x & + & y & - & z & = & 2 & \\ 2x & - & 3y & + & 5z & = & 1 & |+(\mathbf{-2})\cdot\mathbf{I} \end{array} \quad\Leftrightarrow$$

$$\begin{array}{rcrcrcrl} x & + & y & - & z & = & 2 & \\ & & -5y & + & 7z & = & -3 & |\cdot(-\frac{1}{5}) \end{array} \quad\Leftrightarrow\quad \begin{array}{rcrcrcrl} x & + & y & - & z & = & 2 & |+(\mathbf{-1})\cdot\mathbf{II} \\ & & y & - & \frac{7}{5}z & = & \frac{3}{5} & \end{array} \quad\Leftrightarrow$$

$$\begin{array}{rcrcrcr} x & & & + & \frac{2}{5}z & = & \frac{7}{5} \\ & & y & - & \frac{7}{5}z & = & \frac{3}{5} \end{array} \quad\Leftrightarrow\quad \begin{array}{rcl} x & = & -\frac{2}{5}z + \frac{7}{5} \\ y & = & \frac{7}{5}z + \frac{3}{5} \end{array}$$

Wir sehen, dass im letzten System der Wert der Variable $x$ durch keine Gleichung festgelegt ist; er ist „frei wählbar“. Wir erhalten diesmal eine unendliche Lösungsmenge der Form

$$\mathscr{L} \quad = \quad \{(-\tfrac{2}{5}z + \tfrac{7}{5}, \tfrac{7}{5}z + \tfrac{3}{5}, z) \mid z \in \mathbb{R}\}.$$

Setzt man für $z$ konkrete Werte ein, so erhält man jedesmal ein anderes Lösungstupel des ursprünglichen linearen Gleichungssystems, für $z = 0$ beispielsweise das Tupel $(\frac{7}{5}, \frac{3}{5}, 0)$, und für $z = 1$ das Tupel $(1, 2, 1)$. Häufig wird eine Lösungsmenge wie diese fälschlicherweise in der Form

$$\mathscr{L} \quad = \quad \{(-\tfrac{2}{5}z + \tfrac{7}{5}, \tfrac{7}{5}z + \tfrac{3}{5})\} \quad , \quad z \in \mathbb{R} \qquad (*)$$

angegeben. Warum ist diese Angabe definitiv falsch? Zunächst einmal ist (*) gar keine klar formulierte mathematische Aussage, weil der Ausdruck „$z$“ weder zuvor definiert noch mit einem Quantor ($\forall$ oder $\exists$) versehen wurde. Letzteres könnte man nachholen, indem man statt desssen

$$\forall z \in \mathbb{R} : \mathscr{L} = \{(-\tfrac{2}{5}z + \tfrac{7}{5}, \tfrac{7}{5}z + \tfrac{3}{5}, z)\} \qquad \text{oder} \qquad \exists z \in \mathbb{R} : \mathscr{L} = \{(-\tfrac{2}{5}z + \tfrac{7}{5}, \tfrac{7}{5}z + \tfrac{3}{5})\}$$

schreibt. Dies sind jetzt klar formulierte Aussagen, aber beide treffen nicht zu. Die erste Aussage würde bedeuten, dass die Gleichung $\mathscr{L} = \{(-\frac{2}{5}z + \frac{7}{5}, \frac{7}{5}z + \frac{3}{5})\}$ für *jedes* $z \in \mathbb{R}$ erfüllt ist. Dann würde beispielsweise $\mathscr{L} = \{(\frac{7}{5}, \frac{3}{5}, 0)\}$ und zugleich $\mathscr{L} = \{(1, 2, 2)\}$ gelten, denn laut Annahme gilt die Gleichung sowohl für $z = 0$ als auch für $z = 1$. Das ist natürlich absurd; es kann nach Definition nur eine Lösungsmenge des linearen Gleichungssystem geben. Aber auch durch die zweite Aussage wird die Lösungsmenge nicht korrekt angebenen. Denn sie bedeutet, dass $\mathscr{L} = \{(-\frac{2}{5}z + \frac{7}{5}, \frac{7}{5}z + \frac{3}{5}, z)\}$ zumindest für ein $z \in \mathbb{R}$ erfüllt ist, beispielsweise für $z = 0$ oder $z = -\frac{3}{4}$. In jedem Fall wäre die Lösungsmenge einelementig, würde also nur ein Zahlentupel enthalten. Wie wir aber oben festgestellt haben, besteht $\mathscr{L}$ aus unendlich vielen Zahlentupeln.

Nach diesen kurzen Wiederholungen kommen wir nun zu der Frage, was uns im kommenden Semester erwartet.

- Zunächst einmal werden wir lineare Gleichungssysteme und ihre Lösungsmengen systematisch untersuchen, und zwar nicht nur über den reellen Zahlen, sondern über einem beliebigen Grundkörper $K$. Dabei werden wir beispielsweise feststellen, dass über $K = \mathbb{R}$ (und ebenso über jedem anderen unendlichen Körper) nur die drei schon beobachteten Fälle eintreten können: Die Lösungsmenge eines linearen Gleichungssystems ist entweder leer, oder sie enthält genau ein Lösungstupel, oder sie besteht aus unendlich vielen Lösungstupeln.
- Um linearen Gleichungssysteme effizient zu lösen, aber auch für andere Aufgaben, werden wir die ***Vektor- und Matrixschreibweise*** einführen. Dabei wird sich herausstellen, dass man in gewissen Grenzen mit Matrizen wie mit Zahlen rechnen kann: Man kann sie addieren, subtrahieren und multiplizieren, einige Matrixen besitzen sogar einen Kehrwert. Es gibt allerdings auch Unterschiede; beispielsweise ist bei Matrizen die Gleichung $AB = BA$ in der Regel nicht erfüllt. (Wir werden auch sehen, dass $A$ und $B$ ein bestimmtes Format haben müssen, damit die Produkte $AB$ und $BA$ überhaupt definiert sind.)
- Um Lösungsmengen von linearen Gleichungssystem qualitativ beschreiben zu können, werden wir eine ganze Reihe wichtiger algebraischer Grundbegriffe kennenlernen, zum Beispiel den Begriff des Vektorraums, affine Unterräume, das Konzept der linearen Abhängkeit, lineare Abbildungen und den Dimensionsbegriff. Wir werden sehen, dass viele dieser Begriffe eine anschaulich-geometrische Interpretation besitzen; beispielsweise kann mit der ***Determinante*** das Volumen eines Paralletops oder die Orientierung eines Koordinatensystems (Rechts- und Linkshändigkeit) beschrieben werden. Viele dieser Begriffe werden auch in weiterführenden Vorlesungen verwendet.
- Neben dem Lösen von linearen Gleichungssystemen lernen wir viele weitere Rechentechniken kennen, beispielsweise Basisbestimmung, Koordinatentransformationen, Determinanten- und Eigenwertberechnung.

Viele Konzepte, die wir in der Linearen Algebra kennenlernen, haben in anderen Teilgebieten der Mathematik wichtige Anwendungen.

- Die Geometrie wurde oben bereits angesprochen; viele geometrische Eigenschaften von Strukturen lassen sich mit Hilfsmitteln aus der linearen Algebra untersuchen. Beispielsweise sind Symmetrieoperationen wie Drehungen, Translationen oder Spiegelungen nichts anderes als lineare Abbildungen oder affine Transformationen. Bilineare Abbildungen benötigt man für Winkel-, Flächen- oder Abstandsberechnungen. Auch bei der Klassifikation geometrischer Strukturen, zum Beispiel von Kegelschnitten oder sog. qudratischen Flächen (Quadriken) im $\mathbb{R}^3$, kommt die Lineare Algebra zum Einsatz.
- Auch für die mehrdimensionle Analysis, das Thema des nächsten Semesters, spielt die Lineare Algebra eine wichtige Rolle. Beispielsweise werden wir sehen, dass die Ableitung einer mehrdimensionalen Funktion, wie zum Beispiel
$$f : \mathbb{R}^3 \to \mathbb{R}^2 \quad , \quad (x, y, z) \mapsto (3x - 5y + z, 2x + 3z),$$

an jedem einzelnen Punkt des $\mathbb{R}^3$ keine reelle Zahl, sondern eine Matrix ist (genauer gesagt, eine $2 \times 3$-Matrix, mit sechs Einträgen). Auch für die Beschreibung der Lösungsfunktionen von gewöhnlichen Differenzialgleichungen ist die Lineare Algebra essentiell.

- Viele Verfahren der Numerik, zum Beispiel die Interpolation, basieren auf Methoden der Linearen Algebra.
- In der Algebra-Vorlesung werden wir sehen, dass der Vektorraumbegriff der Linearen Algebra beim Verständnis von sog. algebraischen Körpererweiterungen eine wichtige Rolle spielt.
- In der Funktionanalysis betrachtet man Funktionen systematisch als Lösung von Gleichungen, beispielsweise Differential- oder Integralgleichungen. Zum großen Teil geht es dabei im Kern um das Studium unendlichdimensionaler Vektorräume.

Die Liste ließe sich noch um viele weitere Punkte ergänzen. Zum Schluss sollte noch erwähnt werden, dass auch außerhalb der Mathematik, vor allem in der Naturwissenschaften, die Methoden der Linearen Algebra zu einem unverzichtbaren Hilfsmittel geworden sind.

- Am stärksten von der Linearen Algebra durchsetzt ist natürlich die Physik. Mechanische und elektromagnetische Wellen sind Linearkombinationen sogenannter ***ebener Wellen*** und bilden damit einen Vektorraum. Ausschlaggebend hierfür ist die Linearität der ***Wellengleichungen***, die bei der Bestimmung und der Klassifikation der Lösungen eine zentrale Rolle spielt. Der Trägheitstensor, mit dem in der Mechanik die Rotation von starren Körpern beschrieben wird, ist nichts anderes als ein linearer Endomorphismus; seine Eigenvektoren werden die ***Trägheitsachsen*** des Körpers genannt. Sie liefern die Richtungen, um die eine stabile Rotation des Körpers möglich ist.

  Sowohl Wellenphänomene als auch die Eigenwerttheorie sind auch essentiell für die ***Quantenmechanik***. Beispielsweise wird jede Messgröße (Observable) eines quantenmechanischen Systems durch einen linearen Endomorphismus beschrieben, dessen Eigenwerte genau die Werte sind, die bei einer Messung am System potentiell, also mit einer gewissen Wahrscheinlichkeit, auftreten. Die Berechnung von Eigenwerten und -vektoren werden wir am Ende dieses Semesters kennenlernen.

  Die Gesetze ***Speziellen Relativitätstheorie*** lassen sich besonders einfach und elegant mit Hilfe einer indefiniten Bilinearform (der sogenannten Minkowski-Metrik) formulieren. Noch stärker gehen die Methoden der Linearen Algebra (neben der Analysis und der Differentialgeometrie) in die ***Allgemeine Relativitätstheorie*** ein, hier in Form von ***Tensorfeldern*** auf Mannigfaltigkeiten. Beispielsweise ist die berühmte Einsteinsche Feldgleichung eine Gleichung zwischen zwei Tensorfeldern, dem Einsteinschen Krümmungstensor und dem Energie-Impuls-Tensor.

- Lineare Gleichungssysteme treten auch in anderen Fachgebieten in großer Zahl auf. Beispielsweise benötigt man sie in der Chemie zur Berechnung von Gleichgewichtsreaktionen. In der Elektrotechnik bestimmt man Ströme und Spannungen in Gleich- und in Wechselstromkreisen durch Lösung solcher Systeme. (Bei Gleichstromkreisen handelt es sich um Systeme über $\mathbb{R}$, bei Wechselstromkreisen um Systeme über dem Körper $\mathbb{C}$ der komplexen Zahlen. Dort werden beispielsweise Spulen und Kondensatoren als rein „imaginäre" Widerstände behandelt, und auch Strom und Spannung werden zu komplexen Zahlen, weil man neben der Amplitude auch die Phase berücksichtigt.) Auch in den Wirtschafts- und Sozialwissenschaften hat man es mit linearen Gleichungsystemen zu tun, beispielsweise bei Optimierungsproblemen oder statistischen Berechnungen.

# § 1. Matrizen und Lineare Gleichungssysteme

***Inhaltsübersicht***

In diesem Abschnitt definieren wir lineare Gleichungssysteme (LGS) über einem beliebigen Körper $K$ in einer beliebigen Anzahl $n$ von Variablen und legen fest, was wir unter einer Lösung bzw. der Lösungsmenge eines solchen Systems verstehen. Wir unterscheiden homogene und inhomogene LGS und untersuchen den Zusammenhang zwischen den Lösungsmengen solcher Systeme. Zum Schluss führen wir die sog. Matrixschreibweise ein, mit der sich LGS in einer kompakten Form darstellen lassen.

***Wichtige Definitionen und Sätze***

- Linearform auf dem $K^n$ ($n \in \mathbb{N}$, $K$ Körper)
- lineares Gleichungssystem (LGS), homogen bzw. inhomogen
- Lösungsmenge eines LGS
- elementare Umformung eines LGS, Äquivalenzumformung
- $m \times n$-Matrix über einem Körper $K$
- Matrix-Vektor-Produkt
- (erweiterte) Koeffizientenmatrix eines LGS

Im gesamten Text bezeichnet $K$ stets einen beliebigen Körper, solange nichts genaueres festgelegt wird. Mit $K^\times$ bezeichnen wir die Menge der Elemente in $K$ ungleich $0_K$. Für jedes $n \in \mathbb{N}$ ist $0_{K^n}$ das Tupel $(0_K, ..., 0_K) \in K^n$, dessen sämtliche Einträge gleich $0_K$ sind. Ist $x \in K^n$, dann bezeichnen wir mit $x_1, ..., x_n$ die Komponenten von $x$. Es gilt also jeweils $x = (x_1, ..., x_n)$.

Die Elemente in $K^n$ werden auch ***Vektoren*** genannt. Sie lassen sich komponentenweise addieren oder mit einem „Skalar" $\lambda \in K$ multiplizieren. Für beliebige $v, w \in K^n$ definieren wir

$$v + w = (v_1 + w_1, ..., v_n + w_n) \qquad \text{und} \qquad \lambda v = (\lambda v_1, ..., \lambda v_n).$$

**(1.1) Definition** Sei $n \in \mathbb{N}$. Eine ***Linearform*** auf dem $K^n$ ist eine Abbildung $\phi : K^n \to K$, die in der Form $\phi(x) = \sum_{j=1}^n a_j x_j$ mit geeigneten Koeffizienten $a_1, ..., a_n \in K$ dargestellt werden kann.

Beispielsweise sind $\phi(x_1, x_2) = 3x_1 - 2x_2$ und $\psi(x_1, x_2) = x_1 + x_2$ Linearformen.

**(1.2) Definition** Seien $m, n \in \mathbb{N}$. Ein ***lineares Gleichungssystem*** (LGS) bestehend aus $m$ Gleichungen in $n$ Unbekannten ist ein Paar

$$((\phi_1, ..., \phi_m), b)$$

bestehend aus einem Tupel von $m$ Linearformen $\phi_i$ und einem Vektor $b \in K^m$. Ist $b = 0_{K^m}$, dann spricht man von einem ***homogenen***, sonst von einem ***inhomogenen*** LGS.

Sind beispielsweise $\phi$ und $\psi$ die oben angegebenen Linearformen, dann ist $\left((\phi,\psi),\binom{7}{6}\right)$ eine Kurzschreibweise für das inhomogene lineare Gleichungssystem

$$\begin{array}{rcrcl} 3x_1 & - & 2x_2 & = & 7 \\ x_1 & + & x_2 & = & 6. \end{array}$$

Häufig werden lineare Gleichungssysteme in der folgenden ausgeschriebenen Form dargestellt: Sind die $m$ Linearformen gegeben durch $\phi_i(x) = \sum_{j=1}^n a_{ij}x_j$ für $1 \le i \le m$ und ist $b = (b_1,...,b_m)$, dann schreibt man das LGS $((\phi_1,...,\phi_m),b)$ auch in der Gestalt

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & ... & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & ... & + & a_{2n}x_n & = & b_2 \\ \vdots & & & & & & & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & ... & + & a_{mn}x_n & = & b_m \end{array} \tag{1.1}$$

**(1.3) Definition** Ein Element $v \in K^n$ bezeichnet man als ***Lösung*** des LGS $((\phi_1,...,\phi_m),b)$, wenn die Gleichungen $\phi_i(v) = b_i$ für $1 \le i \le m$ erfüllt sind. Die Teilmenge

$$\mathscr{L} \quad = \quad \{v \in K^n \mid \phi_i(v) = b_i \text{ für } 1 \le i \le m\} \quad \subseteq \quad K^n$$

wird die ***Lösungsmenge*** des LGS genannt.

Ein Element $v = (v_1,...,v_n) \in K^n$ liegt nach Definition genau dann in der $\mathscr{L}$, wenn alle $m$ Gleichungen in (1.1) erfüllt sind, nachdem man $x_j$ für $1 \le j \le n$ durch $v_j$ ersetzt hat. Wir notieren nun einige grundsätzliche Beobachtungen zur Form solcher Lösungsmengen. Zur Vorbereitung halten wir fest

**(1.4) Lemma** Ist $\phi : K^n \to K$ eine Linearform und sind $v,w \in K^n$ und $\lambda \in K$, dann gilt $\phi(v+w) = \phi(v)+\phi(w)$ und $\phi(\lambda v) = \lambda\phi(v)$. (Abbildungen mit dieser Eigenschaft werden wir später als *linearen Abbildungen* bezeichnen.)

***Beweis:*** Beide Gleichungen erhält man durch einfaches Nachrechnen. Nach Definition der Linearformen hat die Abbildung $\phi$ die Form $\phi(x) = \sum_{j=1}^n a_j x_j$ mit $a_1,...,a_n \in K$. Man erhält nun

$$\phi(v+w) \quad = \quad \sum_{j=1}^n a_j(v_j+w_j) \quad = \quad \sum_{j=1}^n a_j v_j + \sum_{j=1}^n a_j w_j \quad = \quad \phi(v)+\phi(w)$$

und ebenso $\phi(\lambda v) = \sum_{j=1}^n a_j(\lambda v_j) = \lambda \sum_{j=1}^n a_j v_j = \lambda\phi(v)$. □

**(1.5) Proposition** Sei $((\phi_1,...,\phi_m),0_{K^m})$ ein homogenes LGS mit Lösungsmenge $\mathscr{L} \subseteq K^n$. Ist $\lambda \in K$ und sind $v,w \in \mathscr{L}$, dann gilt auch $v+w \in \mathscr{L}$ und $\lambda v \in \mathscr{L}$. (Dies ist, wie wir später sehen werden, die charakteristische Eigenschaft eines *Untervektorraums* des $K^n$.)

***Beweis:*** Aus $v,w \in \mathscr{L}$ folgt nach Definition $\phi_i(v) = \phi_i(w) = 0_K$ für $1 \le i \le m$. Nach (1.4) ist damit auch $\phi_i(v+w) = \phi_i(v)+\phi_i(w) = 0_K+0_K = 0_K$ für $1 \le i \le m$ und somit $v+w \in \mathscr{L}$. Ebenso ist $\phi_i(\lambda v) = \lambda\phi_i(v) = \lambda\cdot 0_K = 0_K$ und damit $\lambda v \in \mathscr{L}$. □

**(1.6) Proposition** Sei $((\phi_1,...,\phi_m),b)$ ein beliebiges LGS, $\mathscr{L}$ die Lösungsmenge dieses Systems und $\mathscr{L}^{\mathrm{h}}$ die Lösungsmenge des zugehörigen homogenen LGS gegeben durch $((\phi_1,...,\phi_m),0_{K^m})$. Ist $v \in \mathscr{L}$ ein beliebiges Element, dann gilt $\mathscr{L} = v + \mathscr{L}^{\mathrm{h}}$, also

$$\mathscr{L} \quad = \quad \{v+w \mid w \in \mathscr{L}^{\mathrm{h}}\}.$$

***Beweis:*** Aus $v \in \mathscr{L}$ folgt $\phi_i(v) = b_i$ für $1 \le i \le m$. Beweisen wir nun zunächst die Inklusion „$\supseteq$“. Ist $w \in \mathscr{L}^{\mathrm{h}}$, dann gilt $\phi_i(w) = 0_K$ für $1 \le i \le m$. Wir erhalten $\phi_i(v+w) = \phi_i(v)+\phi_i(w) = b_i + 0_K = b_i$ für $1 \le i \le m$ und somit $v+w \in \mathscr{L}$.

Zum Nachweis von „$\subseteq$“ sei $u \in \mathscr{L}$ vorgegeben. Dann gilt $\phi_i(u) = b_i$ für $1 \le i \le m$. Setzen wir $w = u - v$, dann folgt $\phi_i(w) = \phi_i(u)+\phi_i(-v) = \phi_i(u)-\phi_i(v) = b_i - b_i = 0_K$ für alle $i$. Dies zeigt, dass $w$ in $\mathscr{L}^{\mathrm{h}}$ enthalten ist. Also ist $u$ in der Form $v+w$ mit $w \in \mathscr{L}^{\mathrm{h}}$ darstellbar. □

Um die Lösungsmenge eines LGS konkret auszurechnen, benötigt man Operationen, die zur Vereinfachung des Systems genutzt werden können, ohne die Lösungsmenge zu ändern. Zunächst bemerken wir

**(1.7) Lemma** Seien $\phi,\psi$ Linearformen auf dem $K^n$ und $\lambda \in K$. Dann sind auch die beiden Abbildungen $\phi+\psi$ und $\lambda\phi$ gegeben durch $(\phi+\psi)(x) = \phi(x)+\psi(x)$ und $(\lambda\phi)(x) = \lambda\phi(x)$ Linearformen.

***Beweis:*** Dies sieht man unmittelbar durch Einsetzen. Seien $\phi$ und $\psi$ gegeben durch $\phi(x) = \sum_{j=1}^n a_j x_j$ und $\psi(x) = \sum_{j=1}^n b_j x_j$ mit $a_j, b_j \in K$ für $1 \le j \le n$. Dann gilt für jedes $x \in K^n$ jeweils

$$(\phi+\psi)(x) = \sum_{j=1}^n (a_j+b_j)x_j \qquad \text{und} \qquad (\lambda\phi)(x) = \sum_{j=1}^n (\lambda a_j)x_j.$$

Also sind auch $\phi+\psi$ und $\lambda\phi$ Linearformen. □

**(1.8) Definition** Unter einer ***elementaren Umformung*** eines LGS $((\phi_1,...,\phi_m),b)$ versteht man eine der folgenden Operationen.

($\mathrm{M}_{k,\lambda}$) Ersetzung von $\phi_k$ durch $\lambda\phi_k$ und von $b_k$ durch $\lambda b_k$, für ein $k \in \{1,...,m\}$ und ein $\lambda \in K^\times$

($\mathrm{A}_{k,\ell,\lambda}$) Ersetzung von $\phi_\ell$ durch $\lambda\phi_k + \phi_\ell$ und von $b_\ell$ durch $\lambda b_k + b_\ell$, für $k,\ell \in \{1,...,m\}$ mit $k \neq \ell$ und beliebiges $\lambda \in K$

Als weiteren Umformungstyp betrachtet man häufig noch

($\mathrm{V}_{k,\ell}$) Vertauschung von $\phi_k$ und $\phi_\ell$ sowie von $b_k$ und $b_\ell$, für $k,\ell \in \{1,...,m\}$ mit $k \neq \ell$

Diese Umformungen betrachten wir aber nicht als elementar, weil sie aus den Umformungen vom Typ ($\mathrm{M}_{k,\lambda}$) und ($\mathrm{A}_{k,\ell,\lambda}$) zusammengesetzt werden kann, wie das folgende Schema zeigt.

$$\begin{pmatrix} \phi_k, b_k \\ \phi_\ell, b_\ell \end{pmatrix} \xrightarrow{\mathrm{A}_{k,\ell,1}} \begin{pmatrix} \phi_k, b_k \\ \phi_k+\phi_\ell, b_k+b_\ell \end{pmatrix} \xrightarrow{\mathrm{M}_{k,-1}} \begin{pmatrix} -\phi_k, -b_k \\ \phi_k+\phi_\ell, b_k+b_\ell \end{pmatrix} \xrightarrow{\mathrm{A}_{\ell,k,1}} \begin{pmatrix} \phi_\ell, b_\ell \\ \phi_k+\phi_\ell, b_k+b_\ell \end{pmatrix} \xrightarrow{\mathrm{A}_{k,\ell,-1}} \begin{pmatrix} \phi_\ell, b_\ell \\ \phi_k, b_k \end{pmatrix}$$

**(1.9) Proposition** Sei $((\phi_1,...,\phi_m),b)$ ein LGS und $((\phi'_1,...,\phi'_m),b')$ ein weiteres, das aus dem ersten durch Anwendung einer elementaren Umformung entsteht. Bezeichnen $\mathscr{L},\mathscr{L}' \subseteq K^n$ die Lösungsmengen der beiden Systeme, dann gilt $\mathscr{L} = \mathscr{L}'$.

***Beweis:*** Zunächst betrachten wir den Fall, dass $\mathscr{L}'$ durch Anwendung einer elementaren Umformung vom Typ ($\mathrm{M}_{k,\lambda}$) entsteht. Es gilt dann $\phi'_k = \lambda\phi_k$ und $b'_k = \lambda b_k$ sowie $\phi'_i = \phi_i$ und $b'_i = b_i$ für alle $i \neq k$. Für alle $v \in K^n$ und alle $i \neq k$ gilt dann offenbar $\phi_i(v) = b_i \Leftrightarrow \phi'_i(v) = b'_i$, und ebenso

$$\phi_k(v) = b_k \quad \Leftrightarrow \quad \lambda\phi_k(v) = \lambda b_k \quad \Leftrightarrow \quad \phi'_k(v) = b'_k.$$

Insgesamt erhalten wir damit

$$v \in \mathscr{L} \quad \Leftrightarrow \quad \phi_i(v) = b_i \text{ für } 1 \le i \le m \quad \Leftrightarrow \quad \phi'_i(v) = b'_i \text{ für } 1 \le i \le m \quad \Leftrightarrow \quad v \in \mathscr{L}'$$

und somit $\mathscr{L} = \mathscr{L}'$. Betrachten wir nun den Fall einer Umformung vom Typ ($\mathrm{A}_{k,\ell,\lambda}$). Hier gilt $\phi_i = \phi'_i$ und $b_i = b'_i$ für alle $i \neq \ell$, außerdem $\phi'_\ell = \lambda\phi_k + \phi_\ell$ und $b'_\ell = \lambda b_k + b_\ell$. Für jedes $v \in K^n$ und $i \neq k,\ell$ ist die Äquivalenz $\phi_i(v) = b_i \Leftrightarrow \phi'_i(v) = b'_i$ offensichtlich. Für die Indizes $k,\ell$ gilt

$$\phi_k(v) = b_k \ \wedge \ \phi_\ell(v) = b_\ell \quad \Leftrightarrow \quad \phi_k(v) = b_k \ \wedge \ \lambda\phi_k(v) + \phi_\ell(v) = \lambda b_k + b_\ell \quad \Leftrightarrow \quad \phi'_k(v) = b'_k \wedge \phi'_\ell(v) = b'_\ell.$$

Also ist $\mathscr{L} = \mathscr{L}'$ auch für diesen Umformungstyp erfüllt. □

Eine Umformung, die die Lösungsmenge eines LGS nicht ändert, wird ***Äquivalenzumformung*** genannt. Die Umformungen vom Typ ($\mathrm{M}_{k,\lambda}$), ($\mathrm{A}_{k,\ell,\lambda}$) und ($V_{k,\ell}$) sind also alles Äquivalenzumformungen.

Wir werden nun sehen, wie man ein LGS noch kompakter darstellen kann.

**(1.10) Definition** Seien $m, n \in \mathbb{N}$. Eine ***m×n - Matrix*** über $K$ ist eine Abbildung

$$A : \{1, ..., m\} \times \{1, ..., n\} \longrightarrow K.$$

Dabei nennt man $A(i, j)$ den ***Eintrag*** von $A$ an der Stelle $(i, j)$. Die Menge aller $m \times n$-Matrizen über $K$ wird mit $\mathscr{M}_{m\times n,K}$ bezeichnet. An Stelle von $\mathscr{M}_{n\times n,K}$ schreibt man auch kürzer $\mathscr{M}_{n,K}$. Die Elemente dieser Menge werden als ***quadratische*** Matrizen bezeichnet.

Durch die Gleichung $A = (a_{ij})$ ordnet man dem Eintrag $A(i, j)$ der Matrix $A$ die Bezeichnung $a_{ij}$ zu. Allgemein legen wir folgende Konvention fest: Bezeichnet ein Großbuchstabe wie zum Beispiel $A, B, C$ eine Matrix, dann bezeichnen die indizierten Kleinbuchstaben $a_{ij}$, $b_{ij}$, $c_{ij}$ immer automatisch die Einträge dieser Matrix. Man kann eine Matrix auch auf übersichtliche Weise als rechteckiges Schema der Form

$$A \quad = \quad \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \dots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \qquad \text{darstellen.}$$

Allgemein nennt man $a_{i\bullet} = (a_{i1}, ..., a_{in}) \in K^n$ die ***i-te Zeile*** und $a_{\bullet j} = (a_{1j}, ..., a_{mj}) \in K^m$ die ***j-te Spalte*** von $A$. Im weiteren Verlauf wird es sich als praktisch erweisen, auch die Elemente aus $K^n$ als spezielle Matrizen mit nur einer Spalte anzusehen. Wir legen deshalb fest, dass von nun an $K^n = \mathscr{M}_{n\times 1,K}$ für alle $n \in \mathbb{N}$ gilt.

Zur Beschreibung der Einträge verwendet man häufig als hilfreiche Abkürzung das sogenannte ***Kronecker-Delta***. Für jeden Körper $K$ und beliebige $m, n \in \mathbb{N}$ definiert man

$$\delta_{mn} \quad = \quad \delta_{mn,K} \quad = \quad \begin{cases} 1_K & \text{falls} \quad m = n \\ 0_K & \text{falls} \quad m \neq n. \end{cases}$$

Falls aus dem Kontext geschlossen werden kann, welcher Körper gemeint ist, wird der Index $K$ auch oft weggelassen. Die folgenden konkreten Beispiele für Matrizen werden uns in den Anwendungen immer wieder begegnen.

(i) die ***Nullmatrix*** $\mathbf{0} = \mathbf{0}^{(m\times n)}$ in $\mathscr{M}_{m\times n,K}$, deren sämtliche Einträge gleich $0_K$ sind
(Wir bezeichnen mit $\mathbf{0}^{(n)} = \mathbf{0}^{(n\times n)}$ die quadratische Nullmatrix.)

(ii) die ***Einheitsmatrix*** $E = E^{(n)}$ in $\mathscr{M}_{n,K}$ mit den Einträgen $\delta_{ij}$ für $1 \leq i \leq m$ und $1 \leq j \leq n$
(Die Einheitsmatrix ist also immer quadratisch.)

(iii) die ***Basismatrizen*** $B_{k\ell} = B_{k\ell}^{(m\times n)}$ mit den Einträgen $b_{ij} = \delta_{ik}\delta_{j\ell}$ für $1 \leq i \leq m$ und $1 \leq j \leq n$

Beispielsweise ist

$$\mathbf{0}^{(2\times 3)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad , \quad E^{(3)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{und} \quad B_{12}^{(3\times 2)} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

**(1.11) Definition** Seien $A \in \mathscr{M}_{m\times n,K}$ und $v \in K^n$. Dann bezeichnet man den Vektor $w \in K^m$ mit den Komponenten $w_i$ gegeben durch

$$w_i \quad = \quad \sum_{j=1}^{n} a_{ij} v_j \qquad \text{für} \quad 1 \le i \le m$$

als das ***Matrix-Vektor-Produkt*** $Av$ von $A$ und $v$.

Wir geben ein konkretes Beispiel für die Berechnung eines Matrix-Vektor-Produkts an.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 3 \\ -1 \\ 7 \end{pmatrix} \quad = \quad \begin{pmatrix} 1\cdot 3 + 2\cdot(-1) + 3\cdot 7 \\ 4\cdot 3 + 5\cdot(-1) + 6\cdot 7 \\ 7\cdot 3 + 8\cdot(-1) + 9\cdot 7 \end{pmatrix} \quad = \quad \begin{pmatrix} 22 \\ 49 \\ 76 \end{pmatrix}.$$

**(1.12) Proposition** Das Matrix-Vektorprodukt erfüllt die Rechenregeln $A(v+w) = Av + Aw$ und $A(\lambda v) = \lambda(Av)$ für beliebige Matrixen $A \in \mathscr{M}_{m\times n,K}$ und Vektoren $v, w \in K^n$.

***Beweis:*** Wir rechnen beide Gleichungen komponentenweise nach. Sei $A = (a_{ij})$, $v' = Av$, $w' = Aw$, $u = A(v+w)$ und $z = A(\lambda v)$. Zu zeigen ist $u = v' + w'$ und $z = \lambda v'$. Nach Definition des Matrix-Vektor-Produkts gelten die Gleichungen

$$v_i' = \sum_{j=1}^{n} a_{ij} v_j \quad , \quad w_i' = \sum_{j=1}^{n} a_{ij} w_j \quad , \quad u_i = \sum_{j=1}^{n} a_{ij}(v_j + w_j) \quad , \quad z_i = \sum_{j=1}^{n} a_{ij}(\lambda v_j).$$

für $1 \le i \le m$. Daraus folgt

$$u_i \quad = \quad \sum_{j=1}^{n} a_{ij}(v_j + w_j) \quad = \quad \sum_{j=1}^{n} a_{ij} v_j + \sum_{j=1}^{n} a_{ij} w_j \quad = \quad v_i' + w_i'$$

und ebenso $z_i = \sum_{j=1}^{n} a_{ij}(\lambda v_j) = \lambda \sum_{j=1}^{n} a_{ij} v_j = \lambda v_i'$ für $1 \le i \le m$, also $u = v' + w'$ und $z = \lambda v'$. □

Jedem Paar $(A,b)$ bestehend aus einer Matrix $A \in \mathscr{M}_{m\times n,K}$ und einem Vektor $b \in K^m$ kann ein LGS der Form $((\phi_1,...,\phi_m),b)$ zugeordnet werden, indem man die Linearformen $\phi_i$ durch

$$\phi_i(x) \quad = \quad \sum_{j=1}^{n} a_{ij}x_j \qquad \text{für} \quad 1 \le i \le m \quad \text{definiert.}$$

Die Lösungsmenge des LGS ist dann durch $\mathscr{L} = \{ \ v \in K^n \mid Av = b \ \}$ gegeben. Man bezeichnet $A$ als die ***Koeffizientenmatrix*** des LGS. Die Matrix $\tilde{A} = (A \ b) \in \mathscr{M}_{m\times(n+1),K}$, deren erste $n$ Spalten mit denen von $A$ und deren letzte Spalte mit $b$ übereinstimmt, nennt man die ***erweiterte*** Koeffizientenmatrix des Systems. Beispielsweise ist die Koeffizientenmatrix $A$ bzw. die erweiterte Koeffizientenmatrix $\tilde{A}$ des LGS $x_1 - 2x_2 + 5x_3 = 7$, $2x_1 - x_3 = 8$ gegeben durch

$$A = \begin{pmatrix} 1 & -2 & 5 \\ 2 & 0 & -1 \end{pmatrix} \qquad \text{bzw.} \qquad \tilde{A} = \begin{pmatrix} 1 & -2 & 5 & 7 \\ 2 & 0 & -1 & 8 \end{pmatrix}.$$

Die Aussagen in Proposition (1.5) und (1.6) ergeben sich nun auch unmittelbar aus den Rechenregeln für das Matrix-Vektor-Produkt: Ist beispielsweise $\mathscr{L} \subseteq K^n$ die Lösungsmenge eines homogenen LGS gegeben durch $(A, 0_{K^m})$, und sind $v,w \in \mathscr{L}$ vorgegeben, dann gilt $Av = 0_{K^m}$ und $Aw = 0_{K^m}$. Es folgt $A(v+w) = Av + Aw = 0_{K^m} + 0_{K^m} = 0_{K^m}$ und somit $v + w \in \mathscr{L}$. Ebenso gilt $A(\lambda v) = \lambda Av = \lambda \cdot 0_{K^m} = 0_{K^m}$ und somit $\lambda v \in \mathscr{L}$, für jedes $\lambda \in K$. Der Beweis von Proposition (1.6) in ähnlicher Form sei dem Leser als (wichtige) Übungsaufgabe überlassen.

Jede elementare Umformung des LGS entspricht einer Umformung der Matrix $A$ bei gleichzeitiger Umformung des Vektors $b$. Beispielsweise wird in $(\mathrm{M}_{k,\lambda})$ die $k$-te Zeile von $A$ und die $k$-te Komponente von $b$ mit dem Wert $\lambda$ multipliziert. Bei $(\mathrm{A}_{k,\ell,\lambda})$ addiert man das $\lambda$-fache der $k$-ten Zeile von $A$ zur $\ell$-ten Zeile von $A$ und ebenso das $\lambda$-fache der Komponente $b_k$ des Vektors $b$ zur Komponente $b_\ell$. Im nächsten Kapitel werden wir sehen, wie sich diese Operationen durch reine Matrizenrechnung bewerkstelligen lassen. Dies wiederum wird bei der Beschreibung und der Analyse des Gaußschen Lösungsverfahrens eine wichtige Rolle spielen.

# § 2. Das Gaußsche Eliminationsverfahren

***Inhaltsübersicht***

In diesem Kapitel behandeln wir ein allgemeines Verfahren, dass die Bestimmung der Lösungsmenge von beliebig großen LGS ermöglichst. Dabei gehen wir davon aus, dass das LGS durch seine erweiterte Koeffizientenmatrix (siehe §1) gegeben ist. Diese Matrix bringt man durch eine fest vorgegebene Folge von Umformungsschritten auf eine sog. normierte Zeilenstufenform. Die Lösungsmenge kann dann auf einfache Weise von der Matrix abgelesen werden, egal ob diese leer oder einelementig ist, oder aus mehreren Elementen besteht. Neben dem Gaußschen Eliminationsverfahren beschäftigen wir uns mit Rechenoperationen für Matrizen und untersuchen, wie diese mit den elementaren Umformungen eines LGS aus dem letzten Kapitel zusammenhängen.

**(2.1) Definition** Eine Matrix $A \in \mathscr{M}_{m\times n,K}$ befindet sich in ***Zeilenstufenform*** (kurz ZSF), wenn $A = \mathbf{0}^{(m\times n)}$ gilt oder folgende Bedingung erfüllt ist: Es gibt ein $r \in \{1,...,m\}$ und $j_1,...,j_r \in \{1,...,n\}$ mit $j_1 < j_2 < ... < j_r$, so dass

(i) $a_{ij_i} \neq 0_K$ für $1 \le i \le r$ und

(ii) $a_{ij} = 0_K$ für $j < j_i$ oder $i > r$

erfüllt ist. Man nennt $r$ den ***Zeilenrang*** einer solchen Matrix. Das Tupel $(r, j_1,..., j_r)$ bezeichnen wir insgesamt als die ***Kennzahlen*** der ZSF.

Die Positionen $(i, j_i)$ mit $1 \le i \le r$ in der Matrix werden ***Zeilenköpfe*** genannt. Die Bedingung (i) besagt, dass die Einträge in den Zeilenköpfen ungleich Null sind. Nach Bedingung (ii) befinden sich links von den Zeilenköpfen nur Nulleinträge; in den „kopflosen“ Zeilen sind alle Einträge gleich Null. Der Zeilenrang kann offenbar nie größer als $\min\{m,n\}$ werden.

**(2.2) Definition** Eine Matrix $A \in \mathscr{M}_{m\times n,K}$ befindet sich in ***normierter*** ZSF, wenn $A = \mathbf{0}^{(m\times n)}$ gilt oder wenn sie in ZSF mit den Kennzahlen $(r, j_1,..., j_r)$ vorliegt und außerdem die folgenden Bedingungen erfüllt sind: Es gilt $a_{ij_i} = 1_K$ für $1 \le i \le r$ und $a_{kj_i} = 0_K$ für $1 \le i \le r$ und $1 \le k < i$.

Bei der normierten ZSF kommen also folgende Bedingungen hinzu: Die Einträge in den Zeilenköpfen sind gleich $1_K$, und oberhalb der Zeilenköpfe befinden sich nur Nulleinträge. Bei einer Matrix $A$ in normierter ZSF gilt also insgesamt $a_{ij} = 0$ in jedem der drei Fälle

(1) $i > r$ (2) $i \le r$ und $j < j_i$ (3) $j = j_k$ für ein $k \in \{1,...,r\} \setminus \{i\}$;

in Worten, die Einträge der Matrix sind Null (1) unterhalb der $r$-ten Zeile, (2) links von den Spaltenköpfen und (3) in jeder Spalte, in der sich ein Spaltenkopf befindet, sind alle anderen Einträge gleich Null. Abgesehen davon können aber durchaus noch weitere Einträge von $A$ gleich Null sein. Wir bemerken außerdem, dass eine Matrix $A \in \mathscr{M}_{m\times n,K}$ in normierter ZSF mit Zeilenrang $r = n$ in den oberen $n$ Zeilen mit der Einheitsmatrix $E^{(n)}$ übereinstimmt, denn in diesem Fall muss $j_k = k$ für $1 \le k \le n$ gelten.

Wir geben einige konkrete Beispiele für Matrizen in Zeilenstufenform an.

$$A = \begin{pmatrix} 0 & 2 & 3 & 4 & 0 & 6 \\ 0 & 0 & 3 & 0 & 8 & 1 \\ 0 & 0 & 0 & 2 & 2 & 9 \\ 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Diese Matrix liegt in Zeilenstufenform vor, mit zugehörigen Kennzahlen $r = 4$, $j_1 = 2$, $j_2 = 3$, $j_3 = 4$ und $j_4 = 6$. Es besteht aber keine normierte Zeilenstufenform, denn beispielsweise sind die Einträge in den Zeilenköpfen $(1,2)$, $(2,3)$, $(3,4)$ und $(4,6)$ ungleich 1. Außerdem gibt es Einträge ungleich Null oberhalb der Zeilenköpfe.

$$B = \begin{pmatrix} 1 & 0 & 7 & 0 & 5 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Dies ist eine Matrix in normierter Zeilenstufenform, mit den Kennzahlen $r = 3$, $j_1 = 1$, $j_2 = 2$ und $j_3 = 4$. Auch die Einheitsmatrix

$$E^{(4)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

liegt in normierter Zeilenstufenform vor. Die Kennzahlen lauten $r = 4$ und $j_i = i$ für $1 \le i \le 4$.

**(2.3) Definition** Für $1 \le \ell \le n$ bezeichnet $e_\ell \in K^n$ jeweils den ***$\ell$-ten Einheitsvektor*** mit den Einträgen $e_{\ell j} = \delta_{\ell j}$.

In $K^3$ sind die drei Einheitsvektoren beispielsweise gegeben durch

$$e_1 = (1_K, 0_K, 0_K) \quad , \quad e_2 = (0_K, 1_K, 0_K) \quad \text{und} \quad e_3 = (0_K, 0_K, 1_K).$$

Sei nun $A$ eine Matrix in normierter ZSF mit Kennzahlen $(r, j_1, ..., j_r)$. Unser erstes Ziel in diesem Abschnitt besteht darin, die Lösungsmenge $\mathscr{L}^{\mathrm{h}}$ des homogenen LGS mit Koeffizientenmatrix $A$, gegeben durch $Ax = 0_{K^m}$, explizit zu beschreiben. Dazu definieren wir

$$S = \{1, ..., n\} \setminus \{j_1, ..., j_r\}$$

und definieren für jede Zahl $\ell \in S$ einen Vektor $b_\ell \in K^m$ durch

$$b_\ell \quad = \quad e_\ell - \sum_{k=1}^{r} a_{k\ell} e_{j_k}.$$

Der Vektor $b_\ell$ entsteht also aus dem Nullvektor dadurch, dass man die $\ell$-te Komponente auf $1_K$ setzt und die Einträge $-a_{1\ell}, ..., -a_{r\ell}$ der $\ell$-ten Spalte auf die Positionen $j_1, ..., j_r$ des Vektors verteilt. Es gilt also

$$b_{\ell j_k} = -a_{k\ell} \text{ für } 1 \le k \le r \quad \text{und} \quad b_{\ell j} = \delta_{\ell j} \text{ für alle } \ell \in S.$$

Mit Hilfe dieser Vektoren lässt sich nun die Lösungsmenge $\mathscr{L}^{\mathrm{h}}$ folgendermaßen darstellen.

**(2.4) Satz** Sei $\mathscr{L}^{\mathrm{h}} \subseteq K^n$ die Lösungsmenge eines homogenen LGS mit Koeffizientenmatrix $A$, und seien $S$ und die Vektoren $b_\ell$ für $\ell \in S$ definiert wie oben.

(i) Im Fall $S = \varnothing$ gilt $\mathscr{L}^{\mathrm{h}} = \{ 0_{K^n} \}$.

(ii) Ist $S$ nichtleer, dann ist die Lösungsmenge gegeben durch

$$\mathscr{L}^{\mathrm{h}} \quad = \quad \left\{ \sum_{\ell \in S} \lambda_\ell b_\ell \;\middle|\; \lambda_\ell \in K \ \forall \ \ell \in S \right\}.$$

***Beweis:*** zu (i) Unter dieser Vorausssetzung gilt $\{j_1, ..., j_r\} = \{1, ..., n\}$, woraus wiederum $r = n$ und somit $m \ge n$ folgt. Wie oben ausgeführt, stimmen bei Zeilenrang $n$ die ersten $n$ Zeilen von $A$ mit der Einheitsmatrix $E^{(n)}$ überein. Es gilt also $a_{ij} = \delta_{ij}$ für $1 \le i, j \le n$ und $a_{ij} = 0_K$ falls $i > n$. Wir erinnern außerdem daran, dass nach Definition $\mathscr{L}^{\mathrm{h}} = \{v \in K^n \mid Av = 0_{K^m}\}$ gilt. Für jeden Vektor $v \in K^n$ erhalten wir die Äquivalenz

$$v \in \mathscr{L}^{\mathrm{h}} \quad \Leftrightarrow \quad Av = 0_{K^m} \quad \Leftrightarrow \quad (Av)_k = 0 \text{ für } 1 \le k \le m \quad \Leftrightarrow \quad \sum_{j=1}^{n} a_{kj} v_j = 0_K \text{ für } 1 \le k \le m$$

$$\Leftrightarrow \quad \sum_{j=1}^{n} \delta_{kj} v_j = 0_K \text{ für } 1 \le k \le n \quad \Leftrightarrow \quad v_k = 0_K \text{ für } 1 \le k \le n \quad \Leftrightarrow \quad v = 0_{K^n}.$$

zu (ii) Zunächst zeigen wir, dass $\mathscr{L}^{\mathrm{h}}$ genau aus den Vektoren besteht, die genau $r$ Gleichungen erfüllen, welche durch die Matrix $A$ vorgegeben sind. Weil (1) die Einträge $a_{ij}$ von $A$ unterhalb der $r$-ten Zeile gleich Null sind und es (3) für $1 \le k \le r$ in der $j_k$-ten Spalte außer $a_{kj_k} = 1_K$ keinen Eintrag ungleich Null gibt, gilt für alle $v \in K^n$ die Äquivalenz

$$v \in \mathscr{L}^{\mathrm{h}} \quad \Leftrightarrow \quad \sum_{j=1}^{n} a_{kj} v_j = 0_K \text{ für } 1 \le k \le m \quad \overset{(1)}{\Leftrightarrow} \quad \sum_{j=1}^{n} a_{kj} v_j = 0_K \text{ für } 1 \le k \le r$$

$$\overset{(3)}{\Leftrightarrow} \quad \sum_{j \in S \cup \{j_k\}} a_{kj} v_j = 0_K \quad \Leftrightarrow \quad v_{j_k} = -\sum_{j \in S} a_{kj} v_j \text{ für } 1 \le k \le r.$$

Für die oben konstruierten Vektoren $b_\ell$ mit $\ell \in S$ muss also die folgende Gleichung überprüft werden.

$$\left\{ v \in K^n \;\middle|\; v_{j_k} = -\sum_{\ell \in S} a_{k\ell} v_\ell \text{ für } 1 \le k \le r \right\} \quad = \quad \left\{ \sum_{\ell \in S} \lambda_\ell b_\ell \;\middle|\; \lambda_\ell \in K \;\forall\; \ell \in S \right\}$$

„$\supseteq$" Sei $v = \sum_{\ell \in S} \lambda_\ell b_\ell$ mit $\lambda_\ell \in K$ für alle $\ell \in S$. Dann sind die Komponenten von $v$ gegeben durch $v_{j_k} = \sum_{\ell \in S} \lambda_\ell b_{\ell j_k} = \sum_{\ell \in S} \lambda_\ell (-a_{k\ell})$ für $1 \le k \le r$ und $v_j = \sum_{\ell \in S} \lambda_j \delta_{\ell j} = \lambda_j$ für $j \in S$. Wir erhalten für $1 \le k \le r$ also

$$v_{j_k} \quad = \quad \sum_{\ell \in S} (-\lambda_\ell) a_{k\ell} \quad = \quad -\sum_{\ell \in S} a_{k\ell} v_\ell \quad ,$$

und damit ist $v$ in der Menge auf der linken Seite enthalten.

„$\subseteq$" Sei $v \in K^n$ ein Vektor mit $v_{j_k} = -\sum_{\ell \in S} a_{k\ell} v_\ell$ für $1 \le k \le r$. Wir definieren $\lambda_\ell = v_\ell$ für alle $\ell \in S$ und setzen $w = \sum_{\ell \in S} \lambda_\ell b_\ell$. Nun rechnen wir nach, dass die Vektoren $v$ und $w$ übereinstimmen. Für $1 \le k \le r$ gilt

$$w_{j_k} \quad = \quad \sum_{\ell \in S} \lambda_\ell b_{\ell j_k} \quad = \quad \sum_{\ell \in S} \lambda_\ell (-a_{k\ell}) \quad = \quad -\sum_{\ell \in S} a_{k\ell} v_\ell \quad = \quad v_{j_k}.$$

Für die Komponenten $j \in S$ erhalten wir ebenso

$$w_j \quad = \quad \sum_{\ell \in S} \lambda_\ell b_{\ell j} \quad = \quad \sum_{\ell \in S} \lambda_\ell \delta_{\ell j} \quad = \quad \lambda_j \quad = \quad v_j$$

für alle $j \in S$. Insgesamt gilt also $v = w$, und damit ist $v$ in der Menge auf der rechten Seite enthalten. □

Wir diskutieren eine Reihe von Anwendungsbeispielen für die soeben bewiesene Lösungsformel.

(i) Das homogene lineare Gleichungssystem $x_1 = 0$, $x_2 + 2x_3 = 0$ hat die Koeffizientenmatrix

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

Es handelt sich um eine Matrix in normierter Zeilenstufenform mit den Kennzahlen $r = 2$, $j_1 = 1$, $j_2 = 2$. Es ist $S = \{1,2,3\} \setminus \{j_1, j_2\} = \{3\}$. Die Lösungsmenge ist somit $\mathscr{L}^{\mathrm{h}} = \{\lambda_3 b_3 \mid \lambda_3 \in \mathbb{R}\}$ mit dem Lösungsvektor

$$b_3 \quad = \quad \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}.$$

(ii) Das homogene lineare Gleichungssystem $x_1 = 0$, $x_3 = 0$ hat die Koeffizientenmatrix

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

Die Kennzahlen dieser normierten Zeilenstufenform lauten $r = 2$, $j_1 = 1$, $j_2 = 3$. Es ist $S = \{2\}$, und die Lösungsmenge ist gegeben durch $\mathscr{L}^{\mathrm{h}} = \{\lambda_2 b_2 \mid \lambda_2 \in \mathbb{R}\}$ mit

$$b_2 \quad = \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

(iii) Das homogene lineare Gleichungssystem $x_1 - 4x_3 + 5x_3 = 0$, $x_2 + 2x_3 = 0$, $x_4 = 0$ hat die Koeffizientenmatrix

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

eine normierte ZSF mit den Kennzahlen $r = 3$, $j_1 = 1$, $j_2 = 2$, $j_3 = 4$. Hier ist $S = \{1, ..., 5\} \setminus \{j_1, j_2, j_3\} = \{3, 5\}$. Der Lösungsraum $\mathscr{L}^{\mathrm{h}} = \{\lambda_3 b_3 + \lambda_5 b_5 \mid \lambda_3, \lambda_5 \in K\}$ wird diesmal aufgespannt von den Vektoren

$$b_3 = \begin{pmatrix} 4 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad , \quad b_5 = \begin{pmatrix} -5 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

**(2.5) Satz** Sei also $\tilde{A} = (A\ b) \in \mathscr{M}_{m \times (n+1),K}$ die erweiterte Koeffizientenmatrix eines LGS und $\mathscr{L} \subseteq K^n$ dessen Lösungsmenge. Wir setzen voraus, dass $\tilde{A}$ in normierter ZSF vorliegt, mit den Kennzahlen $r$ und $j_1, ..., j_r$.

(i) Ist $j_r = n + 1$, dann gilt $\mathscr{L} = \varnothing$.

(ii) Sei nun $j_r \le n$. Wir definieren einen Vektor $w \in K^n$ durch $\sum_{k=1}^{r} b_k e_{j_k}$. Dann gilt $w \in \mathscr{L}$.

Der spezielle Lösungsvektor $w$ entsteht also einfach dadurch, dass man die Werte $b_1, ..., b_r$ auf die Positionen $j_1, ..., j_r$ verteilt und die übrigen Komponenten auf Null setzt. Es gilt also $w_{j_k} = b_k$ für $1 \le k \le r$ und $w_\ell = 0$ für alle $\ell \in S$.

***Beweis:*** zu (i) Nehmen wir an, dass $\mathscr{L}$ nichtleer und $w$ ein Element aus $\mathscr{L}$ ist. Dann gilt insbesondere $\sum_{j=1}^{n} a_{rj} w_j = b_r$. Wegen $j_r = n + 1$ gilt aber $a_{rj} = 0_K$ für $1 \le j \le n$ und $b_r = a_{r,n+1} = a_{r,j_r} = 1_K$. Setzen wir dies in die Gleichung ein, so erhalten wir $\sum_{j=1}^{n} 0_K w_j = 1_K$. Der Widerspruch $0_K = 1_K$ zeigt, dass unsere Annahme falsch war.

zu (ii) Zu zeigen ist $\sum_{j=1}^{n} a_{kj} w_j = b_k$ für $1 \le k \le m$. Auf Grund der Eigenschaft (3) der normierten ZSF ist $a_{kj_k} = 1_K$ jeweils der einzige Eintrag ungleich Null in der $j_k$-ten Spalte, für $1 \le k \le r$. Nach Definition des Vektors $w$ erhalten wir für $1 \le k \le r$ somit

$$\sum_{j=1}^{n} a_{kj} w_j \quad = \quad \sum_{i=1}^{r} a_{kj_k} w_{j_k} \quad \overset{(3)}{=} \quad a_{kj_k} w_{j_k} \quad = \quad w_{j_k} \quad = \quad b_k.$$

Für $r < k \le m$ gilt nach Eigenschaft (1) der normierten ZSF (Einträge unterhalb der $r$-ten Zeile gleich Null) sowohl $a_{kj} = 0$ für $1 \le j \le n$ als auch $b_k = 0$, also ebenfalls $\sum_{j=1}^{n} a_{kj} w_j = 0 = b_k$. Insgesamt ist die Gleichung $\sum_{j=1}^{n} a_{kj} w_j = b_k$ also tatsächlich für $1 \le k \le m$ erfüllt. □

Wir demonstrieren die Anwendung der Lösungsformel an einem konkreten Beispiel. Das inhomogene LGS $x_1 - 4x_3 + 5x_5 = -2$, $x_2 + 2x_3 = 7$, $x_4 = 5$ besitzt die erweiterte Koeffizientenmatrix

$$\tilde{A} \quad = \quad \begin{pmatrix} 1 & 0 & -4 & 0 & 5 & -2 \\ 0 & 1 & 2 & 0 & 0 & 7 \\ 0 & 0 & 0 & 1 & 0 & 5 \end{pmatrix}.$$

Es handelt sich um eine normierte Zeilenstufenform mit den Kennzahlen $r = 3$ und $j_1 = 1$, $j_2 = 2$, $j_3 = 4$. Nach Satz (2.5) ist $w = (-2, 7, 0, 5, 0) \in \mathbb{R}^5$ ein Lösungsvektor, was man durch Einsetzen in die Gleichungen des Systems unmittelbar überprüft: Es ist $(-2) - 4 \cdot 0 + 5 \cdot 0 = -2$, $7 + 2 \cdot 0 = 7$ und $5 = 5$.

Um nun ein Lösungsverfahren für beliebige LGS zu erhalten, brauchen wir also nur noch ein Verfahren, mit dem wir beliebige Matrizen in normierte Zeilenstufenform überführen können. Um dieses Problem auf systematische Weise lösen zu können, benötigen wir eine Reihe von Rechenoperationen für Matrizen.

(i) Seien $m, n \in \mathbb{N}$ und $A, B \in \mathscr{M}_{m\times n,K}$ mit Einträgen $A = (a_{ij})$ und $B = (b_{ij})$. Dann nennt man die Matrix $C = (c_{ij})$ mit $c_{ij} = a_{ij} + b_{ij}$ für $1 \le i \le m$ und $1 \le j \le n$ die ***Summe*** von $A$ und $B$. Wir bezeichnen diese Matrix mit $A + B$. Beispielsweise gilt

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

(ii) Sei $A \in \mathscr{M}_{m\times n,K}$ mit $A = (a_{ij})$ und $\lambda \in K$. Dann ist die Matrix $C = (c_{ij})$ mit $c_{ij} = \lambda a_{ij}$ ein ***skalares Vielfaches*** von $A$, das wir mit $\lambda A$ bezeichnen. Beispielsweise ist

$$7 \cdot \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad = \quad \begin{pmatrix} 7 & 14 & 21 \\ 28 & 35 & 42 \end{pmatrix}.$$

(iii) Seien nun $m, n, r \in \mathbb{N}$ und $A \in \mathscr{M}_{m\times n,K}$, $B \in \mathscr{M}_{n\times r,K}$. Dann heißt die Matrix $C \in \mathscr{M}_{m\times r,K}$ mit den Einträgen

$$c_{ij} \quad = \quad \sum_{k=1}^{n} a_{ik} b_{kj}$$

***Produkt*** der Matrizen $A$ und $B$ und wird mit $AB$ bezeichnet. Auch hierzu ein konkretes Beispiel:

$$\begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 3 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} \quad = \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & -1 & -3 \\ 3 & 6 & 9 & 12 \end{pmatrix}$$

Den Eintrag an der Position $(2, 2)$ erhält man zum Beispiel durch Multiplikation der zweiten Zeile der ersten Matrix mit der zweiten Spalte der zweiten Matrix, also durch die Rechung $(-1) \cdot 2 + 1 \cdot 3 = 1$. Der Eintrag an der Position $(3, 3)$ kommt entsprechend durch $3 \cdot 3 + 0 \cdot 2 = 9$ zu Stande.

(iv) Sei $A \in \mathscr{M}_{m\times n,K}$. Die Matrix $B \in \mathscr{M}_{n\times m,K}$ mit den Einträgen $b_{ij} = a_{ji}$ für $1 \le i \le n$ und $1 \le j \le m$ wird die zu $A$ ***transponierte*** Matrix ${}^{\mathrm{t}}A$ genannt. Zum Beispiel ist

$${}^{\mathrm{t}}\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad = \quad \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Die Berechnung des Matrixprodukts ähnelt dem Matrix-Vektor-Produkt aus dem ersten Kapitel. Um den Eintrag $c_{ij}$ der Matrix $C = AB$ an der Position $(i, j)$ zu erhalten, muss die $i$-te Zeile der Matrix $A$ mit der $j$-ten Spalte der Matrix $B$ multipliziert werden. Man beachte, dass das Produkt $AB$ nur gebildet werden kann, wenn die Spaltenzahl von $A$ mit der Zeilenzahl von $B$ übereinstimmt. Die Summe $A + B$ ist nur dann definiert, wenn $A$ und $B$ dasselbe Format, also dieselbe Anzahl Zeilen *und* Spalten besitzen.

**(2.6) Proposition** Seien $A, A' \in \mathscr{M}_{m\times n,K}$, $B, B' \in \mathscr{M}_{n\times r,K}$ und $C \in \mathscr{M}_{r\times s,K}$. Dann gilt

(i) $A(B + B') = AB + AB'$ und $(A + A')B = AB + A'B$

(ii) $A(\lambda B) = (\lambda A)B = \lambda(AB)$

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

(iv) ${}^{\mathrm{t}}(AB) = {}^{\mathrm{t}}B\,{}^{\mathrm{t}}A$

***Beweis:*** zu (i) Den Beweis der ersten Formel überlassen wir dem Leser als Übungsaufgabe und beweisen statt dessen direkt die zweite. Es sei $C = A + A'$, $D = CB$, $F = AB$, $G = A'B$ und $H = F + G$. Dann ist $D = H$ zu zeigen. Es gilt $c_{ij} = a_{ij} + a'_{ij}$ und

$$d_{ij} \quad = \quad \sum_{k=1}^{n} c_{ik} b_{kj} \quad = \quad \sum_{k=1}^{n} (a_{ik} + a'_{ik}) b_{kj} \quad = \quad \sum_{k=1}^{n} a_{ik} b_{kj} + \sum_{k=1}^{n} a'_{ik} b_{kj}.$$

Für die Einträge der Matrizen $F$ und $G$ erhalten wir

$$f_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \qquad \text{und} \qquad g_{ij} = \sum_{k=1}^{n} a'_{ik} b_{kj}.$$

Es folgt

$$h_{ij} \quad = \quad f_{ij} + g_{ij} \quad = \quad \sum_{k=1}^{n} a_{ik} b_{kj} + \sum_{k=1}^{n} a'_{ik} b_{kj} \quad = \quad d_{ij}$$

also insgesamt $D = H$.

zu (ii) Wir definieren $C = \lambda B$, $D = AC$, $F = \lambda A$, $G = FB$, $H = AB$ und $U = \lambda H$. Zu zeigen ist dann $D = G = U$. Nach Definition gilt $c_{ij} = \lambda b_{ij}$ und

$$d_{ij} \quad = \quad \sum_{k=1}^{n} a_{ik} c_{kj} \quad = \quad \sum_{k=1}^{n} \lambda a_{ik} b_{kj}.$$

Andererseits gilt auch $f_{ij} = \lambda a_{ij}$ und $g_{ij} = \sum_{k=1}^{n} f_{ik} b_{kj} = \sum_{k=1}^{n} \lambda a_{ik} b_{kj} = d_{ij}$, womit die Gleichung $D = G$ bewiesen ist. Nun gilt außerdem $h_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$ und

$$u_{ij} \quad = \quad \lambda h_{ij} \quad = \quad \sum_{k=1}^{n} \lambda a_{ik} b_{kj} \quad = \quad d_{ij} \quad ,$$

wodurch auch die Gleichung $U = D$ bewiesen ist.

zu (iii) Wir definieren $D = AB$, $F = DC$, $G = BC$ und $H = AG$. Dann gilt $d_{k\ell} = \sum_{i=1}^n a_{ki} b_{i\ell}$, und für die Matrix $F$ erhalten wir

$$f_{k\ell} \quad = \quad \sum_{i=1}^r d_{ki} c_{i\ell} \quad = \quad \sum_{i=1}^r \sum_{j=1}^n a_{kj} b_{ji} c_{i\ell}.$$

Andererseits hat $G$ die Einträge $g_{k\ell} = \sum_{i=1}^r b_{ki} c_{i\ell}$, und für die Matrix $H$ gilt

$$h_{k\ell} \quad = \quad \sum_{i=1}^n a_{ki} g_{i\ell} \quad = \quad \sum_{i=1}^n \sum_{j=1}^r a_{ki} b_{ij} c_{j\ell} \quad , \quad \text{also insgesamt} \quad F = H.$$

zu (iv) Hier definieren wir die Hilfsmatrizen $C = AB$, $D = {}^{\mathrm{t}}C$, $F = {}^{\mathrm{t}}A$, $G = {}^{\mathrm{t}}B$ und $H = GF$. Dann müssen wir $D = H$ nachrechnen. Es gilt

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj} \qquad \text{und} \qquad d_{ij} = c_{ji} = \sum_{k=1}^n a_{jk} b_{ki}.$$

Wegen $f_{ij} = a_{ji}$ und $g_{ij} = b_{ji}$ gilt außerdem

$$h_{ij} \quad = \quad \sum_{k=1}^n g_{ik} f_{kj} \quad = \quad \sum_{k=1}^n b_{ki} a_{jk} \quad = \quad d_{ij}$$

also $H = D$ wie gewünscht. □

**(2.7) Proposition** Sei $A \in \mathscr{M}_{m\times n,K}$. Dann gilt $E^{(m)}A = AE^{(n)} = A$.

***Beweis:*** Sei $B = E^{(m)}A$. Für $1 \le i \le m$ und $1 \le j \le n$ gilt dann

$$b_{ij} \quad = \quad \sum_{k=1}^m \delta_{ik} a_{kj} \quad = \quad a_{ij}.$$

Damit ist $B = A$ bewiesen. Sei nun $C = AE^{(n)}$. Für $1 \le i \le m$ und $1 \le j \le n$ gilt dann ebenfalls

$$c_{ij} \quad = \quad \sum_{k=1}^n a_{ik} \delta_{kj} \quad = \quad a_{ij} \quad , \qquad \text{also} \quad C = A. \qquad \square$$

Bei Rechnungen mit Matrizen ist es oft günstig, diese in mehrere Bereiche aufzuteilen. Sei $A \in \mathscr{M}_{m\times n,K}$, seien $k_1, k_2, \ell_1, \ell_2$ natürliche Zahlen mit $1 \le k_1 \le k_2 \le m$, $1 \le \ell_1 \le \ell_2 \le n$, und außerdem $r = k_2 - k_1 + 1$, $s = \ell_2 - \ell_1 + 1$. Dann nennt man die Matrix $B \in \mathscr{M}_{r\times s,K}$ mit den Einträgen $b_{ij} = a_{k_1+i-1,\ell_1+j-1}$ für $1 \le i \le r$, $1 \le j \le s$ eine ***Teilmatrix*** von $A$; es handelt sich um einen „rechteckigen Ausschnitt" der Matrix $A$.

Häufig verwendet man die sogenannte ***Blockschreibweise***, um Matrizen darzustellen, die aus bestimmten Teilmatrizen aufgebaut sind. So steht beispielsweise der Ausdruck

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

für die Matrix, deren linker oberer Teil aus den Einträgen von $A$ und entsprechend in den übrigen drei Bereichen aus den Einträgen von $B$, $C$ und $D$ besteht. Dabei wird vorausgesetzt, dass untereinander stehende Matrizen (hier: $A, C$ bzw.

$B, D$) stets dieselbe Spaltenzahl und nebeneinander stehende Matrizen ($A, B$ bzw. $C, D$) dieselbe Zeilenzahl haben. Das Rechnen mit Matrizen in Blockschreibweise wird durch eine Reihe von Rechenregeln vereinfacht.

**(2.8) Proposition** Seien $A, B, C, D$ Matrizen über $K$.

(i) Stimmt die Spaltenzahl von $A$ und $B$ mit der Zeilenzahl von $C$ überein, dann gilt

$$\begin{pmatrix} A \\ B \end{pmatrix} C \quad = \quad \begin{pmatrix} AC \\ BC \end{pmatrix}.$$

(ii) Stimmt die Spaltenzahl von $A$ mit der Zeilenzahl von $B$ und $C$ überein, dann gilt

$$A \begin{pmatrix} B & C \end{pmatrix} \quad = \quad \begin{pmatrix} AB & AC \end{pmatrix}.$$

(iii) Stimmt die Spaltenzahl von $A$ mit der Zeilenzahl von $C$ und die Spaltenzahl von $B$ mit der Zeilenzahl von $D$ überein, dann gilt

$$\begin{pmatrix} A & B \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} \quad = \quad AC + BD.$$

***Beweis:*** Wir beschränken uns auf den Beweis von (iii). Nach Voraussetzung gilt $A \in \mathscr{M}_{m \times n_1, K}$, $B \in \mathscr{M}_{m \times n_2, K}$, $C \in \mathscr{M}_{n_1 \times r, K}$ und $D \in \mathscr{M}_{n_2 \times r, K}$ für geeignete $m, n_1, n_2, r \in \mathbb{N}$. Die Matrix $AC + BD$ auf der rechten Seite ist in $\mathscr{M}_{m \times r, K}$ enthalten. Seien nun $k, \ell$ mit $1 \le k \le m$ und $1 \le \ell \le r$ vorgegeben. Zu zeigen ist, dass der Eintrag des Matrixprodukts links an der Position $(k, \ell)$ mit dem Eintrag der Matrix $AC + BD$ an derselben Stelle übereinstimmt. Um den Eintrag auf der linken Seite auszurechnen, muss die $k$-te Zeile des Faktors $(A \quad B)$ mit der $\ell$-ten Spalte des zweiten Faktors multipliziert werden. Dies liefert den Wert

$$\sum_{j=1}^{n_1} a_{kj} c_{j\ell} + \sum_{j=1}^{n_2} b_{kj} d_{j\ell}.$$

Die erste Summe entspricht dem Eintrag von $AC$ an der Stelle $(k, \ell)$, die zweite Summe dem Eintrag von $BD$ an derselben Position. Insgesamt erhalten wir also den Eintrag von $AC + BD$ an der Stelle $(k, \ell)$. □

Wir demonstrieren die Funktionsweise der Rechenregel (2.8) (iii) für Blockmatrizen anhand eines Beispiels und betrachten dazu die vier Matrizen

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad , \quad B = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \quad , \quad C = \begin{pmatrix} 9 & 10 \\ 11 & 12 \end{pmatrix} \quad , \quad D = \begin{pmatrix} 13 & 14 \\ 15 & 16 \end{pmatrix}.$$

Es gilt

$$AC = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 9 & 10 \\ 11 & 12 \end{pmatrix} = \begin{pmatrix} 31 & 34 \\ 71 & 78 \end{pmatrix} \qquad \text{und} \qquad BD = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \begin{pmatrix} 13 & 14 \\ 15 & 16 \end{pmatrix} = \begin{pmatrix} 155 & 166 \\ 211 & 226 \end{pmatrix}$$

und somit

$$AC + BD \quad = \quad \begin{pmatrix} 155 & 166 \\ 211 & 226 \end{pmatrix} + \begin{pmatrix} 31 & 34 \\ 71 & 78 \end{pmatrix} \quad = \quad \begin{pmatrix} 186 & 200 \\ 282 & 304 \end{pmatrix}.$$

Eine direkte Multiplikation der zusammengesetzten Matrizen liefert dasselbe Ergebnis:

$$\begin{pmatrix} AB \end{pmatrix}\begin{pmatrix} C \\ D \end{pmatrix} \quad = \quad \begin{pmatrix} 1 & 2 & 5 & 6 \\ 3 & 4 & 7 & 8 \end{pmatrix}\begin{pmatrix} 9 & 10 \\ 11 & 12 \\ 13 & 14 \\ 15 & 16 \end{pmatrix} \quad = \quad \begin{pmatrix} 186 & 200 \\ 282 & 304 \end{pmatrix}.$$

Allgemeiner kann gezeigt werden, dass man Matrizen mit beliebiger Aufteilung „blockweise" multiplizieren kann, wobei lediglich vorausgesetzt werden muss, dass die Teilmatrizen, die dabei multipliziert werden sollen, „zusammenpassen".

**(2.9) Proposition** Seien $m,n,r \in \mathbb{N}$, seien $A^{(i,j)}$ für $1 \le i \le m$, $1 \le j \le n$ und $B^{(j,k)}$ für $1 \le j \le n$ und $1 \le k \le r$ Matrizen mit der Eigenschaft, dass die Spaltenzahl von $A^{(i,j)}$ jeweils mit der Zeilenzahl von $B^{(j,k)}$ übereinstimmt, für alle $i,j,k$. Außerdem setzen wir voraus, dass die Zeilenzahlen von $A^{(i,j)}$ für festes $i$ und die Spaltenzahlen von $B^{(j,k)}$ für festes $k$ jeweils gleich sind. Dann gilt

$$\begin{pmatrix} A^{(1,1)} & \cdots & A^{(1,n)} \\ \vdots & & \vdots \\ A^{(m,1)} & \cdots & A^{(m,n)} \end{pmatrix}\begin{pmatrix} B^{(1,1)} & \cdots & B^{(1,r)} \\ \vdots & & \vdots \\ B^{(n,1)} & \cdots & B^{(n,r)} \end{pmatrix} \quad = \quad \begin{pmatrix} C^{(1,1)} & \cdots & C^{(1,r)} \\ \vdots & & \vdots \\ C^{(m,1)} & \cdots & C^{(m,r)} \end{pmatrix}$$

mit $C^{(i,k)} = \sum_{j=1}^{n} A^{(i,j)}B^{(j,k)}$ für $1 \le i \le m$ und $1 \le k \le r$.

Wir geben den Beweis nur der Vollständigkeit halber an, für den weiteren Verlauf ist er ohne Belang. Für alle $i,j,k$ sei $m_i \times n_j$ jeweils das Format der Matrix $A^{(i,j)}$ und $n_j \times r_k$ das Format der Matrix $B^{(j,k)}$. Dann hat die Matrix $C^{(i,k)}$ jeweils das Format $m_i \times r_k$. Wir bezeichnen die Matrix auf der rechten Seite der zu beweisenden Gleichung mit $D$ und die Matrix auf der linken Seite mit $C$. Beide Matrizen haben das Format $m_0 \times r_0$ mit $m_0 = \sum_{i=1}^{m} m_i$ und $r_0 = \sum_{k=1}^{r} r_k$. Außerdem sei $A$ die Matrix mit den Blöcken $A^{(i,j)}$ und $B$ die Matrix mit den Blöcken $B^{(j,k)}$. Nach Definition gilt $D = AB$.

Seien nun $p \in \{1,...,m_0\}$ und $q \in \{1,...,r_0\}$ vorgegeben. Zu zeigen ist $c_{pq} = d_{pq}$. Der Eintrag $d_{pq}$ kommt dadurch zu Stande, dass die $p$-te Zeile von $A$ mit der $q$-ten Spalte von $B$ multipliziert wird. Wir nehmen nun an, dass $i \in \{1,...,m\}$ und $k \in \{1,...,r\}$ so gewählt sind, dass die $p$-te Zeile der Matrix $A$ durch die $f$-ten Zeilen der Matrizen $A^{(i,1)}, A^{(i,2)}, ..., A^{(i,n)}$ läuft, und ebenso die $q$-te Spalte von $B$ durch die $g$-ten Spalten der Matrizen $B^{(1,k)}, B^{(2,k)}, ..., B^{(n,k)}$. Dabei ist $f \in \{1,...,m_i\}$ und $g \in \{1,...,r_k\}$. Setzen wir $n_0 = \sum_{j=1}^{n} n_j$, dann gilt

$$d_{pq} \quad = \quad \sum_{j=1}^{n_j} a_{pj} b_{jq} \quad = \quad \sum_{j=1}^{n} \sum_{\ell=1}^{n_j} a_{f\ell}^{(i,j)} b_{\ell g}^{(j,k)}.$$

Nun läuft die $p$-te Zeile von $C$ auch durch die $f$-ten Zeilen der Matrizen $C^{(i,1)}, C^{(i,2)}, ..., C^{(i,n)}$, und die $q$-te Spalte von $C$ entsprechend durch die $g$-ten Spalten der Matrizen $C^{(1,k)}, C^{(2,k)}, ..., C^{(n,k)}$. Wegen $C^{(i,k)} = \sum_{j=1}^{n} A^{(i,j)}B^{(j,k)}$ für

$1 \le i \le m$ und $1 \le k \le r$ erhalten wir dann wie gewünscht

$$c_{pq} \quad = \quad c^{(i,k)}_{fg} \quad = \quad \sum_{j=1}^{n} (A^{(i,j)} B^{(j,k)})_{fg} \quad = \quad \sum_{j=1}^{n} \sum_{\ell=1}^{n_j} a^{(i,j)}_{f\ell} b^{(j,k)}_{\ell g} \quad = \quad d_{pq}. \qquad \square$$

**(2.10) Definition** Eine Matrix aus $\mathscr{M}_{m,K}$ der Form $\mathrm{M}_{k,\lambda} = E^{(m)} + (\lambda - 1) B^{(m\times m)}_{kk}$ mit $k \in \{1,...,m\}$ und $\lambda \in K^\times$ oder der Form $\mathrm{A}_{k,\ell,\lambda} = E^{(m)} + \lambda B^{(m\times m)}_{\ell k}$ mit $k,\ell \in \{1,...,m\}$ und $\lambda \in K$ wird ***Elementarmatrix*** genannt.

In Blockschreibweise hat die Elementarmatrix $\mathrm{M}_{k,\lambda}$ die Form

$$\mathrm{M}_{k,\lambda} \quad = \quad \begin{pmatrix} E^{(k-1)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \lambda & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & E^{(m-k)} \end{pmatrix}$$

wobei die Einträge $\mathbf{0}$ jeweils für Nullmatrizen der passenden Größe stehen. Die Elementarmatrix $\mathrm{A}_{k,\ell,\lambda}$ hat im Fall $k < \ell$ bzw. $k > \ell$ die Form

$$\mathrm{A}_{k,\ell,\lambda} \quad = \quad \begin{pmatrix} E^{(k-1)} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & E^{(\ell-k-1)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \lambda & \mathbf{0} & 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & E^{(m-\ell)} \end{pmatrix}$$

beziehungsweise

$$\mathrm{A}_{k,\ell,\lambda} \quad = \quad \begin{pmatrix} E^{(\ell-1)} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 & \mathbf{0} & \lambda & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & E^{(k-\ell-1)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & E^{(m-k)} \end{pmatrix}$$

**(2.11) Proposition** Sei $A \in \mathscr{M}_{m\times n,K}$.

(i) Sei $\lambda \in K^\times$ und $k \in \{1,...,m\}$. Multipliziert man die Matrix $A$ von links mit der Elementarmatrix $\mathrm{M}_{k,\lambda}$, so bewirkt dies eine Multiplikation der $k$-ten Zeile mit dem Wert $\lambda$.

(ii) Seien $k,\ell \in \{1,...,m\}$ mit $k \neq \ell$ und $\lambda \in K$. Multipliziert man die Matrix $A$ mit der Elementarmatrix $\mathrm{A}_{k,\ell,\lambda}$, dann wird das $\lambda$-fache der $k$-ten Zeile zur $\ell$-ten Zeile von $A$ addiert.

***Beweis:*** Beide Aussagen lassen sich durch die blockweise Multiplikation von Matrizen unmittelbar nachrechnen.

zu (i) Sei $B \in \mathscr{M}_{(k-1)\times n,K}$ die Teilmatrix bestehend aus den oberen $k-1$ und $C \in \mathscr{M}_{(m-k)\times n,K}$ die Teilmatrix bestehend aus den unteren $m-k$ Zeilen von $A$. Ferner sei $z \in \mathscr{M}_{1\times n,K}$ die $k$-te Zeile von $A$. Dann gilt

$$\mathrm{M}_{k,\lambda}A = \begin{pmatrix} E^{(k-1)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \lambda & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & E^{(m-k)} \end{pmatrix} \begin{pmatrix} B \\ z \\ C \end{pmatrix} = \begin{pmatrix} E^{(k-1)}B + \mathbf{0}z + \mathbf{0}C \\ \mathbf{0}B + \lambda z + \mathbf{0}C \\ \mathbf{0}B + \mathbf{0}z + E^{(m-k)}C \end{pmatrix} = \begin{pmatrix} B \\ \lambda z \\ C \end{pmatrix}$$

zu (ii) Hier beschränken wir uns auf den Fall $k < \ell$ und teilen die Matrix $A$ auf in die Matrix $B \in \mathscr{M}_{(k-1)\times n,K}$ bestehend aus den ersten $k-1$ Zeilen, der Matrix $C \in \mathscr{M}_{(\ell-k-1)\times n,K}$ bestehend aus der $(k+1)$-ten bis zur $(\ell-1)$-ten Zeile und der Matrix $D \in \mathscr{M}_{(m-\ell)\times n,K}$ bestehend aus den unteren $m-\ell$ Zeilen. Ferner seien $z_k, z_\ell \in \mathscr{M}_{1\times n,K}$ die $k$-te und $\ell$-te Zeile von $A$. Dann erhalten wir

$$\mathrm{A}_{k,\ell,\lambda}A = \begin{pmatrix} E^{(k-1)} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & E^{(\ell-k-1)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \lambda & \mathbf{0} & 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & E^{(m-\ell)} \end{pmatrix} \begin{pmatrix} B \\ z_k \\ C \\ z_\ell \\ D \end{pmatrix} = \begin{pmatrix} B \\ z_k \\ C \\ \lambda z_k + z_\ell \\ D \end{pmatrix}$$

□

Wir zeigen anhand zweier Beispiele, dass die Multiplikation mit Elementarmatrizen tatsächlich den angegebenen Effekt hat. Die Multiplikation einer dreizeiligen Matrix mit $\mathrm{M}_{2,3}$ von links bewirkt eine Multiplikation der zweiten Zeile mit dem Wert 3. Zum Beispiel gilt

$$\mathrm{M}_{2,3}\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 15 & 18 & 21 & 24 \\ 9 & 10 & 11 & 12 \end{pmatrix}.$$

Ebenso bewirkt die Multiplikation mit $\mathrm{A}_{1,3,2}$ von links, dass das zweifache der ersten Zeile zur dritten addiert wird, zum Beispiel

$$\mathrm{A}_{1,3,2}\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 11 & 14 & 17 & 20 \end{pmatrix}.$$

Jede Zeilenumformung einer Matrix $A$ lässt sich also durch Multiplikation mit einer Elementarmatrix von links realisieren. Dementsprechend führt die Multiplikation von $A$ mit einem Produkt $E_m \cdot E_{m-1} \cdot \ldots \cdot E_1$ von Elementarmatrizen dazu, dass $A$ einer Folge von $m$ Zeilenumformungen unterworfen wird. Wir bezeichnen die Menge aller Matrizen in $\mathscr{M}_{m,K}$, die sich als Produkt von Elementarmatrizen schreiben lassen, mit $\mathscr{E}_m(K)$.

Mit Hilfe des Matrixkalküls werden wir nun zeigen, dass sich jede Matrix durch eine endliche Anzahl von Zeilenumformungen auf normierte Zeilenstufenform bringen lässt.

**(2.12) Lemma** Sei $A \in \mathscr{M}_{m\times 1,K}$ eine Matrix, die aus einer einzigen Spalte besteht, also eine Matrix der Form $A = {}^{\mathrm{t}}(a_1\ a_2\ \ldots\ a_m)$. Sind nicht alle Einträge von $A$ gleich Null, dann gibt es ein Produkt $E \in \mathscr{E}_m(K)$ von Elementarmatrizen mit $EA = {}^{\mathrm{t}}(1_K\ 0_K\ 0_K\ \ldots\ 0_K)$.

***Beweis:*** Auf Grund unserer Vorbemerkung genügt es zu zeigen, dass $A$ durch eine endliche Abfolge von elementaren Zeilenumformungen auf die Gestalt ${}^{\mathrm{t}}(1\ 0\ \ldots\ 0)$ gebracht werden kann. Auch Vertauschungen von Zeilen sind zulässig, weil diese (wie oben gesehen) durch endlich viele elementare Umformungen realsierbar sind. Nach Voraussetzung gibt es ein $k \in \{1, \ldots, m\}$ mit $a_k \neq 0_K$. Nach Multiplikation der $k$-ten Zeile mit $a_k^{-1}$ und Vertauschung der $k$-ten mit der ersten Zeile gilt $a_1 = 1_K$. Nun addieren wir für $\ell = 2, \ldots, m$ jeweils das $(-a_\ell)$-fache der ersten Zeile zur $\ell$-ten. Dies führt dazu, dass sämtliche Einträge der Matrix mit Ausnahme des ersten zu Null werden. □

**(2.13) Satz** Jede Matrix $A \in \mathscr{M}_{m\times n,K}$ kann durch endlich viele elementare Zeilenumformungen auf normierte ZSF gebracht werden. Eine äquivalente Formulierung dieser Aussage lautet: Es gibt eine Matrix $E \in \mathscr{E}_m(K)$, so dass $EA$ in normierter ZSF vorliegt.

***Beweis:*** Wir zeigen zunächst, dass $A$ auf ZSF gebracht werden kann und führen den Beweis durch vollständige Induktion über die Anzahl $n$ der Spalten. Der Fall $n = 1$ ist mit (2.12) bereits erledigt, denn nach Definition ist ${}^{\mathrm{t}}(1_K\ 0_K\ \ldots\ 0_K)$ eine Matrix in ZSF (mit den Kennzahlen $r = j_1 = 1$). Sei nun $n \in \mathbb{N}$, und setzen wir die Aussage für dieses $n$ voraus. Sei außerdem $A \in \mathscr{M}_{m\times(n+1),K}$ eine beliebige Matrix. Wir müssen zeigen, dass $A$ auf ZSF gebracht werden kann und unterscheiden dafür zwei Fälle.

*1. Fall:* Die erste Spalte von $A$ hat nur Nulleinträge.
Dann hat $A$ die Form $(\mathbf{0}^{(m\times 1)}\ B)$ mit einer Matrix $B \in \mathscr{M}_{m\times n,K}$. Nach Induktionsvoraussetzung gibt es eine Matrix $E \in \mathscr{E}_m(K)$, so dass $B' = EB$ in ZSF vorliegt, mit gewissen Kennzahlen $r, j_1, \ldots, j_r$. Es gilt

$$EA \quad = \quad E(\mathbf{0}^{(m\times 1)}\ B) \quad = \quad (\mathbf{0}^{(m\times 1)}\ EB) \quad = \quad (\mathbf{0}^{(m\times 1)}\ B').$$

Wie man leicht überprüft, liegt auch $(\mathbf{0}^{(m\times 1)}\ B')$ die Matrix in ZSF vor, mit den Kennzahlen $r, j_1 + 1, \ldots, j_r + 1$.

*2. Fall:* Die erste Spalte von $A$ hat Einträge ungleich Null.
In diesem Fall kann $A$ in der Blockgestalt

$$A \quad = \quad \begin{pmatrix} a_{11} & z \\ s & C \end{pmatrix}$$

dargestellt werden, mit $a_{11} \in K$, $z \in \mathscr{M}_{1\times n,K}$, $s \in \mathscr{M}_{(m-1)\times 1,K}$ und $C \in \mathscr{M}_{(m-1)\times n,K}$, wobei die Teilmatrix ${}^{\mathrm{t}}(a_{11} s)$ nicht nur Nulleinträge enthält. Nach (2.12) gibt es eine Matrix $E \in \mathscr{E}_m(K)$ mit

$$E\begin{pmatrix} a_{11} \\ s \end{pmatrix} = \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix}$$

und wir erhalten

$$EA = \begin{pmatrix} 1 & z' \\ \mathbf{0} & C' \end{pmatrix}$$

mit geeigneten Matrizen $z' \in \mathscr{M}_{1\times n,K}$ und $C' \in \mathscr{M}_{(m-1)\times n,K}$. Nach Induktionsvoraussetzung existiert nun eine Matrix $E' \in \mathscr{E}_{m-1}(K)$, so dass $E'C'$ in ZSF vorliegt, mit gewissen Kennzahlen $r, j_1, ..., j_r$. Außerdem gilt

$$\begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & E' \end{pmatrix}\begin{pmatrix} 1 & z' \\ \mathbf{0} & C' \end{pmatrix} = \begin{pmatrix} 1 & z' \\ \mathbf{0} & E'C' \end{pmatrix}$$

Wieder überprüft man, dass sich die Matrix rechts in ZSF befindet, mit Kennzahlen $r+1, 1, j_1+1, ..., j_r+1$. Anhand der Gleichung

$$\begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & U \end{pmatrix}\begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & V \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & UV \end{pmatrix}$$

für Blockmatrizen sieht man, dass mit $E'$ auch die Matrix

$$\begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & E' \end{pmatrix}$$

als Produkt von Elementarmatrizen darstellbar ist.

Zu zeigen bleibt, dass jede Matrix in ZSF durch elementare Zeilenumformungen auf *normierte* ZSF gebracht werden kann. Dazu setzen wir voraus, dass $A$ bereits in ZSF mit den Kennzahlen $r, j_1, ..., j_r$ vorliegt. Um $a_{ij_i} = 1$ für $1 \le i \le r$ zu erreichen, dividiert man einfach für jedes $i$ die $i$-te Zeile durch $a_{ij_i}$. Die ZSF der Matrix wird durch diese Operation nicht zerstört, da die Eigenschaft eines Eintrages, gleich Null oder ungleich Null zu sein, dadurch nicht verändert wird.

Die Bedingung $a_{kj_i} = 0$ für $k < i$ kann dadurch erfüllt werden, dass man nacheinander für die Zeilennummern $i = r, r-1, r-2, ..., 1$ jeweils das $a_{kj_i}$-fache der $i$-ten Zeile von der $k$-ten Zeile subtrahiert, für $1 \le k < i$. Dabei ist darauf zu achten, dass in keinem Schritt die ZSF beeinträchtigt wird und die erreichte Form für die Spalten $j_\ell$ mit $\ell > i$ erhalten bleibt. Die ZSF bleibt erhalten, da die $i$-te Zeile ihren ersten Eintrag $\neq 0$ erst in der Spalte $j_i$ hat und $j_i > j_k$ für $1 \le k < i$ gilt. Somit werden weder die Zeilenköpfe der darüberliegenden Zeilen noch die Einträge links davon verändert. Die Zeilen unterhalb der $i$-ten bleiben völlig unverändert. Auch die Bedingung $a_{kj_\ell} = 0$ für $\ell > i$ und $k < \ell$ bleibt erhalten, da der einzige Eintrag ungleich Null in der $j_\ell$-ten Spalte der Eintrag $a_{\ell j_\ell} = 1$ ist, und dieser spielt wegen $\ell > i$ bei der Zeilenumformung keine Rolle. □

Man kann sich an diesem „induktiven" Beweisschema orientieren, um eine beliebige, konkret vorgegebene Matrix $A \in \mathscr{M}_{m\times n,K}$ zunächst auf Zeilenstufenform und dann auf normierte Zeilenstufenform zu bringen. Diese systematische Vorgehensweise bezeichnet man dann als das ***Gaußsche Eliminationsverfahren***.

# § 3. Die allgemeine lineare Gruppe

***Inhaltsübersicht***

In diesem Abschnitt definieren wir zunächst zwei grundlegende algebraische Strukturen, auf die wir im weiteren Verlauf der Vorlesung häufig zurückgreifen werden: Monoide und Gruppen. Diese Begriffe werden anschließend eingesetzt, um die algebraischen Eigenschaften der Matrizen genauer zu untersuchen. Insbesondere wird dabei ein Kriterium für die *Invertierbarkeit* von Matrizen herauskommen. Wie wir sehen werden, bilden die invertierbaren Matrizen eine Gruppe, die sogenannte *allgemeine lineare Gruppe*, die in der gesamten Mathematik eine wichtige Rolle spielt.

***Wichtige Definitionen und Sätze***

- Verknüpfung auf einer Menge
- assoziative und kommutative Verknüpfungen
- Monoid, Gruppe, Neutralelement, inverses Element
- Abgeschlossenheit einer Teilmenge unter einer Verknüpfung
- Neudefinition der Ringe und Körper

**(3.1) Definition** Eine ***Verknüpfung*** auf einer Menge $A$ ist eine Abbildung $A \times A \to A$.

Als Bezeichnungen für eine Verknüpfung sind die Symbole $\cdot$, $\odot$, $*$, $+$, $\oplus$ und einige Varianten üblich. Wird eines der Symbole $\cdot$, $\odot$, $*$ verwendet, dann spricht man von einer ***multiplikativen*** Verknüpfung, bei $+$ oder $\oplus$ nennt man sie ***additiv***. Die beiden Typen unterscheiden sich aber außschließlich durch das verwendete Symbol, mathematisch gesehen besteht zwischen einer additiven und einer multiplikativen Verknüpfung keinerlei Unterschied.

Multiplikative Verknüpfungssymbole werden zur Vereinfachung der Notation häufig auch weggelassen, d.h. an Stelle von $a \cdot b$ schreibt man einfach $ab$. Sollen mehrere Elemente miteinander verknüpft werden, so ist die Verwendung von ***Klammern*** üblich, um die Reihenfolge der angewendeten Verknüpfungen anzuzeigen. So bedeutet zum Beispiel der Ausdruck $a(b(cd))$, dass zunächst das Element $x_1 = cd$ gebildet wird, anschließend $x_2 = bx_1$ und schließlich $x_3 = ax_2$.

**(3.2) Definition** Eine Verknüpfung $\cdot$ auf einer Menge $A$ bezeichnet man als

(i) ***kommutativ***, wenn $ab = ba$ für alle $a, b \in A$

(ii) ***assoziativ***, wenn $a(bc) = (ab)c$ für alle $a, b, c \in A$ erfüllt ist.

Bei assoziativen Verknüpfungen können die Klammern auch weggelassen werden. Für beliebige Elemente $a, b, c, d \in A$ ist dann zum Beispiel $abcd$ eine Kurzschreibweise für das Element $a(b(cd))$, welches auf Grund der Assoziativität mit jedem anders geklammerten Ausdruck, etwa $(ab)(cd)$ oder $a((bc)d)$, übereinstimmt.

**(3.3) Definition** Eine ***Monoid*** ist ein Paar $(G, \cdot)$ bestehend aus einer Menge $G \neq \emptyset$ und einer Verknüpfung $\cdot : G \times G \to G$ mit folgenden Eigenschaften.

(i) Die Verknüpfung ist assoziativ.

(ii) Es gibt ein Element $e \in G$ mit $ae = ea = a$ für alle $a \in G$.

Ist die Verknüpfung $\cdot$ auch kommutativ, dann wird $(G, \cdot)$ ein ***kommutatives*** oder ***abelsches*** Monoid genannt.

Ein Element $e$ mit der unter (ii) genannten Eigenschaft bezeichnet man als ***Neutralelement*** des Monoids. Jeder Monoid $(G, \cdot)$ besitzt nur ein Neutralelement. Ist nämlich $e'$ ein weiteres, dann gilt $e = ee' = e'$. Aus diesem Grund kann von *dem* Neutralelement des Monoids gesprochen werden. Man verwendet die feststehende Bezeichung $e_G$ für dieses ausgezeichnete Element. Bei einer additiven Verknüpfung wird im allgemeinen an Stelle von $e_G$ das Symbol $0_G$ verwendet.

**(3.4) Definition** Sei $(G, \cdot)$ ein Monoid. Wenn für jedes $a \in G$ ein Element $b \in G$ existiert, so dass die Gleichungen

$$ab \quad = \quad ba \quad = \quad e_G$$

erfüllt sind, so bezeichnet man $(G, \cdot)$ als eine ***Gruppe***. Wie bei den Monoiden spricht man von einer kommutativen oder abelschen Gruppe, wenn die Verknüpfung kommutativ ist.

Für jedes $a \in G$ bezeichnet man ein Element $b$ mit der angegebenen Eigenschaft als ein zu $a$ ***inverses*** Element. Jedes Element $a$ in einem Monoid $(G, \cdot)$ besitzt höchstens ein Inverses. Sind nämlich $b, c \in G$ beide zu $a$ invers, dann gilt $ba = e_G$ und $ac = e_G$, insgesamt also

$$b \quad = \quad be_G \quad = \quad b(ac) \quad = \quad (ba)c \quad = \quad e_G c \quad = \quad c.$$

Auf Grund dieser Eindeutigkeit ist es zulässig, für das Inverse von $a$ die feststehende Bezeichnung $a^{-1}$ zu verwenden. Bei einer additiven Verknüpfung ist die Bezeichnung $-a$ für das Inverse üblich. Diejenigen Elemente in einem Monoid, die ein Inverses besitzen, bezeichnet man als ***invertierbare*** Elemente. Eine Gruppe ist also ein Monoid, in dem jedes Element invertierbar ist.

Für die invertierbaren Elemente eines Monoids gelten eine Reihe von wichtigen Rechenregeln.

**(3.5) Lemma** Sei $(G, \cdot)$ ein Monoid, und seien $a, b \in G$ invertierbare Elemente.

(i) Ist $c \in G$ ein Element mit $ac = e_G$, dann gilt $c = a^{-1}$. Ebenso folgt aus $ca = e_G$ bereits die Gleichheit $c = a^{-1}$.

(ii) Das Neutralelement $e_G$ ist invertierbar, und es gilt $e_G^{-1} = e_G$.

(iii) Das Element $ab$ ist invertierbar, und es gilt $(ab)^{-1} = b^{-1}a^{-1}$.

(iv) Auch das Inverse $a^{-1}$ von $a$ ist invertierbar, und es gilt $(a^{-1})^{-1} = a$.

***Beweis:*** zu (i) Aus $ac = e_G$ und $a^{-1}a = e_G$ folgt insgesamt $c = e_G c = (a^{-1}a)c = a^{-1}(ac) = a^{-1}e_G = a^{-1}$. Setzen wir $ca = e_G$ voraus, so erhalten wir die Gleichheit entsprechend durch $aa^{-1} = e_G$ und die Rechnung $c = ce_G = c(aa^{-1}) = (ca)a^{-1} = c(aa^{-1}) = (ca)a^{-1} = e_G a^{-1} = a^{-1}$.

zu (ii) Ein Element $c \in G$ ist ein Inverses von $e_G$, wenn die Gleichungen $e_G c = ce_G = e_G$ erfüllt sind. Aus der Gleichung $e_G e_G = e_G$ folgt also, dass das Neutralelement $e_G$ sein eigenes Inverses ist.

zu (iii) Dies folgt aus den Gleichungen $(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = ae_G a^{-1} = aa^{-1} = e_G$ und $(b^{-1}a^{-1})(ab) = b^{-1}(aa^{-1})b = b^{-1}e_G b = b^{-1}b = e_G$.

zu (iv) Ein Element $c \in G$ ist ein Inverses von $a^{-1}$, wenn die Gleichungen $a^{-1}c = e_G$ und $ca^{-1} = e_G$ gelten. Also folgt die Aussage aus den Gleichungen $a^{-1}a = e_G$ und $aa^{-1} = e_G$. □

Sei $*$ eine Verknüpfung auf einer Menge $A$. Eine Teilmenge $U \subseteq A$ bezeichnet man als ***abgeschlossen*** unter $*$, wenn für alle $a, b \in U$ auch das Element $a * b$ in $U$ enthalten ist. Dies bedeutet, dass die Einschränkung der Verknüpfungsabbildung auf die Teilmenge $U \times U \subseteq A \times A$ eine Abbildung mit dem Bildbereich $U$ liefert. Man erhält also eine neue Verknüpfung $*_U$, die nun auf der Menge $U$ definiert ist. Um die Notation aber nicht zu aufwändig werden zu lassen, behält man meistens auch auf $U$ das alte Verknüpfungssymbol bei. Gelegentlich ist die Schreibweise $*_U$ aber hilfreich, wenn betont werden soll, dass die Verknüpfungsabbildung auf $U$ gemeint ist.

Beispielsweise ist nach (3.5) (iii) in jedem Mononoid $(G, \cdot)$ die Menge der invertierbaren Elemente unter der Verknüpfung $\cdot$ abgeschlossen. Also ist auch auf der Menge dieser Elemente, die man üblicherweise mit $G^\times$ bezeichnet, eine Verknüpfung definiert. Dasselbe Lemma zeigt auch, dass das Neutralelement $e_G$ des Monoids in $G^\times$ enthalten ist. Mit jedem Element $a$ liegt auch das Inverse $a^{-1}$ in $G^\times$. Daraus folgt

**(3.6) Satz** Die Menge $G^\times$ der invertierbaren Elemente eines Monoids $(G, \cdot)$ bildet zusammen mit der auf $G^\times$ eingeschränkten Verknüpfung eine Gruppe.

Die aus dem ersten Semester bekannte Definitionen des ***Rings*** und des ***Körpers*** lassen sich mit dem Begriff des Monoids und der Gruppe nun kürzer formulieren.

**(3.7) Definition** Ein ***Ring*** ist ein Tripel $(R, +, \cdot)$ bestehend aus einer Menge $R$ und zwei Verknüpfungen $+$ und $\cdot$ auf $R$ mit folgenden Eigenschaften.

(i) Das Paar $(R, +)$ ist eine abelsche Gruppe.

(ii) Das Paar $(R, \cdot)$ ist ein abelsches Monoid.

(iii) Es gilt das Distributivgesetz $a(b + c) = ab + ac$ für alle $a, b, c \in R$.

Wir erinnern daran, dass das Neutralelement der Gruppe $(R, +)$ das ***Nullelement*** $0_R$ und das Neutralelement des Monoids $(R, \cdot)$ das ***Einselement*** $1_R$ des Rings genant wird. Ist darüber hinaus die Gruppe $R^\times$ der bezüglich Multiplikation invertierbaren Elemente gleich $R \setminus \{0_R\}$, dann nennt man den Ring $(R, +, \cdot)$ einen ***Körper***.

Wir verwenden die neu eingeführten Begriffe nun, um unser bisherigen Wissen über die Regeln zur Matrizenrechnung besser zu strukturieren. Wie immer bezeichnen wir mit $K$ einen beliebigen Körper.

**(3.8) Satz** Sei $K$ ein Körper, und seien $m, n \in \mathbb{N}$.

(i) Die Menge $\mathscr{M}_{m\times n,K}$ bildet mit der Addition von Matrizen eine abelsche Gruppe, deren Neutralelement die Nullmatrix $\mathbf{0}$ ist.

(ii) Die Menge $\mathscr{M}_{n,K}$ bildet mit der Multiplikation von Matrizen ein Monoid, mit der Einheitsmatrix $E$ als Neutralelement. Dieses ist kommutativ für $n = 1$ und nicht-kommutativ für $n \geq 2$.

***Beweis:*** zu (i) Zu zeigen ist, dass für beliebige Matrizen $A, B, C \in \mathscr{M}_{m\times n,K}$ das Assoziativgesetz $A + (B + C) = (A + B) + C$ und das Kommutativgesetz $A + B = B + A$ sowie die Gleichung $A + \mathbf{0} = A$ gültig ist. Das Nachrechnen dieser Rechenregeln ist einfacher als in (2.6) und wird deshalb hier nicht ausgeführt. Sie zeigen, dass $(\mathscr{M}_{m\times n,K}, +)$ ein abelsches Monoid ist. Ebenso leicht weist man nach, dass für jedes $A \in \mathscr{M}_{m\times n,K}$ mit der Matrix $-A = (-1_K)A$ die Gleichung $A + (-A) = \mathbf{0}$ erfüllt ist. Also ist unser Monoid sogar eine abelsche Gruppe.

zu (ii) Seien $A, B, C \in \mathscr{M}_{n,K}$ vorgegeben. Bereits in (2.6) wurde das Assoziativgesetz $A(BC) = (AB)C$ wurde nachgewiesen, und nach (2.7) gilt jeweils $AE = EA = A$. Also ist $(\mathscr{M}_{n,K}, \cdot)$ ein Monoid. Im Fall $n = 1$ gilt $(a)(b) = (ab) = (ba) = (b)(a)$ für alle $a, b \in K$, d.h. die Kommutativität von $(\mathscr{M}_{1,K}, \cdot)$ folgt direkt aus der Kommutativität von $(K, \cdot)$. Um zu zeigen, dass das Monoid $(\mathscr{M}_{n,K}, \cdot)$ für $n = 2$ nicht kommutativ ist, betrachten wir die Matrizen

$$A = \begin{pmatrix} 0_K & 1_K \\ 1_K & 0_K \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1_K & 1_K \\ 0_K & 1_K \end{pmatrix}$$

Es gilt

$$AB \quad = \quad \begin{pmatrix} 0_K & 1_K \\ 1_K & 0_K \end{pmatrix} \begin{pmatrix} 1_K & 1_K \\ 0_K & 1_K \end{pmatrix} \quad = \quad \begin{pmatrix} 0_K & 1_K \\ 1_K & 1_K \end{pmatrix}$$

aber

$$BA \quad = \quad \begin{pmatrix} 1_K & 1_K \\ 0_K & 1_K \end{pmatrix} \begin{pmatrix} 0_K & 1_K \\ 1_K & 0_K \end{pmatrix} \quad = \quad \begin{pmatrix} 1_K & 1_K \\ 1_K & 0_K \end{pmatrix}.$$

Weil in jedem Körper Null- und Einselement verschieden sind, gilt $0_K \neq 1_K$ und somit $AB \neq BA$. Für den Nachweis der Nichtkommutativität im Fall $n > 2$ greifen wir auf die Blockdarstellung von Matrizen zurück. Mit Hilfe der soeben definierten $2 \times 2$-Matrizen $A, B$ setzen wir

$$\tilde{A} \quad = \quad \begin{pmatrix} A & \mathbf{0}^{(2\times(n-2)} \\ \mathbf{0}^{((n-2)\times 2} & \mathbf{0}^{(n-2)\times(n-2)} \end{pmatrix} \qquad \text{und} \qquad \tilde{B} \quad = \quad \begin{pmatrix} B & \mathbf{0}^{(2\times(n-2)} \\ \mathbf{0}^{((n-2)\times 2} & \mathbf{0}^{(n-2)\times(n-2)} \end{pmatrix}.$$

Es gilt dann

$$\tilde{A}\tilde{B} \quad = \quad \begin{pmatrix} A & \mathbf{0}^{(2\times(n-2))} \\ \mathbf{0}^{((n-2)\times 2)} & \mathbf{0}^{((n-2)\times(n-2))} \end{pmatrix} \begin{pmatrix} B & \mathbf{0}^{(2\times(n-2))} \\ \mathbf{0}^{((n-2)\times 2)} & \mathbf{0}^{((n-2)\times(n-2))} \end{pmatrix} \quad = \quad \begin{pmatrix} AB & \mathbf{0}^{(2\times(n-2))} \\ \mathbf{0}^{((n-2)\times 2)} & \mathbf{0}^{((n-2)\times(n-2))} \end{pmatrix}$$

und

$$\tilde{B}\tilde{A} \quad = \quad \begin{pmatrix} B & \mathbf{0}^{(2\times(n-2))} \\ \mathbf{0}^{((n-2)\times 2)} & \mathbf{0}^{((n-2)\times(n-2))} \end{pmatrix} \begin{pmatrix} A & \mathbf{0}^{(2\times(n-2))} \\ \mathbf{0}^{((n-2)\times 2)} & \mathbf{0}^{((n-2)\times(n-2))} \end{pmatrix} \quad = \quad \begin{pmatrix} BA & \mathbf{0}^{(2\times(n-2))} \\ \mathbf{0}^{((n-2)\times 2)} & \mathbf{0}^{((n-2)\times(n-2))}. \end{pmatrix}$$

Wegen $AB \neq BA$ gilt auch $\tilde{A}\tilde{B} \neq \tilde{B}\tilde{A}$. Somit ist das Monoid $(\mathscr{M}_{n,K}, \cdot)$ auch im Fall $n > 2$ nicht kommutativ. □

Eine Matrix $A \in \mathscr{M}_{n,K}$ wird ***invertierbar*** genannt, wenn eine Matrix $B \in \mathscr{M}_{n,K}$ mit $AB = BA = E^{(n)}$ existiert. Beispielsweise sind die Elementarmatrizen $\mathrm{M}_{k,\lambda}$ und $\mathrm{A}_{k,\ell,\lambda}$ invertierbar. Für alle $\lambda \in K^\times$ und $k \in \{1,...,n\}$ gilt

$$\mathrm{M}_{k,\lambda}\mathrm{M}_{k,\lambda^{-1}} \quad = \quad \begin{pmatrix} E^{(k-1)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \lambda & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & E^{(n-k)} \end{pmatrix} \begin{pmatrix} E^{(k-1)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \lambda^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & E^{(n-k)} \end{pmatrix}$$

$$= \quad \begin{pmatrix} E^{(k-1)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & E^{(n-k)} \end{pmatrix} \quad = \quad E^{(n)}$$

und ebenso $\mathrm{M}_{k,\lambda^{-1}}\mathrm{M}_{k,\lambda} = E^{(n)}$. Die Invertierbarkeit von $\mathrm{A}_{k,\ell,\lambda}$ zeigen wir nur im Fall $k < \ell$: Das Produkt $\mathrm{A}_{k,\ell,\lambda}\mathrm{A}_{k,\ell,-\lambda}$ ist gegeben durch

$$\begin{pmatrix} E^{(k-1)} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & E^{(\ell-k-1)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \lambda & \mathbf{0} & 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & E^{(n-\ell)} \end{pmatrix} \begin{pmatrix} E^{(k-1)} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & E^{(\ell-k-1)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\lambda & \mathbf{0} & 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & E^{(n-\ell)} \end{pmatrix}$$

$$= \quad \begin{pmatrix} E^{(k-1)} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & E^{(\ell-k-1)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & E^{(n-\ell)} \end{pmatrix} \quad = \quad E^{(n)}$$

Aus Satz (3.6), angewendet auf das Monoid $(\mathscr{M}_{n,K}, \cdot)$ und die Teilmenge der invertierbaren Matrizen, folgt

**(3.9) Folgerung** Die Menge der invertierbaren $n \times n$-Matrizen über einem Körper $K$ bildet mit der Multiplikation von Matrizen eine Gruppe. Man nennt sie die ***allgemeine lineare Gruppe*** und bezeichnet sie mit $\mathrm{GL}_n(K)$.

Weil $\mathrm{GL}_n(K)$ unter Matrizenmultiplikation abgeschlossen ist und alle Elementarmatrizen in $\mathrm{GL}_n(K)$ liegen, ist auch die Menge $\mathscr{E}_n(K)$ aller Matrizen, die als Produkte von Elementarmatrizen dargestellt werden können, in $\mathrm{GL}_n(K)$ enthalten. Ist allgemein $B \in \mathrm{GL}_n(K)$, dann gilt für jedes $A \in \mathscr{M}_{n,K}$ offenbar die Äquivalenz

$$A \in \mathrm{GL}_n(K) \quad \Longleftrightarrow \quad AB \in \mathrm{GL}_n(K) \quad \Longleftrightarrow \quad BA \in \mathrm{GL}_n(K).$$

Setzen wir nämlich $A \in \mathrm{GL}_n(K)$ voraus, dann ist $AB$ invertierbar, weil die Menge der invertierbaren Matrizen unter Multiplikation abgeschlossen ist. Setzen wir umgekehrt die Invertierbarkeit von $AB$ voraus, dann folgt mit demselben Argument wegen $(AB)B^{-1} = A(BB^{-1}) = AE^{(n)} = A$ die Invertierbarkeit von $A$. Der Beweis der Äquivalenz $A \in \mathrm{GL}_n(K) \Longleftrightarrow BA \in \mathrm{GL}_n(K)$ läuft völlig analog.

Wir beschäftigen uns nun mit der Frage, wie man die Invertierbarkeit von Matrizen nachweist, und wie gegebenenfalls die inverse Matrix berechnet werden kann.

**(3.10) Satz** Lässt sich eine Matrix $A \in \mathscr{M}_{n,K}$ durch endliche viele elementare Zeilenumformungen in eine Matrix $A'$ in normierter ZSF mit Zeilenrang $r = n$ umwandeln, so ist $A$ invertierbar.

***Beweis:*** Bereits in §2 haben wir bemerkt, dass eine Matrix $A'$ in normierter ZSF mit Zeilenrang $r = n$ zwangsläufig mit der Einheitsmatrix $E^{(n)}$ übereinstimmt. Weil $A$ durch elementare Zeilenumformungen in $A' = E^{(n)}$ überführt werden kann, gibt es eine Matrix $T \in \mathscr{E}_n(K) \subseteq \mathrm{GL}_n(K)$ mit $TA = E^{(n)}$. Es folgt $A = E^{(n)}A = (T^{-1}T)A = T^{-1}(TA) = T^{-1}E^{(n)} = T^{-1}$. Damit ist die Invertierbarkeit von $A$ bewiesen. □

Die Beweisidee in (3.10) kann genutzt werden, um die zu $A$ inverse Matrix auszurechnen. Wendet man die Zeilenumformungen im Beweis statt auf $A$ auf die Blockmatrix $(A \;\; E^{(n)})$ an, so erhält man die Matrix

$$T(A \;\; E^{(n)}) \quad = \quad (TA \;\; TE^{(n)}) \quad = \quad (E^{(n)} \;\; T).$$

Aus der rechten Hälfte der umgeformten Matrix kann die Inverse von $A$ abgelesen werden, denn es gilt die Äquivalenz $A = T^{-1} \Longleftrightarrow A^{-1} = T$. Wir demonstrieren dieses Berechnungsverfahren, indem wir $A^{-1}$ für die Matrix

$$A \quad = \quad \begin{pmatrix} 3 & 0 & 2 \\ -1 & 1 & -1 \\ 7 & 0 & 5 \end{pmatrix}$$

bestimmen. Dazu schreiben wir die Einheitsmatrix $E^{(3)}$ neben unsere Matrix $A$ und formen auf normierte ZSF um.

$$\left(\begin{array}{ccc|ccc} 3 & 0 & 2 & 1 & 0 & 0 \\ -1 & 1 & -1 & 0 & 1 & 0 \\ 7 & 0 & 5 & 0 & 0 & 1 \end{array}\right) \mapsto \left(\begin{array}{ccc|ccc} -1 & 1 & -1 & 0 & 1 & 0 \\ 3 & 0 & 2 & 1 & 0 & 0 \\ 7 & 0 & 5 & 0 & 0 & 1 \end{array}\right) \mapsto$$

$$\left(\begin{array}{ccc|ccc} 1 & -1 & 1 & 0 & -1 & 0 \\ 3 & 0 & 2 & 1 & 0 & 0 \\ 7 & 0 & 5 & 0 & 0 & 1 \end{array}\right) \mapsto \left(\begin{array}{ccc|ccc} 1 & -1 & 1 & 0 & -1 & 0 \\ 0 & 3 & -1 & 1 & 3 & 0 \\ 0 & 7 & -2 & 0 & 7 & 1 \end{array}\right) \mapsto$$

$$\left(\begin{array}{ccc|ccc} 1 & -1 & 1 & 0 & -1 & 0 \\ 0 & 1 & -\frac{1}{3} & \frac{1}{3} & 1 & 0 \\ 0 & 7 & -2 & 0 & 7 & 1 \end{array}\right) \mapsto \left(\begin{array}{ccc|ccc} 1 & -1 & 1 & 0 & -1 & 0 \\ 0 & 1 & -\frac{1}{3} & \frac{1}{3} & 1 & 0 \\ 0 & 0 & \frac{1}{3} & -\frac{7}{3} & 0 & 1 \end{array}\right) \mapsto$$

$$\left(\begin{array}{ccc|ccc} 1 & -1 & 1 & 0 & -1 & 0 \\ 0 & 1 & -\frac{1}{3} & \frac{1}{3} & 1 & 0 \\ 0 & 0 & 1 & -7 & 0 & 3 \end{array}\right) \mapsto \left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 7 & -1 & -3 \\ 0 & 1 & 0 & -2 & 1 & 1 \\ 0 & 0 & 1 & -7 & 0 & 3 \end{array}\right) \mapsto$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 5 & 0 & -2 \\ 0 & 1 & 0 & -2 & 1 & 1 \\ 0 & 0 & 1 & -7 & 0 & 3 \end{array}\right)$$

Als Ergebnis erhalten wir also

$$A^{-1} \quad = \quad \begin{pmatrix} 3 & 0 & 2 \\ -1 & 1 & -1 \\ 7 & 0 & 5 \end{pmatrix}^{-1} \quad = \quad \begin{pmatrix} 5 & 0 & -2 \\ -2 & 1 & 1 \\ -7 & 0 & 3 \end{pmatrix}.$$

Offen bleibt hierbei die Frage, wie es zu interpretieren ist, wenn die Matrix $A$ zwar auf normierte ZSF, aber mit Zeilenrang $r < n$, gebracht werden kann. Dafür ist es notwendig, dass wir uns neben den Zeilen- auch mit ***Spaltenumformungen*** einer Matrix befassen. Unter einer ***elementaren*** Spaltenumformungen verstehen wir, dass die zu den Zeilenumformungen analogen Operationen auf die Spalten einer Matrix $A \in \mathscr{M}_{m\times n,K}$ angewendet werden, also im einzelnen

(i) die Multiplikation der $k$-ten Spalten einer Matrix mit einem Wert $\lambda$, wobei $\lambda \in K^\times$ und $k \in \{1,...,n\}$ ist
(ii) die Addition des $\lambda$-fachen der $k$-ten Spalte zur $\ell$-ten, mit $\lambda \in K$ und $k,\ell \in \{1,...,n\}$, $k \neq \ell$.

**(3.11) Lemma** Die Multiplikationen einer Matrix $A \in \mathscr{M}_{m\times n,K}$ mit den Transponierten von Elementarmatrizen von rechts bewirken elementare Spaltenumformungen. Genauer gilt:

(i) Die Matrix $A\,{}^t\mathrm{M}_{k,\lambda}$ entsteht aus der Matrix $A$ durch Multiplikation der $k$-ten Spalte mit $\lambda$.

(ii) Die Matrix $A\,{}^t\mathrm{A}_{k,\ell,\lambda}$ entsteht aus der Matrix $A$ durch Addition des $\lambda$-fachen der $k$-ten Spalte zur $\ell$-ten Spalte.

***Beweis:*** Wir beschränken uns auf den Beweis der Aussage (i). Nach der Rechenregel (iv) in (2.6) gilt $A\,{}^{\mathrm{t}}\mathrm{M}_{k,\lambda} = {}^{\mathrm{t}}({}^{\mathrm{t}}A)\,{}^{\mathrm{t}}\mathrm{M}_{k,\lambda} = {}^{\mathrm{t}}(\mathrm{M}_{k,\lambda}\,{}^{\mathrm{t}}A)$. Der Übergang ${}^{\mathrm{t}}A \mapsto \mathrm{M}_{k,\lambda}\,{}^{\mathrm{t}}A$ bewirkt nach (2.11) die Multiplikation der $k$-ten Zeile von ${}^{\mathrm{t}}A$ mit dem Wert $\lambda$. Für jedes $\ell$ ist die $\ell$-te Spalte von $A$ gleich der $\ell$-ten Zeile von ${}^{\mathrm{t}}A$, und entsprechend ist die $\ell$-te Spalte von $A\,{}^{\mathrm{t}}\mathrm{M}_{k,\lambda} = {}^{\mathrm{t}}(\mathrm{M}_{k,\lambda}\,{}^{\mathrm{t}}A)$ gleich der $\ell$-ten Zeile von $\mathrm{M}_{k,\lambda}\,{}^{\mathrm{t}}A$. Also stimmt die $\ell$-te Spalte von $A$ mit der $\ell$-ten Spalte von $A\,{}^{\mathrm{t}}\mathrm{M}_{k,\lambda}$ für $\ell \neq k$ überein, und für $\ell = k$ unterscheiden sie sich um den Faktor $\lambda$. □

Wir bemerken noch, dass mit jeder Matrix $A \in \mathrm{GL}_n(K)$ auch die Transponierte ${}^{\mathrm{t}}A$ invertierbar ist, mit $({}^{\mathrm{t}}A)^{-1} = {}^{\mathrm{t}}(A^{-1})$. Dies folgt direkt aus der Rechnung

$${}^{\mathrm{t}}A\,{}^{\mathrm{t}}(A^{-1}) \quad = \quad {}^{\mathrm{t}}(A^{-1}A) \quad = \quad {}^{\mathrm{t}}E^{(n)} \quad = \quad E^{(n)}$$

und einer analogen Rechnung, die ${}^{\mathrm{t}}(A^{-1})\,{}^{\mathrm{t}}A = E^{(n)}$ liefert.

**(3.12) Satz** Für jede Matrix $A \in \mathscr{M}_{m\times n,K}$ gibt es invertierbare Matrizen $T \in \mathrm{GL}_m(K)$ und $U \in \mathrm{GL}_n(K)$ und ein $r \in \{1,...,n\}$, so dass die Matrix $TAU$ die Blockgestalt

$$\begin{pmatrix} E^{(r)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \qquad \text{besitzt.}$$

***Beweis:*** Wir wissen bereits, dass eine Matrix $T \in \mathrm{GL}_m(K)$ existiert, so dass $B = TA$ in normierter ZSF vorliegt, mit gewissen Kennzahlen $r, j_1, ..., j_r$. Nach (3.11) genügt es nun zu zeigen, dass $B$ durch elementare Spaltenumformungen auf die angegebene Blockgestalt gebracht werden kann. Nach Definition der normierten ZSF befindet sich für $1 \leq k \leq r$ in der $j_k$-ten Spalten von $B$ jeweils der $k$-te Einheitsvektor $e_k \in K^m$. Nun führt man nacheinander für $1 \leq k \leq r$ die folgende Operation aus:

$$\text{Addition des } (-a_{k\ell})\text{-fachen der } j_k\text{-ten Spalte zur } \ell\text{-ten, für } j_k < \ell \leq n$$

Durch diese Operation werden die Einträge rechts von der Position $(k, j_k)$ zu Null, während alle übrigen Einträge der Matrix unverändert bleiben.

Nach Durchführung dieser Schritte enthält die modifizierte Matrix $B'$ in den Spalten $j_1, ..., j_r$ die Einheitsvektoren $e_1, ..., e_r$, alle übrigen Spalten sind Null. Nun vertauscht man die Spalten noch so, dass sich die Einheitsvektoren in den ersten $r$ Spalten befinden. Dann hat die Matrix die gewünschte Form. □

Auch hier lassen sich die Matrizen $T$ und $U$, die die angegebene Blockgestalt erzeugen, explizit berechnen. Zunächst wendet man die erforderlichen Zeilenumformungen statt auf $A$ auf die Blockmatrix $(A \;\; E^{(m)})$ an und erhält so eine Matrix der Form $(B \;\; T)$ mit $T \in \mathrm{GL}_m(K)$, wobei $B = TA$ sich in normierter ZSF befindet. Anschließend wendet man auf die linke Teilmatrix von $(B \;\; E^{(n)})$ Spaltenumformungen an, die $B$ auf die Blockgestalt bringen, und ***dieselben*** Spaltenumformungen auch auf die rechte Teilmatrix. Man erhält damit eine Matrix der Form $(C \;\; U)$ mit $U \in \mathrm{GL}_m(K)$, wobei $C = BU$ die angegebene Blockgestalt hat. Die Matrizen $T$ und $U$ haben die gewünschte Umformungeigenschaft.

Durch dieses Rechenverfahren wird auch eine Möglichkeit aufgezeigt, eine Matrix auf Invertierbarkeit zu testen.

**(3.13) Satz** Sei $A \in \mathscr{M}_{n,K}$ eine Matrix, die durch elementare Zeilenumformungen auf normierte ZSF mit Zeilenrang $r < n$ gebracht werden kann. Dann ist $A$ *nicht* invertierbar.

***Beweis:*** Nehmen wir an, dass die Voraussetzung erfüllt ist, die Matrix $A$ aber dennoch in $\mathrm{GL}_n(K)$ liegt. Nach (3.12) gibt es Matrizen $T, U \in \mathrm{GL}_n(K)$ mit

$$TAU \quad = \quad \begin{pmatrix} E^{(r)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

Mit $A$ wäre dann auch $TAU$ invertierbar. Aber eine Matrix mit Nullzeilen kann nicht invertierbar sein, denn für beliebiges $B \in \mathscr{M}_{r \times n,K}$ und $V \in \mathrm{GL}_n(K)$ gilt

$$\begin{pmatrix} B \\ \mathbf{0} \end{pmatrix} V \quad = \quad \begin{pmatrix} BV \\ \mathbf{0}V \end{pmatrix} \quad = \quad \begin{pmatrix} BV \\ \mathbf{0} \end{pmatrix} \quad ,$$

wobei die letzte Matrix offensichtlich *nicht* mit der Einheitsmatrix übereinstimmt. Der Widerspruch zeigt, dass die Annahme falsch war. □

Unser Rechenverfahren zur Bestimmung der Inversen einer Matrix $A$ liefert also zugleich ein Entscheidungskriterium für die Invertierbarkeit: Kommt bei der Rechnung eine normierte ZSF mit Zeilenrang $r < n$ heraus, dann existiert die Inverse von $A$ nicht.

# § 4. Vektorräume und lineare Abbildungen

***Inhaltsübersicht***

Der Begriff des $K$-Vektorrraums ist für die Lineare Algebra von zentraler Bedeutung. Es handelt sich dabei um eine Menge $V$ mit einer Verknüpfung, der *Vektoraddition*, und einer Abbildung $K \times V \to V$, der *skalaren Multiplikation*. Das wichtigstes Beispiel sind die Vektorräume der Form $K^n$, für die wir die Vektoraddition und skalare Multiplikation bereits in §2 definiert haben. Aber auch viele weitere Objekte der Mathematik besitzen eine Vektorraumstruktur, zum Beispiel die Menge $\mathscr{M}_{m\times n,K}$ der $(m \times n)$-Matrizen über einem Körper $K$, oder die Menge der $K$-wertigen Abbildungen auf einer Menge $X$.

Eine Abbildung $V \to W$ zwischen Vektorräumen bezeichnet man als *lineare Abbildung*, wenn sie „verträglich" mit der Vektoraddition und der skalaren Multiplikation der beiden Vektorräume ist. Beispielsweise ist die Matrix-Vektor-Multiplikation aus §2 ein Beispiel für eine lineare Abbildung $K^n \to K^m$. Lineare Abbildungen sind ein wichtiges Hilfsmittel, um Beziehungen zwischen verschiedenen Vektorräumen herzustellen. Oft haben sie eine geometrische Interpretation; zum Beispiel ist die Spiegelung im $\mathbb{R}^2$ an einer Gerade durch den Ursprung $(0,0)$ eine lineare Abbildung, ebenso jede Drehung um $(0,0)$. Auch bei der Untersuchung von linearen Gleichungssystemen spielen die linearen Abbildungen eine wichtige Rolle, wie wir im weiteren Verlauf noch sehen werden.

***Wichtige Definitionen und Sätze***

- $K$-Vektorraum, Vektoraddition, skalare Multiplikation, Vektor, Nullvektor
- Ist $X$ eine Menge und $V$ ein $K$-Vektorraum, dann bilden die Abbildungen $X \to V$ ebenfalls einen $K$-Vektorraum.
- lineare Abbildung / Homomorphismus von $K$-Vektorräumen
- Mono-, Epi-, Iso-, Endo- und Automorphismen

**(4.1) Definition** Sei $K$ ein Körper. Ein $K$-***Vektorraum*** ist ein Tripel $(V,+,\cdot)$ bestehend aus einer nichtleeren Menge $V$ und Abbildungen $+ : V\times V \to V$ und $\cdot : K\times V \to V$ genannt ***Vektoraddition*** und ***skalare Multplikation***, so dass folgende Bedingungen erfüllt sind.

(i) Das Paar $(V,+)$ ist eine abelsche Gruppe.

(ii) Für alle $v,w \in V$ und $\lambda,\mu \in K$ gelten die Rechenregeln

(a) $(\lambda+\mu)\cdot v = (\lambda\cdot v)+(\mu\cdot v)$

(b) $\lambda\cdot(v+w) = (\lambda\cdot v)+(\lambda\cdot w)$

(c) $(\lambda\mu)\cdot v = \lambda\cdot(\mu\cdot v)$

(d) $1_K \cdot v = v$

Die Elemente der Menge $V$ werden ***Vektoren*** genannt.

Bei der skalaren Multiplikation wird häufig auf das Abbildungssymbol $\cdot$ verzichtet. Das Neutralelement der Gruppe $(V,+)$ bezeichnet man als den ***Nullvektor*** $0_V$ des Vektorraums. Das Inverse eines Vektors $v \in V$ bezüglich der Vektoraddition bezeichnet man mit $-v$ und verwendet $v-w$ als abkürzende Schreibweise für $v+(-w)$. Per Konvention bindet die skalare Multiplikation stärker als die Vektoraddition, d.h. der Ausdruck $\lambda v + w$ ist gleichbedeutend mit $(\lambda v) + w$ für $\lambda \in K$, $v, w \in V$.

**(4.2) Proposition** Sei $(V, +, \cdot)$ ein $K$-Vektorraum. Dann gilt für alle $\lambda \in K$ und $v \in V$ die Äquivalenz

$$\lambda v = 0_V \quad \Leftrightarrow \quad \lambda = 0_K \text{ oder } v = 0_V \quad ,$$

außerdem $(-1_K)v = -v$ für alle $v \in V$.

***Beweis:*** Zunächst beweisen wir die Äquivalenz. „$\Leftarrow$“ Ist $\lambda = 0_K$, dann gilt $\lambda v = 0_K v = (0_K + 0_K)v = 0_K v + 0_K v = \lambda v + \lambda v$. Addition von $-\lambda v$ auf beiden Seiten dieser Gleichung liefert

$$\lambda v + (-\lambda v) = \lambda v + \lambda v + (-\lambda v) \quad \Leftrightarrow \quad 0_V = \lambda v + 0_V \quad \Leftrightarrow \quad 0_V = \lambda v.$$

Setzen wir nun $v = 0_V$ voraus, dann erhalten wir $\lambda v = \lambda 0_V = \lambda(0_V + 0_V) = \lambda 0_V + \lambda 0_V = \lambda v + \lambda v$. Wieder führt die Addition von $-\lambda v$ auf beiden Seiten zum gewünschten Ergebnis.

„$\Rightarrow$“ Setzen wir $\lambda v = 0_V$ voraus, und nehmen wir an, es ist $\lambda \neq 0_K$. Dann gilt

$$v \quad = \quad 1_K v \quad = \quad (\lambda^{-1}\lambda)v \quad = \quad \lambda^{-1}(\lambda v) \quad = \quad \lambda^{-1}0_V \quad = \quad 0_V \quad ,$$

wobei im letzten Schritt die bereits bewiesene Rechenregel $\mu 0_V = 0_V$ für alle $\mu \in K$ verwendet wurde. Beweisen wir nun noch die Gleichung $(-1_K)v = -v$. Es gilt $v + (-1_K)v = 1_K v + (-1_K)v = (1_K + (-1_K))v = 0_K v = 0_V$. Addition von $-v$ auf beiden Seiten liefert

$$v + (-1_K)v + (-v) = 0_V + (-v) \quad \Leftrightarrow \quad v + (-v) + (-1_K)v = -v \quad \Leftrightarrow$$

$$0_V + (-1_K)v = -v \quad \Leftrightarrow \quad (-1_K)v = -v \qquad \square$$

$\square$

**(4.3) Proposition** Sei $K$ ein Körper.
Die folgenden Strukturen sind Beispiele für $K$-Vektorräume.

(i) für jedes $n \in \mathbb{N}$ das Tripel $(K^n, +, \cdot)$ mit den Abbildungen

$$+ : K^n \times K^n \to K^n \quad , \quad ((a_1, ..., a_n), (b_1, ..., b_n)) \mapsto (a_1 + b_1, ..., a_n + b_n)$$

und

$$\cdot : K \times K^n \to K^n \quad , \quad (\lambda, (a_1, ..., a_n)) \mapsto (\lambda a_1, ..., \lambda a_n)$$

(ii) für alle $m, n \in \mathbb{N}$ das Tripel $(\mathscr{M}_{m\times n,K}, +, \cdot)$, wobei $+ : \mathscr{M}_{m\times n,K} \times \mathscr{M}_{m\times n,K} \to \mathscr{M}_{m\times n,K}$ für die Addition und $\cdot : K \times \mathscr{M}_{m\times n,K} \to \mathscr{M}_{m\times n,K}$ für die skalare Multiplikation von Matrizen steht

(iii) das Tripel $(K, +, \cdot)$, wobei $+$ die Addition und $\cdot$ die Multiplikation von $K$ bezeichnet

(iv) das Tripel $(\{0_V\}, +, \cdot)$, wobei $+ : \{0_V\} \times \{0_V\} \to \{0_V\}$ durch $(0_V, 0_V) \mapsto 0_V$ und $\cdot : K \times \{0_V\} \to \{0_V\}$ durch $(\lambda, 0_V) \mapsto 0_V$ für alle $\lambda \in K$ gegeben ist

***Beweis:*** Punkt (i) kann als Spezialfall von (ii) angesehen werden, da wir in §1 festgelegt haben, dass die Elemente aus $K^n$ Matrizen mit einer einzigen Spalte sind. Beweisen wir nun die Vektorraum-Eigenschaft von Beispiel (ii). Nach (3.8) (i) ist $(\mathscr{M}_{n,K},+)$ eine abelsche Gruppe, mit der Nullmatrix als Neutralelement. Zu überprüfen sind noch die Gleichungen $(\lambda+\mu)A = \lambda A + \mu A$, $\lambda(A+B) = \lambda A + \lambda B$, $(\lambda\mu)A = \lambda(\mu A)$ und $1_K A = A$ für beliebige Körperelemente $\lambda,\mu \in K$ und beliebige Matrizen $A,B \in \mathscr{M}_{m\times n,K}$. Der Beweis dieser Rechenregeln erfolgt ähnlich wie in (2.6) und wird hier nicht ausgeführt.

Kommen wir nun zu Beispiel (iii). Dass $(K,+)$ eine abelsche Gruppe mit $0_K$ als Neutralelement ist, ergibt sich direkt aus (3.7) (i). Seien nun $\lambda,\mu \in K$ und $a,b \in K$ vorgegeben, wobei $\lambda,\mu$ die Rolle der Körperelemente und $a,b$ die Rolle der „Vektoren" übernehmen. Die Rechenregel (ii)(b) in (4.1) gegeben durch $\lambda(a+b) = \lambda a + \lambda b$ entspricht dem Distributivgesetz im Körper $K$. Die Regel (ii)(a) erhält man durch Kommutativ- und Distributivgesetz, denn es gilt

$$(\lambda+\mu)a \quad = \quad a(\lambda+\mu) \quad = \quad a\lambda + \mu\lambda \quad = \quad \lambda a + \mu a.$$

Nach dem Assoziativgesetz der Multiplikation gilt $(\lambda\mu)a = \lambda(\mu a)$, also ist auch (ii)(c) gültig. Nach Definition des Einselements in einem Körper gilt schließlich noch $1_K a = a$, also Regel (ii)(d).

Betrachten wir nun Beispiel (iv) und überprüfen zunächst, dass $(\{0_V\},+)$ eine abelsche Gruppe ist. Assoziativ- und Kommutativgesetz sind wegen $0_V + 0_V = 0_V + 0_V$ und $(0_V + 0_V) + 0_V = 0_V + 0_V = 0_V + (0_V + 0_V)$ erfüllt. Wegen $a + 0_V = 0_V = a$ und $0_V + a = 0_V = a$ für alle $a \in \{0_V\}$ ist $0_V$ das Neutralelement der Gruppe. Die Gleichung $0_V + 0_V = 0_V + 0_V = 0_V$ zeigt, dass $0_V$ sein eigenes Inverses ist. Damit ist der Nachweis der Gruppeneigenschaften abgeschlossen. Zur Überprüfung der Rechenregeln (ii)(a) bis (d) seien $\lambda,\mu \in K$ vorgegeben. Es gilt

$$(\lambda+\mu)0_V \quad = \quad 0_V \quad = \quad 0_V + 0_V \quad = \quad \lambda 0_V + \mu 0_V \quad ,$$

also ist (ii)(a) erfüllt. Die Regel (ii)(b) erhält man durch $\lambda(0_V + 0_V) = \lambda 0_V = 0_V = 0_V + 0_V = \lambda 0_V + \mu 0_V$, und (ii)(c) durch die Rechnung $(\lambda\mu)\cdot 0_V = 0_V = \lambda \cdot 0_V = \lambda\cdot(\mu\cdot 0_V)$. Schließlich ist wegen $1_K 0_V = 0_V$ auch (ii)(d) erfüllt. □

Nach (4.3) (i) ist $\mathbb{C}^n$ für jedes $n \in \mathbb{N}$ mit der komponentenweisen Addition $+$ und skalaren Multiplikation $\cdot$ ein $\mathbb{C}$-Vektorraum. Wir bemerken noch, dass die skalare Multiplikation $\cdot : \mathbb{C}\times\mathbb{C}^n \to \mathbb{C}^n$ zu einer Abbildung $\cdot_{\mathbb{R}} : \mathbb{R}\times\mathbb{C}^n \to \mathbb{C}^n$ eingeschränkt werden kann. Das Tripel $(\mathbb{C}^n,+,\cdot_{\mathbb{R}})$ ist dann ein $\mathbb{R}$-Vektorraum, denn wenn die Rechenregeln (ii)(a) bis (d) für alle $v,w \in \mathbb{C}^n$ und $\lambda,\mu \in \mathbb{C}$ gültig sind, dann gelten sie erst recht für alle $v,w \in \mathbb{C}^n$ und $\lambda,\mu \in \mathbb{R}$. Auch jeder andere $\mathbb{C}$-Vektorraum kann durch Einschränkung der skalaren Multiplikation auf diese Weise zu einem $\mathbb{R}$-Vektorraum gemacht werden.

Wir betrachten noch ein weiteres Beispiel für Vektoräume, das auf den ersten Blick etwas exotisch erscheinen mag.

**(4.4) Proposition** Sei $K$ ein Körper, $(V,+_V,\cdot_V)$ ein $K$-Vektorraum und $X$ eine beliebige Menge. Mit $A = \mathrm{Abb}(X,V)$ bezeichnen wir die Menge der Abbildungen $X \to V$. Außerdem

(i) definieren wir auf $A$ eine Verknüpfung $\oplus$, indem wir jedem Paar $(f,g)$ von Elementen aus $A$ die Abbildung gegeben durch $(f\oplus g)(x) = f(x) +_V g(x)$ für alle $x \in X$ zuordnen

(ii) definieren wir eine Abbildung $\odot : K \times A \to A$, indem wir jedem Paar $(\lambda, f)$ die Abbildung gegeben durch $(\lambda f)(x) = \lambda \cdot_V f(x)$ für alle $x \in X$ zuordnen.

Dann ist $(A,\oplus,\odot)$ ein $K$-Vektorraum.

***Beweis:*** Zunächst zeigen wir, dass $(A,\oplus)$ eine abelsche Gruppe ist. Seien $f,g,h \in A$ vorgegeben. Für ein beliebiges $x \in X$ gilt

$$((f \oplus g)\oplus h)(x) \quad = \quad (f \oplus g)(x) +_V h(x) \quad = \quad (f(x) +_V g(x)) +_V h(x) \quad = \quad f(x) +_V (g(x) +_V h(x))$$
$$= \quad f(x) +_V (g \oplus h)(x) \quad = \quad (f \oplus (g \oplus h))(x) \quad ,$$

wobei wir im dritten Schritt die Assoziativität der Verknüpfung $+_V$ auf $V$ verwendet haben. Weil $x \in X$ beliebig vorgegeben war, folgt daraus die Gleichheit der Abbildungen, also $(f \oplus g) \oplus h = f \oplus (g \oplus h)$. Ebenso folgt aus $(f \oplus g)(x) = f(x) +_V g(x) = g(x) +_V f(x) = (g \oplus f)(x)$ für $f,g \in V$ und $x \in X$ die Gleichung $f \oplus g = g \oplus f$ für alle $f,g \in A$ und somit die Kommutativität der Verknüpfung $\oplus$ auf $A$.

Sei nun $0_A \in A$ die Abbildung gegeben durch $0_A(x) = 0_V$ für alle $x \in X$. Für alle $f \in A$ und $x \in X$ gilt dann

$$(f \oplus 0_A)(x) \quad = \quad f(x) +_V 0_A(x) \quad = \quad f(x) +_V 0_V \quad = \quad f(x) \quad ,$$

also $f \oplus 0_V = f$, und auf Grund der bereits bewiesenen Kommutativität auch $0_V \oplus f = f$. Schließlich definieren wir für jedes $f \in A$ die Abbildung $-f \in A$ durch $(-f)(x) = -f(x)$ für alle $x \in X$. Für jedes $x \in X$ gilt dann

$$(f \oplus (-f))(x) \quad = \quad f(x) +_V (-f)(x) \quad = \quad f(x) +_V (-f(x)) \quad = \quad 0_V \quad = \quad 0_A(x)$$

und somit $f \oplus (-f) = 0_A$. Auf Grund der Kommutativität von $\oplus$ gilt damit auch $(-f) \oplus f = 0_A$. Damit ist insgesamt nachgewiesen, dass es sich bei $(A,\oplus)$ um eine abelsche Gruppe handelt.

Nun müssen wir noch die Gleichungen (a) bis (d) in (4.1) (ii) überprüfen. Seien dazu $f,g \in A$ und $\lambda,\mu \in K$ vorgegeben. Für alle $x \in X$ gilt

$$((\lambda+\mu)\odot f)(x) \quad = \quad (\lambda+\mu)\cdot_V f(x) \quad = \quad \lambda \cdot_V f(x) +_V \mu \cdot_V f(x) \quad =$$

$$(\lambda \odot f)(x) +_V (\mu \odot f)(x) \quad = \quad (\lambda \odot f \oplus \mu \odot f)(x)$$

und somit $(\lambda+\mu)\odot f = \lambda \odot f \oplus \mu \odot f$ Ebenso gilt für alle $x \in X$ die Gleichung

$$(\lambda \odot (f \oplus g))(x) \quad = \quad \lambda \cdot_V ((f \oplus g)(x)) \quad = \quad \lambda \cdot_V (f(x) +_V g(x)) \quad = \quad \lambda \cdot_V f(x) +_V \lambda \cdot_V g(x)$$
$$= \quad (\lambda \odot f)(x) +_V (\lambda \odot g)(x) \quad = \quad (\lambda \odot f \oplus \lambda \odot g)(x) \quad ,$$

also $\lambda \odot (f \oplus g) = \lambda \odot f \oplus \lambda \odot g$. Zum Nachweis von Bedingung (c) rechnen wir nach, dass für alle $x \in X$ die Gleichung

$$((\lambda\mu)\odot f)(x) \quad = \quad (\lambda\mu)\cdot_V f(x) \quad = \quad \lambda \cdot_V (\mu \cdot_V f(x)) \quad = \quad \lambda \cdot_V (\mu \odot f)(x) \quad = \quad (\lambda \odot (\mu \odot f))(x)$$

und somit $(\lambda\mu)\odot f = \lambda \odot (\mu \odot f)$ gilt; dabei haben wir im zweiten Schritt die Rechenregel (ii)(c) im Vektorraum $(V,+_V,\cdot_V)$ verwendet. Schließlich gilt für alle $f \in A$ und $x \in X$ noch

$$(1_K \odot f)(x) \quad = \quad 1_K \cdot_V f(x) \quad = \quad f(x) \quad ,$$

also $1_K \odot f = f$ für alle $f \in A$. Damit haben wir auch die Regel (ii)(d) auf die entsprechende Regel in $(V,+_V,\cdot_V)$ zurückgeführt. $\square$

**(4.5) Definition** Seien $(V, +_V, \cdot_V)$ und $(W, +_W, \cdot_W)$ $K$-Vektorräume. Eine Abbildung $\phi : V \to W$ heißt ***lineare Abbildung*** oder ***Homomorphismus*** von $K$-Vektorräumen, wenn folgende Bedingungen erfüllt sind.

(i) $\phi(v +_V w) = \phi(v) +_W \phi(w)$ für alle $v, w \in V$

(ii) $\phi(\lambda \cdot_V v) = \lambda \cdot_W \phi(v)$ für alle $v \in V$ und $\lambda \in K$

**(4.6) Lemma** Ist $\phi : V \to W$ eine lineare Abbidlung wie in der Definition angegeben. Dann gilt $\phi(0_V) = 0_W$, $\phi(-v) = -\phi(v)$ und $\phi(v - w) = \phi(v) - \phi(w)$ für alle $v, w \in V$.

***Beweis:*** Die erste Gleichung erhält man mit Hilfe der Eigenschaft (ii) von linearen Abbildungen durch $\phi(0_V) = \phi(0_K \cdot_V 0_V) = 0_K \cdot_W \phi(0_V) = 0_W$. Die zweite ergibt sich durch die Rechnung

$$\phi(-v) \quad = \quad \phi((-1_K) \cdot_V v) \quad = \quad (-1)_K \cdot_W \phi(v) \quad = \quad -\phi(v).$$

Die dritte Gleichung schließlich erhält man durch

$$\phi(v - w) \quad = \quad \phi(v +_V (-w)) \quad = \quad \phi(v) +_W \phi(-w) \quad = \quad \phi(v) +_W (-\phi(w)) \quad = \quad \phi(v) - \phi(w). \qquad \square$$

**(4.7) Proposition** Seien $m, n \in \mathbb{N}$ und $A \in \mathscr{M}_{m \times n, K}$. Dann ist durch $\phi_A : K^n \to K^m$, $v \mapsto Av$ eine lineare Abbildung gegeben.

***Beweis:*** Seien $v, w \in K^n$ und $\lambda \in K$ vorgegeben. Auf Grund der Rechenregeln aus (1.12) für das Matrix-Vektor-Produkt gilt

$$\phi_A(v + w) = A(v + w) = Av + Aw = \phi_A(v) + \phi_A(w)$$

und $\phi_A(\lambda v) = A(\lambda v) = \lambda Av = \lambda \phi_A(v)$. $\square$

Sei $V$ ein $K$-Vektorraum, $n \in \mathbb{N}$, und seien $v_1, ..., v_n \in V$ beliebige Vektoren. Wie bei Körpern verwenden wir den Ausdruck $\sum_{k=1}^n v_k$ als Kurzschreibweise für die Summe $v_1 + ... + v_n$ in $V$.

**(4.8) Lemma** Seien $V, W$ $K$-Vektorräume, $n \in \mathbb{N}$, außerdem $v_1, ..., v_n \in V$ und $\phi : V \to W$ eine lineare Abbildung. Dann gilt

$$\phi\left(\sum_{k=1}^n v_k\right) \quad = \quad \sum_{k=1}^n \phi(v_k).$$

***Beweis:*** Wir beweisen die Aussage durch vollständige Induktion über $n$. Im Fall $n = 1$ lautet die Behauptung nur $\phi(v_1) = \phi(v_1)$ für alle $v_1 \in V$ und ist offensichtlich erfüllt. Sei nun $n \in \mathbb{N}$, und setzen wir die Aussage für dieses $n$ voraus. Seien $v_1, \ldots, v_{n+1} \in V$ beliebige Vektoren. Dann erhalten wir

$$\phi\left(\sum_{k=1}^{n+1} v_k\right) = \phi\left(\sum_{k=1}^{n} v_k + v_{n+1}\right) = \phi\left(\sum_{k=1}^{n} v_k\right) + \phi(v_{n+1}) \overset{(*)}{=}$$

$$\sum_{k=1}^{n} \phi(v_k) + \phi(v_{n+1}) = \sum_{k=1}^{n+1} \phi(v_k) \quad ,$$

wobei an der Stelle (*) die Induktionsvoraussetzung angewendet wurde. □

**(4.9) Proposition** Die Linearformen auf dem $K$-Vektorraum $K^n$ sind genau die linearen Abbildungen $K^n \to K$.

***Beweis:*** Dass jede Linearform $\phi$ auf $K^n$ eine lineare Abbildung $K^n \to K$ ist, folgt direkt aus den Rechenregeln für Linearformen aus (1.4). Sei nun $\phi : K^n \to K$ eine lineare Abbildung. Für $1 \leq k \leq n$ sei $e_k \in K^n$ jeweils der $k$-te Einheitsvektor und $a_k = \phi(e_k)$. Ist nun $v = (\lambda_1, \ldots, \lambda_n) \in K^n$ ein beliebiger Vektor, dann gilt $v = \sum_{k=1}^{n} \lambda_k e_k$. Durch Anwendung von (4.8) erhalten wir

$$\phi(v) = \phi\left(\sum_{k=1}^{n} \lambda_k e_k\right) = \sum_{k=1}^{n} \phi(\lambda_k e_k) = \sum_{k=1}^{n} \lambda_k \phi(e_k) = \sum_{k=1}^{n} \lambda_k a_k.$$

Dies zeigt, dass $\phi$ eine Linearform im Sinne von (1.1) ist. □

Seien $X, Y$ Mengen, sei $V$ ein $K$-Vektorraum, und seien $A = \mathrm{Abb}(X, V)$ und $B = \mathrm{Abb}(Y, V)$. Nach (4.4) gibt es Verknüpfungen $\oplus_A$, $\oplus_B$ auf $A, B$ und Abbildungen $\odot_A : K \times A \to A$ und $\odot_B : K \times B \to B$, so dass $(A, \oplus_A, \odot_A)$ und $(B, \oplus_B, \odot_B)$ zu $K$-Vektorräumen werden.

**(4.10) Proposition** Sei $u : X \to Y$ eine Abbildung. Dann ist durch $\phi_u : B \to A$, $f \mapsto f \circ u$ eine lineare Abbildung zwischen $(A, \oplus_A, \odot_A)$ und $(B, \oplus_B, \odot_B)$ gegeben.

***Beweis:*** Seien $f, g \in B$ und $\lambda \in K$ vorgegeben. Zu zeigen ist $\phi_u(f \oplus_B g) = \phi_u(f) \oplus_A \phi_u(g)$ und $\phi_u(\lambda \odot_B f) = \lambda \odot_A \phi_u(f)$. Zum Beweis der ersten Gleichung sei $x \in X$ ein beliebiges Element. Dann gilt

$$\phi_u(f \oplus_B g)(x) = ((f \oplus_B g) \circ u)(x) = (f \oplus_B g)(u(x)) = f(u(x)) +_V g(u(x))$$

$$= (f \circ u)(x) +_V (g \circ u)(x) = \phi_u(f)(x) +_V \phi_u(g)(x) = (\phi_u(f) \oplus_A \phi_u(g))(x).$$

Weil $x \in X$ beliebig vorgegeben war, folgt daraus $\phi_u(f \oplus_B g) = \phi_u(f) \oplus_A \phi_u(g)$. Zum Beweis der zweiten Gleichung betrachten wir wiederum ein beliebiges $x \in X$. Es gilt

$$\phi_u(\lambda \odot_B f)(x) = ((\lambda \odot_B f) \circ u)(x) = (\lambda \odot_B f)(u(x)) = \lambda \cdot_V f(u(x))$$

$$= \lambda \cdot_V (f \circ u)(x) = \lambda \cdot_V \phi_u(f)(x) = (\lambda \odot_A \phi_u(f))(x)$$

und somit $\phi_u(\lambda \odot_B f) = \lambda \odot_A \phi_u(f)$. □

**(4.11) Proposition** Seien $U, V, W$ drei $K$-Vektorräume und $\phi : U \to V$, $\psi : V \to W$ lineare Abbildungen. Dann ist $\psi \circ \phi$ eine lineare Abbildung von $U$ nach $W$. Ist $\phi$ bijektiv, dann ist $\phi^{-1}$ eine lineare Abbidlung von $V$ nach $U$.

***Beweis:*** Wir überprüfen die Linearität der Abbildung $\psi \circ \phi$. Seien dazu $v, w \in U$ und $\lambda \in K$ vorgegeben. Dann gilt

$$\begin{aligned}(\psi \circ \phi)(v +_U w) &= \psi(\phi(v +_U w)) = \psi(\phi(v) +_V \phi(w)) = \psi(\phi(v)) +_W \psi(\phi(w)) \\ &= (\psi \circ \phi)(v) +_W (\psi \circ \phi)(w)\end{aligned}$$

und $(\psi \circ \phi)(\lambda v) = \psi(\phi(\lambda v)) = \psi(\lambda \phi(v)) = \lambda \psi(\phi(v)) = \lambda(\psi \circ \phi)(v)$. Setzen wir nun voraus, dass $\phi$ bijektiv ist. Um zu zeigen, dass die Abbildung $\phi^{-1}$ linear ist, seien $v, w \in V$ und $\lambda \in K$ vorgegeben. Sei $v' = \phi^{-1}(v)$ und $w' = \phi^{-1}(w)$. Unter Verwendung der Linearität von $\phi$ erhalten wir

$$\begin{aligned}\phi^{-1}(v) +_U \phi^{-1}(w) &= v' +_U w' = \mathrm{id}_U(v' +_U w') = (\phi^{-1} \circ \phi)(v' +_U w') \\ &= \phi^{-1}(\phi(v' +_U w')) = \phi^{-1}(\phi(v') +_V \phi(w')) = \phi^{-1}(v +_V w).\end{aligned}$$

Ebenso gilt $\lambda\phi^{-1}(v) = \lambda v' = \mathrm{id}_U(\lambda v') = (\phi^{-1} \circ \phi)(\lambda v') = \phi^{-1}(\phi(\lambda v')) = \phi^{-1}(\lambda \phi(v')) = \phi^{-1}(\lambda v)$. □

**(4.12) Definition** Eine lineare Abbildung $\phi : V \to W$ heißt

(i) ***Monomorphismus*** (von $K$-Vektorräumen), wenn $\phi$ injektiv ist,
(ii) ***Epimorphismus***, wenn $\phi$ surjektiv ist,
(iii) ***Isomorphismus***, wenn $\phi$ bijektiv ist.

Eine lineare Abbildung $\phi : V \to V$ bezeichnet man als ***Endomorphismus*** von $V$, und ist sie außerdem bijektiv, dann spricht man von einem ***Automorphismus***. Zwei $K$-Vektorräume $V, W$ werden ***isomorph*** genannt, wenn ein Isomorphismus $\phi : V \to W$ existiert.

**(4.13) Folgerung** Die Menge der Automorphismen eines $K$-Vektorraums $V$ ist mit der Komposition $\circ$ von Abbildungen als Verknüpfung eine Gruppe. Man bezeichnet sie mit $\mathrm{GL}(V)$ und nennt sie die ***allgemeine lineare Gruppe*** des Vektorraums $V$.

***Beweis:*** Die vorhergehende Proposition zeigt, dass mit $\phi, \psi : V \to V$ auch die Abbildungen $\psi \circ \phi$ und $\phi^{-1}$ Automorphismen des Vektorraums $V$ sind. Die Assoziativität ergibt sich aus der allgemeinen Regel $h \circ (g \circ f) = (h \circ g) \circ f$ für beliebige Abbildungen zwischen Mengen. Die identische Abbildung $\mathrm{id}_V$ besitzt die definierende Eigenschaft des Neutralelements (es gilt $\phi \circ \mathrm{id}_V = \mathrm{id}_V \circ \phi = \phi$ für jeden Automorphismus $\phi$ von $V$), und die Umkehrabbildung $\phi^{-1}$ von $\phi$ erfüllt die Bedingung $\phi^{-1} \circ \phi = \phi \circ \phi^{-1} = \mathrm{id}_V$ für das inverse Element. □

# § 5. Unterverektorräume

**_Inhaltsübersicht_**

Unter einem *Untervektorraum* versteht man eine Teilmenge $U$ eines Vektorraums $V$, unter der Vektoraddition und der skalaren Multiplikation von $V$ abgeschlossen und bezüglich dieser Operationen selbst ein Vektorraum ist. Beispiele für Untervektorräume des $\mathbb{R}^3$ sind Geraden und Ebenen, die durch die Koordinatenursprung $0_{\mathbb{R}^3} = (0,0,0)$ verlaufen. Geraden und Ebenen, die dies nicht erfüllen, sind „verschobene" Untervektorräume, die als *affine Unterräume* bezeichnet werden.

In diesem Kapitel definieren wir die Untervektorräume für beliebige $K$-Vektorräume $V$ und behandeln auch dazu eine Reihe von Beispielen. Unter anderem sind Lösungsmengen homogener LGS in $n$ Unbekannten über einem Körper $K$ Untervektorräume des $K^n$, bei inhomogenen LGS erhalten wir affine Unterräume des $K^n$. Außerdem sehen wir uns an, durch welche Operationen man aus vorgegebenen Untervektorräumen neue Untervektorräume gewinnt. Hier ist als wichtigste Operation die *Summe* von Untervektorräumen zu nennen. Auch lineare Abbildungen können zur Definition von Untervektorräumen genutzt werden. Hierbei greifen wir auf die aus dem ersten Semester bekannten Konzepte der *Bild-* und *Urbildmengen* von Abbildungen zurück.

**_Wichtige Definitionen_**

- Untervektorraum
- Summe und direkte Summer zweier Untervektorräume
- Kern und Bild einer linearen Abbildung
- affiner Unterraum

**(5.1) Definition** Sei $V$ ein $K$-Vektorraum. Eine Teilmenge $U \subseteq V$ wird ***Untervektorraum*** von $V$ genannt, wenn folgende Bedingungen erfüllt sind.

(i) $0_V \in U$

(ii) $v + w \in U$ für alle $v, w \in U$

(iii) $\lambda v \in U$ für alle $\lambda \in K$ und $v \in U$

**(5.2) Satz** Sei $(V, +, \cdot)$ ein $K$-Vektorraum und $U \subseteq V$ ein Untervektorraum von $V$. Definieren wir Abbildungen $+_U : U \times U \to V$ und $\cdot_U : K \times U \to V$ durch

$$v +_U w \quad = \quad v + w \qquad \text{und} \qquad \lambda \cdot_U v \quad = \quad \lambda \cdot v \qquad \text{für} \quad v, w \in U \text{ und } \lambda \in K,$$

dann ist durch $(U, +_U, \cdot_U)$ ein $K$-Vektorraum gegeben.

***Beweis:*** Weil $U$ ein Untervektorraum ist, gilt $v + w \in U$ für alle $v, w \in U$, also auch $v +_U w = v + w \in U$. Dies zeigt, dass $+_U$ eine Abbildung $U \times U \to U$, also eine Verknüpfung auf $U$ gegeben ist. Ebenso gilt $\lambda \cdot_U v = \lambda \cdot v \in U$ für alle $v \in U$ und $\lambda \in K$. Somit ist $\cdot_U$ eine Abbildung $\cdot_U : K \times U \to U$. Wir müssen nun überprüfen, dass $(U, +_U, \cdot_U)$ die Vektorraum-Bedingungen aus (4.1) erfüllt.

Zunächst überprüfen wir, dass $(U, +_U)$ eine abelsche Gruppe ist. Für alle $u, v, w \in U$ gilt $(u +_U v) +_U w = (u + v) + w = u + (v + w) = u +_U (v +_U w)$, also ist das Assoziativgesetz erfüllt. Nach Voraussetzung liegt $0_V$ in $U$, und für alle $v \in U$ gilt $u +_U 0_V = u + 0_V = u$ und $0_V +_U u = 0_V + u = u$. Damit besitzt $0_V$ die Eigenschaften des Neutralelements in $(U, +_U)$. Sei nun $v \in U$ vorgegeben. Nach Voraussetzung liegt der Vektor $-v = (-1)v$ in $U$. Außerdem gilt $v +_U (-v) = v + (-v) = 0_V$ und $(-v) +_U v = (-v) + v = 0_V$. Also besitzt jedes $v \in U$ in $(U, +_U)$ ein Inverses, nämlich $-v$. Ingesamt bedeutet dies, dass $(U, +_U)$ eine Gruppe ist. Das Kommutativgesetz erhält man durch die Rechnung $v +_U w = v + w = w + v = w +_U v$ für alle $v, w \in U$.

Nun müssen wir noch die Eigenschaften (ii) (a)-(d) aus (4.1) überprüfen. Seien dazu $v, w \in U$ und $\lambda, \mu \in K$ vorgegeben. Es gilt $(\lambda + \mu) \cdot_U v = (\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v = \lambda \cdot_U v +_U \mu \cdot_U v$. Ebenso erhält man $\lambda \cdot_U (v +_U w) = \lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w = \lambda \cdot_U v +_U \lambda \cdot_U w$. Weiter gilt $(\lambda\mu) \cdot_U v = (\lambda\mu) \cdot v = \lambda \cdot (\mu \cdot v) = \lambda \cdot_U (\mu \cdot_U v)$ und schließlich $1_K \cdot_U v = 1_K \cdot v = v$. □

Man sieht, dass das Nachrechnen der Vektorraum-Axiome in $(U, +_U, \cdot_U)$ eine ziemliche Routineangelegenheit war: Überall wurden nur die Symbole $+_U$ und $\cdot_U$ durch $+$ und $\cdot$ ersetzt und anschließend verwendet, dass die Axiome im Vektorraum $V$ gültig sind.

Folgende konkrete Beispiele lassen sich für Untervektorräume angeben.

(i) Ist $V$ ein beliebiger $K$-Vektorraum, dann sind $\{0_V\}$ und $V$ Untervektorräume von $V$.

(ii) Für jedes $v \in V$ ist $\mathrm{lin}(v) = \{\lambda v \mid \lambda \in K\}$ ein Untervektorraum. Im Fall $v \neq 0_V$ bezeichnet man ihn als ***lineare Gerade***. Für beliebige $v, w$ ist auch durch

$$\mathrm{lin}(v, w) \quad = \quad \{\lambda v + \mu w \mid \lambda, \mu \in K\}$$

ein Untervektorraum gegeben. Ist $v \neq 0_V$ und $w \notin \mathrm{lin}(v)$ (oder äquivalent, $v \notin \mathrm{lin}(w)$ und $w \notin \mathrm{lin}(v)$), dann nennt man $\mathrm{lin}(v, w)$ eine ***lineare Ebene***.

(iii) Die Lösungsmenge $\mathscr{L} \subseteq K^n$ eines homogenen linearen Gleichungssystems bestehend aus $m$ Gleichungen in $n$ Unbekannten ist ein Untervektorraum von $K^n$. Denn der Nullvektor $0_{K^n}$ ist immer in $\mathscr{L}$ enthalten, und die Bedingung $v, w \in \mathscr{L}, \lambda \in K \Rightarrow v + w, \lambda w \in \mathscr{L}$ haben wir bereits in (1.5) nachgerechnet.

(iv) Seien $a, b \in \mathbb{R}$, $a < b$. Die Menge $V$ der Funktionen $f : [a, b] \to \mathbb{R}$ besitzt die Struktur eines $\mathbb{R}$-Vektorraums; dies folgt aus Prop. (4.4), angewendet auf die Menge $X = [a, b]$ und den $\mathbb{R}$-Vektorraum $\mathbb{R}$. Die Menge der reellwertigen *stetigen* Funktionen auf $[a, b]$ ist ein Untervektorraum von $V$. Denn aus der Analysis einer Variablen ist bekannt, dass die konstante Funktion $[a, b] \to \mathbb{R}$, $x \mapsto 0$ stetig ist. Diese Funktion ist der Nullvektor des Vektorraus $V$. Außerdem sind mit $f, g$ auch die Funktionen $f + g$ und $\lambda f$ stetig, für jedes $\lambda \in \mathbb{R}$.

Genauso lässt sich auch begründen, dass die Menge der stetigen, auf $]a, b[$ differenzierbaren Funktionen einen Untervektorraum von $V$ bildet, ebenso die Menge der Riemann-integrierbaren Funktionen auf $[a, b]$.

**(5.3) Proposition** Seien $V$ und $W$ zwei $K$-Vektorräume. Nach Prop. (4.4) ist dann auch die Menge $\mathrm{Abb}(V, W)$ aller Abbildungen $V \to W$ ein $K$-Vektorraum. Die linearen Abbildungen $\phi : V \to W$ einen Untervektorraum von $\mathrm{Abb}(V, W)$.

***Beweis:*** Sei $A = \mathrm{Abb}(V, W)$ und $U \subseteq A$ die Teilmenge der linearen Abbildungen zwischen $V$ und $W$. Zunächst müssen wir zeigen, dass die Abbildung $0_A$, die jeden Vektor aus $V$ auf $0_W$ abbildet, in $U$ enthalten ist. Seien dazu $v, v' \in V$ und $\lambda \in K$ vorgegeben. Es gilt $0_A(v + v') = 0_A(v) + 0_A(v') = 0_W + 0_W = 0_W$ und $0_A(\lambda v) = \lambda 0_A(v) = \lambda \cdot 0_W = 0_W$. Also ist $0_A$ in $U$ enthalten.

Seien nun $\phi, \psi \in U$ und $\lambda \in K$ vorgegeben. Zu zeigen ist $\phi + \psi \in U$ und $\lambda\phi \in U$, d.h. wir müssen nachrechnen, dass diese Abbildungen linear sind. Seien dazu $v, v' \in V$ und $\mu \in K$. Weil $\phi$ und $\psi$ linear sind, gilt

$$(\phi + \psi)(v + v') \quad = \quad \phi(v + v') + \psi(v + v') \quad = \quad \phi(v) + \phi(v') + \psi(v) + \psi(v') \quad =$$
$$\phi(v) + \psi(v) + \phi(v') + \psi(v') \quad = \quad (\phi + \psi)(v) + (\phi + \psi)(v')$$

und ebenso

$$(\phi + \psi)(\mu v) \quad = \quad \phi(\mu v) + \psi(\mu v) \quad = \quad \mu\phi(v) + \mu\psi(v) \quad = \quad \mu(\phi(v) + \psi(v)) \quad = \quad \mu(\phi + \psi)(v).$$

Damit ist $\phi + \psi \in U$ nachgewiesen. Nun zeigen wir noch, dass $\lambda\phi$ in $U$ liegt. Es gilt

$$(\lambda\phi)(v + v') \quad = \quad \lambda\phi(v + v') \quad = \quad \lambda(\phi(v) + \phi(v')) \quad = \quad \lambda\phi(v) + \lambda\phi(v') \quad = \quad (\lambda\phi)(v) + (\lambda\phi)(v')$$

und $(\lambda\phi)(\mu v) = \lambda\phi(\mu v) = \lambda\mu\phi(v) = \mu\lambda\phi(v) = \mu(\lambda\phi)(v)$. Also ist $\lambda\phi$ eine lineare Abbildung. □

Wegen (5.2) zeigt (5.3) insbesondere, dass die linearen Abbildungen zwischen zwei $K$-Vektorräumen $V, W$ selbst einen $K$-Vektorraum bilden. Man bezeichnet diesen üblicherweise mit $\mathrm{Hom}_K(V, W)$.

**(5.4) Proposition** Sei $V$ ein $K$-Vektorraum, und seien $U, U'$ Untervektorräume von $V$. Dann sind auch die Mengen

$$U \cap U' \qquad \text{und} \qquad U + U' = \{u + u' \mid u \in U, u' \in U'\} \qquad \text{Untervektorräume von } V.$$

Man bezeichnet $U + U'$ als die ***Summe*** von $U$ und $U'$.

***Beweis:*** Zunächst beweisen wir die Untervektorraum-Eigenschaft von $U \cap U'$. Weil $U, U'$ nach Voraussetzung Untervektorräume sind, gilt $0_V \in U$ und $0_V \in U'$. Es folgt $0_V \in U \cap U'$. Seien nun Elemente $v_1, v_2 \in U \cap U'$ und $\lambda \in K$ beliebig vorgegeben. Dann gilt insbesondere $v_1, v_2 \in U$. Weil $U$ ein Untervektorraum ist, folgt $v_1 + v_2 \in U$ und $\lambda v_1 \in U$, und ebenso gilt $v_1 + v_2 \in U'$ und $\lambda v_1 \in U'$, weil $U'$ ein Untervektorraum ist. Aus $v_1 + v_2 \in U$ und $v_1 + v_2 \in U'$ folgt $v_1 + v_2 \in U \cap U'$, ebenso erhalten wir $\lambda v_1 \in U \cap U'$. Damit sind die Untervektorraum-Eigenschaften für die Menge $U \cap U'$ nachgewiesen.

Nun zeigen wir, dass auch die Menge $U + U'$ ein Untervektorraum von $V$ ist. Wegen $0_V \in U$ und $0_V \in U'$ gilt zunächst $0_V = 0_V + 0_V \in U + U'$. Seien nun $v_1, v_2 \in U + U'$ und $\lambda \in K$ vorgegeben. Dann gibt es $u_1, u_2 \in U$ und $u_1', u_2' \in U'$ mit $v_1 = u_1 + u_1'$ und $v_2 = u_2 + u_2'$. Weil $U$ ein Untervektorraum ist, gilt $u_1 + u_2 \in U$, ebenso gilt $u_1' + u_2' \in U'$. Es folgt $v_1 + v_2 = (u_1 + u_1') + (u_2 + u_2') = (u_1 + u_2) + (u_1' + u_2') \in U + U'$. Aus der Untervektorraum-Eigenschaft von $U$ und $U'$ folgt auch, dass $\lambda u_1 \in U$ und $\lambda u_1' \in U'$ gilt. Wir erhalten $\lambda v_1 = \lambda(u_1 + u_1') = \lambda u_1 + \lambda u_1' \in U + U'$. Damit haben wir auch die Untervektorraum-Eigenschaften von $U + U'$ nachgerechnet. □

Auch aus mehr als zwei Untervektorräumen kann eine Summe gebildet werden. Sei $V$ ein $K$-Vektorraum, und sei $U_1, U_2, U_3, \ldots$ eine beliebige Anzahl von Untervektorräumen von $V$. Man definiert

$$\sum_{k=1}^{1} U_k = U_1 \qquad \text{und} \qquad \sum_{k=1}^{r+1} U_k = \left( \sum_{k=1}^{r} U_k \right) + U_{r+1} \quad \text{für} \quad r \geq 1.$$

Der Nachweis, dass es sich bei $\sum_{k=1}^{r} U_k$ für jedes $r \in \mathbb{N}$ um einen Untervektorraum von $V$ handelt, erfolgt durch vollständige Induktion über $r$, was hier aus Zeitgründen aber nicht ausgeführt wird. Ebenso zeigt man durch vollständige Induktion, dass

$$\sum_{k=1}^{r} U_k \quad = \quad \left\{ \sum_{k=1}^{r} u_k \;\middle|\; u_k \in U_k \text{ für } 1 \leq k \leq r \right\} \qquad \text{gilt.}$$

**(5.5) Definition** Ein $K$-Vektorraum $V$ wird ***direkte Summe*** der Untervektorräume $U, U' \subseteq V$ genannt, wenn die Bedingungen

$$V = U + U' \quad \text{und} \quad U \cap U' = \{0_V\} \qquad \text{erfüllt sind.}$$

Die direkte Summe zweier Untervektorräume $U, U'$ wird mit $U \oplus U'$ bezeichnet.

**(5.6) Lemma** Sei $V$ ein $K$-Vektorraum mit Untervektorräumen $U, U' \subseteq V$. Dann sind die folgenden Bedingungen äquivalent:

(i) $V = U \oplus U'$

(ii) Für jedes $v \in V$ gibt es eindeutig bestimmte Vektoren $u \in U$ und $u' \in U'$ mit $v = u + u'$.

***Beweis:*** „(i) $\Rightarrow$ (ii)" Wegen $V = U + U'$ gibt es für jeden Vektor $v \in V$ Elemente $u \in U$ und $u' \in U'$ mit $v = u + u'$. Wir beweisen nun die Eindeutigkeit. Sei $v \in V$, und seien $u_1, u_2 \in U$ und $u'_1, u'_2 \in U'$ mit $v = u_1 + u'_1 = u_2 + u'_2$. Dann gilt $u_1 - u_2 = u'_2 - u'_1 \in U \cap U'$. Wegen $U \cap U' = \{0_V\}$ folgt $u_1 - u_2 = u'_2 - u'_1 = 0_V$, also $u_1 = u_2$ und $u'_1 = u'_2$.

„(ii) $\Rightarrow$ (i)" Weil jeder Vektor $v \in V$ in der Form $v = u + u'$ mit $u \in U$ und $u \in U'$ geschrieben werden kann, gilt $V = U + U'$. Wir zeigen nun, dass auch $U \cap U' = \{0_V\}$ erfüllt ist. Die Inklusion „$\supseteq$" ist offensichtlich, da $U$ und $U'$ Untervektorräume sind und somit $0_V$ in $U$ und $U'$ enthalten ist. Sei nun $v \in U \cap U'$ vorgegeben. Nach Voraussetzung gibt es eindeutig bestimmte $u \in U$, $u' \in U'$ mit $v = u + u'$. Aus $v = 0_V + v$ mit $0_V \in U$ und $v \in U'$ folgt auf Grund der Eindeutigkeit $u = 0_V$. Ebenso können wir $v$ auch in der Form $v = v + 0_V$ mit $v \in U$ und $0_V \in U'$ schreiben. Diesmal liefert die Eindeutigkeit die Gleichung $u' = 0_V$. Insgesamt erhalten wir $v = u + u' = 0_V + 0_V = 0_V$. □

Auch die direkte Summe von mehreren Untervektorräumen lässt sich rekursiv definieren. Sei $V$ ein $K$-Vektorraum, und seien $U_1, U_2, U_3, ...$ Untervektorräume von $V$. Dann setzt man

$$\bigoplus_{k=1}^{1} U_k = U_1 \qquad \text{und} \qquad \bigoplus_{k=1}^{r+1} U_k = \left( \bigoplus_{k=1}^{r} U_k \right) \oplus U_{r+1} \quad \text{für} \quad r \geq 1.$$

Damit die direkte Summe $\bigoplus_{k=1}^{r+1} U_k$ gebildet werden kann, dürfen sich $\bigoplus_{k=1}^{r} U_k$ und $U_{r+1}$ jeweils nur in $\{0_V\}$ schneiden.

**(5.7) Satz** Sei $V$ ein Vektorraum, $r \in \mathbb{N}$, und seien $U_1, ..., U_r$ Untervektorräume von $V$. Dann sind die folgenden Aussagen äquivalent.

(i) Es gilt $V = \bigoplus_{k=1}^{r} U_k$.

(ii) Jeder Vektor $v \in V$ kann auf eindeutige Weise als Summe $v = \sum_{k=1}^{r} u_k$ dargestellt werden, mit $u_k \in U_k$ für $1 \leq k \leq r$.

(iii) Für $1 \leq k \leq r$ gilt $V = \sum_{j=1}^{r} U_j$ und $U_k \cap \left( \sum_{j \neq k} U_j \right) = \{0_V\}$.

***Beweis:*** Wir beweisen die Äquivalenz der drei Aussagen durch vollständige Induktion über $r$. Im Fall $r = 1$ besteht (i) nur in der Aussage $V = U_1$. Aussage (ii) besagt, dass für jedes $v \in V$ ein eindeutig bestimmter Vektor $u_1 \in U_1$ mit $v = u_1$ existiert, was offenbar zu (i) äquivalent ist. Die Aussage (iii) besteht aus den Gleichungen $V = U_1$ und $U_1 \cap \{0_V\} = \{0_V\}$, und wiederum ist „(i) $\Leftrightarrow$ (iii)" offensichtlich.

Sei nun $r \in \mathbb{N}$ vorgegeben, und setzen wir die Äquivalenz von (i), (ii) und (iii) für dieses $r$ voraus. Seien $U_1, ..., U_{r+1}$ beliebige Untervektorräume von $V$. Wir beginnen mit dem Beweis der Implikation „(i) $\Rightarrow$ (ii)". Hier lautet die Voraussetzung

$$V \quad = \quad \bigoplus_{k=1}^{r+1} U_k \quad = \quad \left( \bigoplus_{k=1}^{r} U_k \right) \oplus U_{r+1}.$$

Insbesondere gilt $V = \sum_{k=1}^{r+1} U_k$; dies bedeutet, dass jedes $v \in V$ jedenfalls als Summe $v = u_1 + ... + u_{r+1}$ dargestellt werden kann, mit $u_k \in U_k$ für $1 \leq k \leq r+1$. Nehmen wir nun an, dass $v = u'_1 + ... + u'_{r+1}$ eine weitere solche Darstellung ist. Weil $V$ nach Voraussetzung die direkte Summe von $\bigoplus_{k=1}^{r} U_k$ und $U_{r+1}$ ist, folgt $u_1 + ... + u_r = u'_1 + ... + u'_r$ und $u_{r+1} = u'_{r+1}$ nach (5.6). Nach Induktionsvoraussetzung besitzt jedes Element in $\bigoplus_{k=1}^{r} U_k$ eine eindeutige Darstellung als Summe von Elementen in $U_1, ..., U_k$. Aus $u_1 + ... + u_r = u'_1 + ... + u'_r$ folgt also $u_k = u'_k$ für $1 \leq k \leq r$.

Beweisen wir nun die Implikation „(ii) $\Rightarrow$ (iii)“ und setzen dazu (ii) voraus. Zunächst zeigen wir die Gleichung $V = \sum_{k=1}^{r+1} U_k$. Die Inklusion „$\supseteq$“ ist nach Definition der Summe offensichtlich. Andererseits hat auf Grund unserer Voraussetzung jedes Element $v \in V$ eine Darstellung $v = u_1 + ... + u_{r+1}$ mit $u_k \in U_k$ für $1 \leq k \leq r+1$. Also gilt auch „$\subseteq$“. Sei nun $k \in \{1, ..., r+1\}$ vorgegeben. Zu zeigen ist die Gleichung

$$U_k \cap \left( \sum_{j \neq k} U_j \right) \quad = \quad \{0_V\}.$$

Hier ist „$\supseteq$“ offensichtlich erfüllt. Zum Beweis von „$\subseteq$“ nehmen wir an, dass ein Vektor $v \neq 0_V$ im Durchschnitt existiert. Dann liegt $v$ einerseits in $U_k$, andererseits gilt $v = \sum_{j \neq k} u_j$ für gewisse Elemente $u_j$ mit $u_j \in U_j$ für $1 \leq j \leq r+1$ und $j \neq k$. Setzen wir $u_k = -v$, dann gilt $\sum_{j=1}^{r+1} u_j = 0_V$. Weil aber der Nullvektor auch in der Form $0_V + ... + 0_V$ mit $0_V \in U_j$ für $1 \leq j \leq r+1$ dargestellt werden kann, und weil diese Darstellung nach (ii) eindeutig ist, folgt $u_j = 0_V$ für $1 \leq j \leq r+1$ mit $j \neq k$ und auch $v = -u_k = 0_V$, im Widerspruch zur Annahme.

Zeigen wir nun noch die Implikation „(iii) $\Rightarrow$ (i)“ und setzen dazu (iii) voraus. Zu zeigen ist

$$V \quad = \quad \bigoplus_{k=1}^{r+1} U_k \quad = \quad \left( \bigoplus_{k=1}^{r} U_k \right) \oplus U_{r+1}.$$

Wir betrachten den Untervektorraum $U = \sum_{k=1}^{r} U_k$. Nach Voraussetzung gilt $U_k \cap (\sum_{j \neq k} U_j) = \{0_V\}$ für $1 \leq k \leq r+1$. Damit ist für $1 \leq k \leq r$ jeweils erst recht der Durchschnitt von $U_k$ mit $\sum_{j \neq k, r+1} U_j$ gleich $\{0_V\}$. Also ist die Bedingung (iii) für den $K$-Vektorraum $U$ und die Untervektorräume $U_1, ..., U_r$ von $U$ erfüllt. Die Induktionsvoraussetzung liefert uns damit

$$U \quad = \quad \bigoplus_{k=1}^{r} U_k.$$

Weiter gilt nach Voraussetzung $V = U + V_{r+1}$, außerdem $U \cap V_{r+1} = \{0_V\}$. Wie gewünscht erhalten wir $V = U \oplus V_{r+1} = \bigoplus_{k=1}^{r+1} U_k$. □

**(5.8) Proposition** Seien $V, W$ $K$-Vektorräume und $\phi : V \to W$ eine lineare Abbildung. Ferner seien $V' \subseteq V$ und $W' \subseteq W$ Untervektorräume. Dann sind die Teilmengen

$$\phi(V') = \{\phi(v) \mid v \in V'\} \qquad \text{und} \qquad \phi^{-1}(W') = \{v \in V \mid \phi(v) \in W'\}$$

Untervektorräume von $W$ bzw. von $V$.

***Beweis:*** Wir rechnen die Untervektorraum-Axiome für beide Teilmengen direkt nach. Seien $w, w' \in \phi(V')$ und $\lambda \in K$. Dann gibt es nach Definition von $\phi(V')$ Vektoren $v, v' \in V'$ mit $w = \phi(v)$ und $w' = \phi(v')$. Da $V'$ ein Untervektorraum ist, gilt $v + v' \in V'$ und damit

$$w + w' = \phi(v) + \phi(v') = \phi(v + v') \quad \in \quad \phi(V').$$

Ebenso gilt $\lambda v \in V'$ auf Grund der Untervektorraum-Eigenschaft und somit $\lambda w = \lambda\phi(v) = \phi(\lambda v) \in \phi(V')$.

Nun zeigen wir, dass auch $\phi^{-1}(W')$ ein Untervektorraum ist. Seien dazu $v, v' \in \phi^{-1}(W')$ und $\lambda \in K$ vorgegeben. Dann gilt $\phi(v), \phi(v') \in W'$ und $\phi(v)+\phi(v') \in W'$, da $W'$ ein Untervektorraum von $W$ ist. Aus $\phi(v+v') = \phi(v)+\phi(v') \in W'$ folgt $v + v' \in \phi^{-1}(W')$. Da auch $\lambda\phi(v)$ in $W'$ liegt, erhalten wir $\phi(\lambda v) = \lambda\phi(v) \in W'$ und somit $\lambda v \in \phi^{-1}(W')$. □

**(5.9) Definition** Seien $V, W$ zwei $K$-Vektorräume und $\phi : V \to W$ eine lineare Abbildung. Dann nennt man

(i) $\ker(\phi) = \phi^{-1}(\{0_W\}) = \{ v \in V \mid \phi(v) = 0_W \}$ den ***Kern*** und

(ii) $\mathrm{im}(\phi) = \phi(V) = \{ \phi(v) \mid v \in V \}$ das ***Bild*** von $\phi$.

Nach (5.8) ist $\ker(\phi)$ ein Untervektorraum von $V$ und $\mathrm{im}(\phi)$ ein Untervektorraum von $W$.

**(5.10) Proposition** Seien $V, W$ $K$-Vektorräume und $\phi : V \to W$ eine lineare Abbildung.

(i) Die Abbildung $\phi$ ist genau dann surjektiv, wenn $\mathrm{im}(\phi) = W$ gilt.

(ii) Sie ist genau dann injektiv, wenn $\ker(\phi) = \{0_V\}$ erfüllt ist.

***Beweis:*** Aussage (i) ist nach Definition der Surjektivität unmittelbar klar. Zum Beweis von (ii) setzen wir zunächst voraus, dass $\phi$ injektiv ist. Die Inklusion $\{0_V\} \subseteq \ker(\phi)$ ist erfüllt, weil der Kern ein Untervektorraum von $V$ ist. Zum Nachweis von $\ker(\phi) \subseteq \{0_V\}$ sei $v \in \ker(\phi)$ vorgegeben. Dann gilt $\phi(v) = 0_W = \phi(0_V)$, und aus der Injektivität von $\phi$ folgt $v = 0_V$.

Setzen wir nun umgekehrt die Gleichung $\ker(\phi) = \{0_V\}$ voraus, und beweisen wir die Injektivität von $\phi$. Seien dazu $v, v' \in V$ mit $\phi(v) = \phi(v')$ vorgegeben. Dann gilt $\phi(v' - v) = \phi(v') - \phi(v) = 0_W$ und somit $v' - v \in \ker(\phi)$. Aus der Voraussetzung an den Kern folgt $v' - v = 0_V \Leftrightarrow v = v'$. □

**(5.11) Definition** Eine Teilmenge $A \subseteq V$ wird ***affiner Unterraum*** von $V$ genannt, wenn entweder $A = \varnothing$ gilt oder ein Untervektorraum $U$ und ein Vektor $v \in V$ existieren, so dass

$$A \quad = \quad v + U \quad = \quad \{v + u \mid u \in U\} \qquad \text{erfüllt ist.}$$

Betrachten wir einige konkrete Beispiele für affine Unterräume.

(i) Seien $u, v \in V$. Dann ist $u + \operatorname{lin}(v) = \{u + \lambda v \mid \lambda \in K\}$ ein affiner Unterraum. Im Fall $v \neq 0_V$ bezeichnet man ihn als ***affine Gerade***. Für beliebige $u, v, w \in V$ ist auch durch

$$u + \operatorname{lin}(v, w) \quad = \quad \{u + \lambda v + \mu w \mid \lambda, \mu \in K\}$$

ein affiner Unterraum gegeben. Ist $v \neq 0_V$ und $w$ kein skalares Vielfaches von $v$ (also $w \neq \lambda v$ für alle $\lambda \in K$), dann nennt man $u + \operatorname{lin}(v, w)$ eine ***affine Ebene***.

(ii) Die Lösungsmenge $\mathscr{L} \subseteq K^n$ eines beliebigen linearen Gleichungssystems bestehend aus $m$ Gleichungen in $n$ Unbekannten ist ein affiner Unterraum von $K^n$. Dies folgt aus der Tatsache, dass für jedes Element $v \in \mathscr{L}$ nach (1.6) die Gleichung $\mathscr{L} = v + \mathscr{L}^h$ gilt, wobei $\mathscr{L}^h$ die Lösungsmenge des zugehörigen *homogenen* LGS bezeichnet, und dass $\mathscr{L}^h$ stets ein Untervektorraum von $K^n$ ist.

**(5.12) Proposition** Sei $V$ ein $K$-Vektorraum und $\varnothing \neq A \subseteq V$ ein affiner Unterraum.

(i) Es gibt einen eindeutig bestimmten Untervektorraum $U$, so dass $A = v + U$ für ein $v \in V$ erfüllt ist.

(ii) Für jeden Vektor $w \in A$ erfüllt der Untervektorraum $U$ aus Teil (i) die Gleichung $A = w + U$.

Wir nennen $U$ den zu $A$ gehörenden Untervektorraum und bezeichnen ihn mit $\mathscr{L}(A)$.

***Beweis:*** zu (i) Nehmen wir an, dass $v, v' \in V$ Vektoren und $U, U'$ Untervektorräume von $V$ mit $v + U = A = v' + U'$ sind. Wegen $v \in A$ gilt $v = v' + u_0$ für ein $u_0 \in U'$, und wegen $v' \in A$ gilt $v' = v + u_1$ für ein $u_1 \in U$. Der Differenzvektor $v' - v$ ist also sowohl in $U$ als auch in $U'$ enthalten. Wir beweisen nun die Gleichung $U = U'$.

„$\subseteq$“ Ist $u \in U$, dann liegt $v + u$ in $A$, und folglich gibt es ein $u' \in U'$ mit $v + u = v' + u'$. Es folgt $u = (v' - v) + u' \in U'$. „$\supseteq$“ Ist $u' \in U'$ vorgegeben, dann gilt $v' + u'$ in $A$, es gibt also ein $u \in U$ mit $v' + u' = v + u$. Daraus folgt $u' = (v - v') + u \in U$.

zu (ii) Sei $U = \mathscr{L}(A)$, $v \in V$ ein Vektor mit $A = v + U$ und $w \in A$ ein beliebiges Element. Dann gibt es ein $u \in U$ mit $w = v + u$. Wir beweisen nun die Gleichung $v + U = w + U$. „$\subseteq$“ Ist $v_1 \in v + U$, dann gibt es ein $u_1 \in U$ mit $v_1 = v + u_1$, und es folgt $v_1 = (w - u) + u_1 = w + (u_1 - u) \in w + U$. „$\supseteq$“ Ist $w_1 \in w + U$, dann existiert ein $u_1 \in U$ mit $w_1 = w + u_1$. Es folgt $w_1 = w + u_1 = (v + u) + u_1 = v + (u + u_1) \in v + U$. □

# § 6. Erzeugendensyteme und lineare Unabhängigkeit

***Inhaltsübersicht***

Bereits im letzten Kapitel haben wir gesehen, wie durch die Angabe von ein oder zwei Elementen $v, w$ eines Vektorraums $V$ Untervektorräume definiert werden können, die wir mit $\text{lin}(v)$ und $\text{lin}(v, w)$ bezeichnet hatten. Wir werden nun sehen, dass jeder beliebigen Teilmenge $S \subseteq V$ ein Untervektorraum $\text{lin}(S)$ zugeordnet werden kann. Man bezeichnet $S$ dann das *Erzeugendensystem* von $\text{lin}(S)$. Die Elemente von $\text{lin}(S)$ sind die *Linearkombinationen* von $S$; dabei handelt es sich um die Summen der Form $\lambda_1 v_1 + ... + \lambda_n v_n$ mit $\lambda_1, ..., \lambda_n \in K$ und $v_1, ..., v_n \in S$.

Die Erzeugendensysteme ermöglichen eine einfache und kompakte Beschreibung aller Untervektorräume eines Vektorraums $V$, denn wie wir sehen werden, benötigt man in der Regel nur Teilmengen $S \subseteq V$ mit einer begrenzten Anzahl von Vektoren. So kann zum Beispiel jeder Untervektorraum von $\mathbb{R}^3$ durch die Angabe von höchstens drei Vektoren beschrieben werden. Neben der Definition von $\text{lin}(S)$ werden wir einige einfache Regeln zum Umgang mit Erzeugendensystemen kennenlernen. Wichtig ist auch die Charakterisierung von $\text{lin}(S)$ als kleinster Untervektorraum $U$ von $V$, der $S$ als Teilmenge enthält.

Eine Teilmenge $S \subseteq V$ heißt *linear abhängig*, wenn der Nullvektor $0_V$ als nichttriviale Linearkombination von Elementen aus $S$ darstellbar ist. Beispielsweise ist $\{e_1, e_2, e_1 + e_2\}$ eine linear abhängige Teilmenge von $\mathbb{R}^2$, denn es gilt die Gleichung $e_1 + e_2 + (-1)(e_1 + e_2) = 0_{\mathbb{R}^2}$. Bei Mengen mit mehr als einem Element ist dies gleichbedeutend damit, dass ein Element in $S$ existiert, dass als Linearkombination der anderen Elemente aus $S$ dargestellt werden kann. Der Begriff der linearen Unabhängigkeit wird im nächsten Kapitel bei der Definition des Dimensionsbegriffs eine wichtige Rolle spielen.

***Wichtige Definitionen***

- Linearkombination eines Tupels $(v_1, ..., v_r)$ von Vektoren
- von einer Teilmenge $S \subseteq V$ aufgespannter Untervektorraum $\text{lin}(S)$
- Erzeugendensystem eines $K$-Vektorraums
- lineare Unabhängigkeit eines Tupels bzw. einer Menge von Vektoren

**(6.1) Definition** Sei $V$ ein $K$-Vektorraum, $r \in \mathbb{N}_0$ und $(v_1, ..., v_r)$ ein Tupel von Elementen aus $V$; im Fall $r = 0$ ist das leere Tupel $()$ ohne Einträge gemeint. Wir bezeichnen einen Vektor $w \in V$ als ***Linearkombination*** des Tupels, wenn ein Tupel $(\lambda_1, ..., \lambda_r) \in K^r$ existiert, so dass

$$w = \sum_{i=1}^{r} \lambda_i v_i \qquad \text{erfüllt ist.}$$

Ist $S \subseteq V$ eine Teilmenge, dann bezeichnen wir $w$ als Linearkombination von $S$, wenn ein Tupel $(v_1, ..., v_r)$ mit $v_i \in S$ für $1 \leq i \leq r$ existiert, so dass $w$ Linearkombination dieses Tupels ist. Die Menge aller Linearkombinationen von $S$ in $V$ bezeichnen wir mit $\text{lin}(S)$.

Für das leere Tupel () existiert nur eine einzige Linearkombination, nämlich der Nullvektor $0_V$. Der Definition unmittelbar zu entnehmen ist, dass aus $S \subseteq S'$ stets $\mathrm{lin}(S) \subseteq \mathrm{lin}(S')$ folgt. Gilt $V = \mathrm{lin}(S)$, dann wird $S$ ein ***Erzeugendensystem*** von $V$ genannt. In diesem Fall sagt man auch, dass der Vektorraum $V$ von $S$ ***erzeugt*** oder ***aufgespannt*** wird.

Für die Vektorräume $K^n$ und $\mathscr{M}_{m\times n,K}$ existieren endliche Erzeugendensysteme, nämlich im ersten Fall die Menge $\{e_1,...,e_n\}$ der Einheitsvektoren und im zweiten die Menge $\{B_{k\ell} \mid 1 \le k \le m, 1 \le \ell \le n\}$ der Basismatrizen. Es gibt aber auch Vektorräume, die durch keine endliche Teilmenge ihrer Vektoren aufgespannt werden können.

**(6.2) Satz** *(ohne Beweis)*

Für jeden Körper $K$ gibt es einen Erweiterungsring $K[x] \supseteq K$ mit einem ausgezeichneten Element $x \notin K$, so dass für jedes $f \in K[x]$ folgende Bedingung erfüllt ist. *Entweder* ist $f = 0_K$, *oder* es gibt ein eindeutig bestimmtes $n \in \mathbb{N}_0$ und eindeutig bestimmte Elemente $a_0,...,a_n \in K$ mit

$$f \quad = \quad \sum_{k=0}^{n} a_k x^k \quad \text{und} \quad a_n \neq 0_K .$$

Man nennt $K[x]$ den ***Polynomring*** über dem Körper $K$, seine Elemente bezeichnet man als *Polynome*. Im Fall $f \neq 0_K$ bezeichnet man $n$ als den ***Grad*** des Polynoms.

Es sei noch einmal hervorgehoben, dass das Element $x$ des Polynomrings $K[x]$ kein Element des Körpers $K$ ist! Man kann $K[x]$ als $K$-Vektorraum betrachten, in dem man die Vektoraddition mit der gewöhnlichen Addition im Ring $K[x]$ gleichsetzt und die skalare Multiplikation $K \times K[x] \to K[x]$ durch Einschränkung der Multiplikationsabbildung

$$K[x] \times K[x] \longrightarrow K[x] \quad , \quad (f,g) \mapsto fg$$

definiert. In diesem Fall ist dann $S = \{x^n \mid n \in \mathbb{N}_0\}$ ein (unendliches) Erzeugendensystem von $K[x]$ als $K$-Vektorraum, wobei $x^0 = 1_K$ ist. Jedes Polynom ist Linearkombination von $S$. Beispielsweise ist das Polynom $f = x^7 - 4x^3 + 5$ eine Linearkombination des Tupels $(1, x^3, x^7)$, und ebenso eine Linearkombination von $(1, x, x^2, ..., x^7)$. Man kann sich aber leicht überzeugen, dass keine *endliche* Teilmenge $T \subseteq K[x]$ mit der Eigenschaft $\mathrm{lin}(T) = K[x]$ existiert. Denn in $T$ gibt es ein Polynom mit maximalem Grad $n$, und folglich kann keine Linearkombination von $T$ einen Grad größer als $n$ haben. Dies bedeutet, dass zum Beispiel $x^{n+1}$ nicht in $\mathrm{lin}(T)$ enthalten ist.

**(6.3) Satz** Sei $V$ ein $K$-Vektorraum und $S \subseteq V$ eine Teilmenge. Dann gilt

(i) Die Menge $\mathrm{lin}(S)$ bildet einen Untervektorraum von $V$ mit $\mathrm{lin}(S) \supseteq S$.

(ii) Ist $U$ ein weiterer Untervektorraum von $V$ mit $U \supseteq S$, dann gilt $U \supseteq \mathrm{lin}(S)$.

Somit ist $\mathrm{lin}(S)$ der kleinste Untervektorraum von $V$, der $S$ als Teilmenge enthält.

***Beweis:*** zu (i) Zunächst beweisen wir, dass $\mathrm{lin}(S)$ ein Untervektorraum von $V$ ist. Der Nullvektor $0_V$ ist eine Linearkombination des leeren Tupels und somit in $\mathrm{lin}(S)$ enthalten. Seien nun $v, w \in \mathrm{lin}(S)$ und $\lambda \in K$ vorgegeben. Zu zeigen ist $v + w \in \mathrm{lin}(S)$ und $\lambda v \in \mathrm{lin}(S)$. Wegen $v \in \mathrm{lin}(S)$ gibt es ein $r \in \mathbb{N}_0$ und ein Tupel $(v_1, ..., v_r)$, so dass $v$ Linearkombination dieses Tupels ist. Es existiert also ein Tupel $(\lambda_1, ..., \lambda_r) \in K^r$ mit $v = \sum_{i=1}^{r} \lambda_i v_i$. Genauso folgt aus $w \in \mathrm{lin}(S)$ die Existenz eines $s \in \mathbb{N}_0$ und von $w_1, ..., w_s \in S$ und $\mu_1, ..., \mu_s \in K$ mit $w = \sum_{j=1}^{s} \mu_j w_j$. Die Gleichung

$$v + w \quad = \quad \sum_{i=1}^{r} \lambda_i v_i + \sum_{j=1}^{s} \mu_j w_j$$

zeigt, dass $v + w$ eine Linearkombination des Tupels $(v_1, ..., v_r, w_1, ..., w_s)$ ist und somit in $\mathrm{lin}(S)$ liegt. Ebenso folgt aus $\lambda v = \sum_{i=1}^{r} (\lambda \lambda_i) v_i$, dass $\lambda v$ eine Linearkombination von $(v_1, ..., v_r)$ ist und $\lambda v$ somit ebenfalls in $\mathrm{lin}(S)$ enthalten ist. Der Nachweis der Untervektorraum-Eigenschaft von $\mathrm{lin}(S)$ ist damit abgeschlossen. Es gilt $S \subseteq \mathrm{lin}(S)$, denn jedes $v \in S$ ist wegen $v = 1_K \cdot v$ jeweils Linearkombination des einelementigen Tupels $(v)$ und somit nach Definition in $\mathrm{lin}(S)$ enthalten.

zu (ii) Sei $U$ ein beliebiger Untervektorraum von $V$ mit $U \supseteq S$. Wir zeigen durch vollständige Induktion über $n \in \mathbb{N}_0$, dass jede Linearkombination jedes $n$-elementigen Tupels $(v_1, ..., v_n)$ mit $v_i \in S$ für $1 \leq i \leq n$ in $U$ enthalten ist. Daraus folgt dann unmittelbar $\mathrm{lin}(S) \subseteq U$. Für $n = 0$ ist die Aussage klar, denn die einzige Linearkombination des leeren Tupels () ist der Nullvektor $0_V$, und es gilt $0_V \in U$, weil $U$ ein Untervektorraum von $V$ ist.

Sei nun $n \in \mathbb{N}_0$, und setzen wir die Aussage für dieses $n$ voraus. Sei $(v_1, ..., v_{n+1})$ ein Tupel mit $v_i \in S$ für $1 \leq i \leq n+1$, und sei $w$ eine Linearkombination dieses Tupels. Es gibt dann also ein Tupel $(\lambda_1, ..., \lambda_{n+1}) \in K^{n+1}$ mit $w = \sum_{i=1}^{n+1} \lambda_i v_i$. Der Vektor $w' = \sum_{i=1}^{n} \lambda_i v_i$ ist eine Linearkombination des $n$-elementigen Tupels $(v_1, ..., v_n)$ und somit nach Induktionsvoraussetzung in $U$ enthalten. Weiter gilt $v_{n+1} \in U$ wegen $S \subseteq U$ und weiter $\lambda_{n+1} v_{n+1} \in U$ und $w = w' + \lambda_{n+1} v_{n+1} \in U$, da $U$ ein Untervektorraum von $V$ ist. Damit ist der Induktionsbeweis abgeschlossen. □

**(6.4) Definition** Sei $V$ ein $K$-Vektorraum. Wir bezeichnen ein Tupel $(v_1, ..., v_r)$ mit $r \in \mathbb{N}_0$, bestehend aus Vektoren $v_i \in V$, als ***linear unabhängig***, wenn für jedes Tupel $(\lambda_1, ..., \lambda_r) \in K^r$ die Implikation

$$\sum_{i=1}^{r} \lambda_i v_i = 0_V \quad \Rightarrow \quad \lambda_1 = ... = \lambda_r = 0_K$$

erfüllt ist. Eine Teilmenge $S \subseteq V$ bezeichnen wir als ***linear unabhängig***, wenn jedes Tupel $(v_1, ..., v_r)$ bestehend aus lauter verschiedenen Elementen der Menge $S$ linear unabhängig ist.

Ein Tupel oder eine Teilmenge von $V$, die nicht linear unabhängig ist, wird ***linear abhängig*** genannt. Jedes Tupel, dass einen Vektor mehrfach enthält, ist automatisch linear abhängig. Sind beispielsweise $v, w \in V$ zwei Vektoren, dann ist das Tupel $(v, w, v)$ linear abhängig, denn es gilt $1_K \cdot v + 0_K \cdot w + (-1_K) \cdot v = 0_V$, aber nicht alle Einträge des Tupels $(1_K, 0_K, -1_K)$ sind gleich Null.

Ebenso kann man sich leicht überzeugen, dass jedes Tupel, und damit auch jede Teilmenge des Vektorraums $V$, die den Nullvektor $0_V$ enthält, linear abhängig ist.

Das leere Tupel ist per Definition linear unabhängig, und dasselbe gilt auch für die leere Menge $\varnothing$. Jede Teilmenge einer linear unabhängigen Menge ist offenbar ebenfalls linear unabhängig. Ist $(v_1, ..., v_n)$ ein linear unabhängiges Tupel, dann ist $\{v_1, ..., v_n\}$ eine $n$-elementige linear unabhängige Teilmenge von $V$.

Beispielsweise ist im $K$-Vektorraum $K^n$ das Tupel $(e_1, ..., e_n)$ bestehend aus den Einheitsvektoren linear unabhängig. Sei nämlich $(\lambda_1, ..., \lambda_n) \in K^n$ ein Tupel mit $\sum_{i=1}^n \lambda_i v_i = 0_V$. Für $1 \leq k \leq n$ ist die $k$-te Komponente von $\sum_{i=1}^n \lambda_i e_i$ jeweils gleich $\sum_{i=1}^n \lambda_i \delta_{ik} = \lambda_k$. Aus $\sum_{i=1}^n \lambda_i v_i = 0_V$ folgt also $\lambda_k = 0_K$ für $1 \leq k \leq n$. Also ist $(e_1, ..., e_n)$ linear unabhängig. Aus unserer letzten Anmerkung im vorherigen Absatz folgt somit, dass $\{e_1, ..., e_n\}$ eine $n$-elementige linear unabhängige Teilmenge von $K^n$ ist.

Ebenso ist die Menge $\{ B_{k\ell}^{(m\times n)} \mid 1 \leq k \leq m, 1 \leq \ell \leq n \}$ der Basismatrizen im $K$-Vektorraum $\mathscr{M}_{m\times n,K}$ der $m \times n$-Matrizen linear unabhängig. Die Teilmenge $\{x^n \mid n \in \mathbb{N}_0\}$ des Polynomrings $K[x]$ ist ein Beispiel für eine unendliche linear unabhängige Menge.

**(6.5) Lemma** Sei $V$ ein $K$-Vektorraum.

(i) Eine einlementige Teilmenge $\{v\} \subseteq V$ ist genau dann linear abhängig, wenn $v = 0_V$ ist, ansonsten linear unabhängig.

(ii) Eine zweielementige Menge $\{v, w\}$ ist genau dann linear abhängig, wenn ein $\lambda \in K$ mit $v = \lambda w$ oder $w = \lambda v$ existiert.

***Beweis:*** zu (i) „$\Leftarrow$“ Wir haben bereits festgestellt, dass ein Tupel oder eine Menge, die den Nullvektor enthält, linear abhängig ist. Insbesondere ist $(0_V)$ linear abhängig (denn es gilt $1_K \cdot 0_V = 0_V$, aber nicht alle Einträge des Tupels $(1_K)$ sind gleich Null), und somit auch $\{0_V\}$ linear abhängig. „$\Rightarrow$“ Wenn $\{v\}$ linear abhängig ist, dann gibt es ein Tupel bestehend aus lauter verschiedenen Vektoren der Menge $\{v\}$, das linear abhängig ist. Hierfür kommt nur das einelementige Tupel $(v)$ in Frage. Weil es linear abhängig ist, gibt es ein $\lambda \in K \setminus \{0_K\}$ mit $\lambda v = 0_V$. Daraus folgt $v = 1_K \cdot v = \lambda^{-1}(\lambda v) = \lambda^{-1} \cdot 0_V = 0_V$.

zu (ii) „$\Rightarrow$“ Wir betrachten nur den Fall, dass $w = \lambda v$ für ein $\lambda \in K$ gilt; im anderen Fall läuft die Argumentation vollkommen analog. Weil $\{v, w\}$ zweielementig ist, gilt $v \neq w$. Somit ist $(v, w)$ ein Tupel bestehend aus verschiedenen Elementen der Menge $\{v, w\}$. Dieses Tupel ist linear abhängig, denn es gilt $(-\lambda)v + 1_K \cdot w = -w + w = 0_V$, aber nicht alle Einträge des Tupels $(-\lambda, 1_K)$ sind gleich Null. Also ist auch die Menge $\{v, w\}$ linear abhängig.

„$\Leftarrow$“ Ist $\{v, w\}$ linear abhängig, dann gibt es ein Tupel bestehend aus verschiedenen Elementen der Menge $\{v, w\}$, das linear abhängig ist. Hierfür kommen nur die Tupel $(v)$, $(w)$ und $(v, w)$ in Frage. Wie wir bereits unter (i) gesehen haben, kann das Tupel $(v)$ nur linear abhängig sein, wenn $v = 0_V$ gilt. In diesem Fall ist $v = 0_K \cdot w$. Ebenso kann $(w)$ nur im Fall $w = 0_V$ linear abhängig sein, und dann folgt $w = 0_V \cdot v$. Ist $(v, w)$ linear abhängig, dann gibt es Koeffizienten $\lambda, \mu \in K$, nicht beide gleich Null, mit $\lambda v + \mu w = 0_V$. Im Fall $\lambda \neq 0_V$ gilt $v = (-\frac{\mu}{\lambda})w$, im Fall $\mu \neq 0_V$ ist $w = (-\frac{\lambda}{\mu})v$. $\square$

**(6.6) Proposition** Sei $V$ ein $K$-Vektorraum und $S \subseteq V$ eine beliebige Teilmenge.

(i) Genau dann ist $S$ linear abhängig, wenn ein $v \in S$ mit $v \in \mathrm{lin}(S \setminus \{v\})$ existiert.

(ii) Ist $S$ linear unabhängig und $v \in V \setminus \mathrm{lin}(S)$, dann ist auch $S \cup \{v\}$ linear unabhängig.

***Beweis:*** zu (i) „⇒" Angenommen, es gibt ein $v \in S$ mit $v \in \mathrm{lin}(S \setminus \{v\})$. Dann gibt es ein $r \in \mathbb{N}_0$ und ein Tupel $(v_1, ..., v_r)$ von Vektoren aus $S \setminus \{v\}$ mit der Eigenschaft, dass $v$ eine Linearkombination dieses Tupels ist. Es gibt also ein Tupel $(\lambda_1, ..., \lambda_r) \in K^r$ mit $v = \sum_{i=1}^r \lambda_i v_i$. Dabei dürfen wir davon ausgehen, dass die Vektoren $v_1, ..., v_r$ alle verschieden sind. Kommt nämlich einer der Vektoren $v_i$ mehrfach im Tupel vor, gilt also $v_j = v_i$ für ein $j \neq i$, dann können wir die Summe $\lambda_i v_i + \lambda_j v_j$ durch $(\lambda_i + \lambda_j)v_i$ ersetzen. Dies zeigt, dass wir $v_j$ aus dem Tupel streichen können, ohne dass die Eigenschaft von $v$, eine Linearkombination des Tupels zu sein, verloren geht. Die Gleichung

$$v = \sum_{i=1}^r \lambda_i v_i \quad \Leftrightarrow \quad (-1_K)v + \sum_{i=1}^r \lambda_i v_i = 0_V$$

zeigt nun, dass das Tupel $(v, v_1, ..., v_r)$ linear abhängig ist. Weil das Tupel aus lauter verschiedenen Elementen der Menge $S$ besteht, folgt daraus, dass $S$ linear abhängig ist.

„⇐" Ist $S$ linear abhängig, dann gibt es ein $r \in \mathbb{N}$ und ein linear abhängiges Tupel $(v_1, ..., v_r)$ bestehend aus lauter verschiedenen Vektoren der Menge $S$. Die lineare Abhängigkeit bedeutet, dass $\lambda_1, ..., \lambda_r \in K$ existieren, nicht alle gleich Null, mit $\sum_{i=1}^r \lambda_i v_i = 0_V$. Nehmen wir an, dass $i \in \{1, ..., r\}$ ein Index mit $\lambda_i \neq 0$ ist. Dann kann die Gleichung umgestellt werden zu

$$v_i \quad = \quad \sum_{j \neq i} \left( -\frac{\lambda_j}{\lambda_i} \right) v_j .$$

Dies zeigt, dass $v_i$ in $\mathrm{lin}(S \setminus \{v_i\})$ enthalten ist.

zu (ii) Nehmen wir an, dass $S$ linear unabhängig ist, dass $v \notin \mathrm{lin}(S)$ gilt, und dass $S \cup \{v\}$ dennoch linear abhängig ist. Dann gibt es ein $r \in \mathbb{N}$ und ein linear abhängiges Tupel $(v_1, ..., v_r)$ bestehend aus lauter verschiedenen Elementen der Menge $S \cup \{v\}$. Einer dieser Vektoren $v_i$ muss mit $v$ übereinstimmen, denn ansonsten wäre $(v_1, ..., v_r)$ ein linear abhängiges Tupel bestehend aus Elementen der Menge $S$, und $S$ somit linear abhängig, im Widerspruch zur Voraussetzung.

Sei $i \in \{1, ..., r\}$ der Index mit $v_i = v$. Auf Grund der linearen Abhängigkeit gibt es Koeffizienten $\lambda_1, ..., \lambda_r \in K$, nicht alle gleich Null, mit $\sum_{j=1}^r \lambda_j v_j = 0_V$. Dabei muss $\lambda_i \neq 0_K$ gelten, denn andernfalls wäre das Tupel ohne den Vektor $v_i$ ebenfalls linear abhängig, was erneut im Widerspruch zur linearen Unabhängigkeit von $S$ stehen würde. So aber können wir die Gleichung wieder zu

$$v \quad = \quad v_i \quad = \quad \sum_{j \neq i} \left( -\frac{\lambda_j}{\lambda_i} \right) v_j$$

umstellen. Aber dies steht im Widerspruch zur Voraussetzung $v \notin \mathrm{lin}(S)$. □

Die Aussage (i) der soeben bewiesenen Proposition wird häufig auch in der Kontraposition verwendet.

**(6.7) Folgerung** Sei $V$ ein $K$-Vektorraum. Eine Teilmenge $S \subseteq V$ ist genau dann linear unabhängig, wenn $v \notin \mathrm{lin}(S \setminus \{v\})$ für alle $v \in S$ erfüllt ist.

Am Ende dieses Kapitels sehen wir uns an, wie sich lineare Unabhängigkeit und die Existenz von Linearkombinationen rechnerisch nachweisen lässt. Sei $V = \mathbb{R}^3$ und $(v_1, v_2, v_3)$ das Tupel bestehend aus den drei Vektoren

$$v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad , \quad v_2 = \begin{pmatrix} 2 \\ 6 \\ 7 \end{pmatrix} \quad , \quad v_3 = \begin{pmatrix} 0 \\ 8 \\ 7 \end{pmatrix}.$$

Sei außerdem $v = (1, 0, 1)$ und $w = (3, 5, 7)$. Unser Ziel ist es zu überprüfen, ob $v$ bzw. $w$ Linearkombinationen des Tupels $(v_1, v_2, v_3)$ sind. Dass es sich bei $v$ um eine Linearkombination des Tupels handelt, ist äquivalent zur Existenz von Koeffizienten $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ mit

$$\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = v \quad \Leftrightarrow \quad \lambda_1 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ 6 \\ 7 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 8 \\ 7 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \Leftrightarrow$$

$$\begin{pmatrix} \lambda_1 + 2\lambda_2 \\ -\lambda_1 + 6\lambda_2 + 8\lambda_3 \\ 7\lambda_2 + 7\lambda_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

$$\Leftrightarrow \quad (\lambda_1, \lambda_2, \lambda_3) \text{ ist Lösungsmenge des LGS } x_1 + 2x_2 = 1 \ , \ -x_1 + 6x_2 + 8x_3 = 0 \ , \ 7x_2 + 7x_3 = 1.$$

Um zu sehen, ob das LGS eine Lösung hat, stellen wir die erweiterte Koeffizientenmatrix auf und bringen sie auf normierte Zeilenstufenform.

$$\begin{pmatrix} 1 & 2 & 0 & 1 \\ -1 & 6 & 8 & 0 \\ 0 & 7 & 7 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 8 & 8 & 1 \\ 0 & 7 & 7 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 7 & 7 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & -2 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Die dritte Zeile in der letzten Matrix entspricht der Gleichung $0 = 1$. Das LGS ist also *unlösbar*. Auf Grund der oben formulierten Äquivalenz folgt daraus, dass $v$ *keine* Linearkombination des Tupels $(v_1, v_2, v_3)$ ist.

Betrachten wir nun an Stelle von $v$ den Vektor $w$. Hier führt die Gleichung $\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = w$ nach dem gleichen Schema auf das LGS

$$x_1 + 2x_2 = 3 \ , \ -x_1 + 6x_2 + 8x_3 = 5 \ , \ 7x_1 + 7_x 3 = 7.$$

Wieder stellen wir die erweiterte Koeffizientenmatrix auf und formen auf normierte ZSF um.

$$\begin{pmatrix} 1 & 2 & 0 & 3 \\ -1 & 6 & 8 & 5 \\ 0 & 7 & 7 & 7 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & 0 & 3 \\ 0 & 8 & 8 & 8 \\ 0 & 7 & 7 & 7 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & 0 & 3 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & 0 & 3 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & -2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Die Gleichungen in der letzten Matrix lauten $x_1 - 2x_3 = 1$ und $x_2 + x_3 = 1$, was zu $x_1 = 1 + 2x_3$ und $x_2 = 1 - x_3$ umgeformt werden kann. Die Lösungsmenge $\mathscr{L}$ des ursprünglichen LGS ist also gegeben durch

$$\mathscr{L} \quad = \quad \left\{ \begin{pmatrix} 1 + 2x_3 \\ 1 - x_3 \\ x_3 \end{pmatrix} \ \middle| \ x_3 \in \mathbb{R} \right\} \quad = \quad \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \ \middle| \ \lambda \in \mathbb{R} \right\}.$$

Jedes Element der Lösungsmenge liefert eine Darstellung von $w$ als Linearkombination des Tupels $(v_1, v_2, v_3)$. Setzt man $\lambda = 0$, so erhält man zum Beispiel das Element $(1,1,0) \in \mathscr{L}$, und $\lambda = 1$ entspricht $(3,0,1) \in \mathscr{L}$. Tatsächlich gilt sowohl

$$1 \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 2 \\ 6 \\ 7 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 \\ 8 \\ 7 \end{pmatrix} \quad = \quad \begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix} \qquad \text{als auch} \qquad 3 \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 2 \\ 6 \\ 7 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 8 \\ 7 \end{pmatrix} \quad = \quad \begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix}.$$

Nach einem ähnlichen Schema lässt sich auch die lineare Unabhängigkeit behandeln. Diesmal betrachten wir in $V = \mathbb{R}^3$ das Tupel $(v_1, v_2, v_3)$ bestehend aus den Vektoren

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \quad , \quad v_2 = \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} \quad , \quad v_3 = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}.$$

Diesmal besteht unser Ziel darin, die lineare Unabhängigkeit von $(v_1, v_2, v_3)$ nachzuweisen. Für jedes Tripel $(\lambda_1, \lambda_2, \lambda_3)$ in $\mathbb{R}^3$ gilt die Äquivalenz

$$\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = 0_{\mathbb{R}^3} \quad \Leftrightarrow \quad \lambda_1 \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} + \lambda_3 \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad \begin{pmatrix} \lambda_1 + 3\lambda_3 \\ 2\lambda_1 + 2\lambda_2 + 3\lambda_3 \\ -\lambda_2 + 3\lambda_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Leftrightarrow \quad (\lambda_1, \lambda_2, \lambda_3) \text{ ist Lösung des LGS } x_1 + 3x_3 = 0 \ , \ 2x_1 + 2x_2 + 3x_3 = 0 \ , \ -x_2 + 3x_3 = 0.$$

Diesmal handelt es sich um ein *homogenes* LGS. Es genügt also, die nicht-erweiterte Koeffizientenmatrix aufzustellen und auf normierte ZSF zu bringen.

$$\begin{pmatrix} 1 & 0 & 3 \\ 2 & 2 & 3 \\ 0 & -1 & 3 \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} 1 & 0 & 3 \\ 0 & 2 & -3 \\ 0 & 1 & -3 \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -3 \\ 0 & 2 & -3 \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -3 \\ 0 & 0 & 3 \end{pmatrix}$$

$$\mapsto \quad \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Die letzte Matrix entspricht dem System $x_1 = 0 \ , \ x_2 = 0 \ , \ x_3 = 0$, die Lösungsmenge $\mathscr{L}$ des Systems ist also gleich $\{(0,0,0)\}$. Daraus ergibt sich insgesamt die Äquivalenz

$$\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = 0_{\mathbb{R}^3} \quad \Leftrightarrow \quad (\lambda_1, \lambda_2, \lambda_3) = (0,0,0) \quad \Leftrightarrow \quad \lambda_1 = \lambda_2 = \lambda_3 = 0.$$

Dies zeigt, dass das Tupel $(v_1, v_2, v_3)$ tatsächlich linear unabhängig ist; dafür ist bereits die Gültigkeit der Implikation „$\Rightarrow$“ hinreichend.

# § 7. Basis und Dimension

***Inhaltsübersicht***

Eine Teilmenge eines $K$-Vektorraums $V$ wird *Basis* genannt, wenn sie zugleich linear unabhängig und ein Erzeugendensystem von $V$ ist. Das wichtigste Ergebnis dieses Abschnitts lautet, dass je zwei Basen eines endlich erzeugten $K$-Vektorraums dieselbe Elementezahl besitzen; diese wird dann die *Dimension* von $V$ genannt.

Wir bestimmen Basen für eine Reihe konkreter $K$-Vektorräume (der $K^n$, Matrizen, Polynome). Außerdem leiten wir einige fundamentale Aussagen über $K$-Vektorräume her. Als wichtigste sind hier der *Basisauswahlsatz* und der *Basisergänzungssatz* zu nennen: Aus jedem Erzeugendensystem eines $K$-Vektorraums kann eine Basis ausgewählt werden, und jede linear unabhängige Teilmenge kann zu einer Basis ergänzt werden.

***Wichtige Definitionen und Sätze***

- (geordnete) Basis eines $K$-Vektorraums $V$
- Austauschsatz
- Je zwei Basen eines endlich erzeugten $K$-Vektorraums haben gleich viele Elemente.
- Dimension eines $K$-Vektorraums
- Basisauswahlsatz und Basisergänzungssatz

**(7.1) Definition** Sei $V$ ein $K$-Vektorraum. Eine Teilmenge $B \subseteq V$ heißt ***Basis*** von $V$, wenn sie linear unabhängig und ein Erzeugendensystem von $V$ ist. Ein Tupel $(v_1, ..., v_n)$ mit $n \in \mathbb{N}_0$, bestehend aus Vektoren $v_i \in V$, wird ***geordnete Basis*** genannt, wenn $\{v_1, ..., v_n\}$ aus $n$ verschiedenen Elementen besteht und eine Basis von $V$ bildet.

Im letzten Kapitel haben wir mehrere Beispiele für Teilmengen eines Vektorraums $V$ gesehen, die einerseits den Vektorraum $V$ erzeugen und andererseits auch linear unabhängig sind. Beispielsweise ist durch die Menge $\{e_1, ..., e_n\}$ der Einheitsvektoren eine Basis des $K^n$ gegeben. Die Menge $\{B_{k\ell}^{(m\times n)} \mid 1 \le k \le m, 1 \le \ell \le n\}$ der Basismatrizen bildet eine Basis des Vektorraums $\mathscr{M}_{m\times n,K}$ (daher der Name). Die Menge $\{x^n \mid n \in \mathbb{N}_0\}$ bildet ein (unendliche) Basis des $K$-Vektorraums $K[x]$.

**(7.2) Satz** Sei $V$ ein $K$-Vektorraum. Für eine Teilmenge $B \subseteq V$ sind die folgenden Aussagen äquivalent.

(i) Sie ist eine Basis von $V$.
(ii) Sie ist ein minimales Erzeugendensystem von $V$.
(iii) Sie ist eine maximale linear unabhängige Teilmenge von $V$.

***Beweis:*** „(i) $\Rightarrow$ (ii)“ Nehmen wir an, dass $B$ kein minimales Erzeugendensystem von $V$ ist. Dann gibt es eine Teilmenge $S \subsetneq B$ mit $V = \text{lin}(S)$. Wählen wir $v \in B \setminus S$ beliebig, dann gilt $v \in \text{lin}(S)$. Nach (6.6) (i) ist $S \cup \{v\}$ also linear abhängig. Wegen $B \supseteq S \cup \{v\}$ ist dann auch $B$ linear abhängig, im Widerspruch zur Voraussetzung.

„(ii) $\Rightarrow$ (iii)“ Gehen wir zunächst davon aus, dass $B$ linear abhängig ist. Dann gibt es nach (6.6) (i) ein $v \in B$ mit $v \in \text{lin}(B \setminus \{v\})$. Aus $B \setminus \{v\} \subseteq \text{lin}(B \setminus \{v\})$ und $v \in \text{lin}(B \setminus \{v\})$ folgt $B \subseteq \text{lin}(B \setminus \{v\})$. Mit Satz (6.3) folgt $V = \text{lin}(B) \subseteq \text{lin}(B \setminus \{v\})$, weil $\text{lin}(B \setminus \{v\})$ ein Untervektorraum von $V$ ist, und damit $V = \text{lin}(B \setminus \{v\})$ Aber dies steht im Widerspruch zu der Vorausetzung, dass $B$ ein *minimales* Erzeugendensystem von $V$ ist.

Nehmen wir nun an, $B$ ist zwar linear unabhängig, aber als linear unabhängige Teilmenge nicht maximal. Dann gibt es eine linear unabhängige Teilmenge $S$ von $V$ mit $S \supsetneq B$. Sei nun $v$ ein beliebiges Element in $S \setminus B$. Wegen $V = \text{lin}(B)$ gilt $v \in \text{lin}(B)$ und wegen $B \subseteq S \setminus \{v\}$ damit erst recht $v \in \text{lin}(S \setminus \{v\})$. Nach (6.6) (i) bedeutet dies, dass $S$ linear abhängig ist. Unsere Annahme hat also erneut zu einem Widerspruch geführt.

„(iii) $\Rightarrow$ (i)“ Nehmen wir an, dass $B$ kein Erzeugendensystem von $V$ ist. Dann existiert ein $v \in V \setminus \text{lin}(B)$. Nach (6.6) (ii) ist deshalb mit $B$ auch $B \cup \{v\}$ linear unabhängig. Aber dies widerspricht der Voraussetzung, dass $B$ maximal als linear unabhängige Teilmenge von $V$ ist. □

**(7.3) Satz** *(Austauschsatz)*
Sei $V$ ein $K$-Vektorraum, $S \subseteq V$ eine linear unabhängige Teilmenge und $E \subseteq V$ ein Erzeugendensystem. Dann gibt es für jede endliche Teilmenge $T \subseteq S$ eine Teilmenge $F \subseteq E$ mit der Eigenschaft, dass $|F| = |T|$ gilt und $(S \setminus T) \cup F$ linear unabhängig ist.

***Beweis:*** Der Hauptteil des Beweises ist der Nachweis der folgenden Hilfsaussage: Für jede endliche Teilmenge $T \subseteq S \setminus E$ gibt es eine Teilmenge $F \subseteq E \setminus S$, so dass $|F| = |T|$ gilt und $(S \setminus T) \cup F$ linear unabhängig ist. Wir beweisen diese Aussage durch vollständige Induktion über $n = |T|$. Ist $n = 0$, dann folgt $T = \varnothing$. Setzen wir $F = \varnothing$, dann gilt $|F| = 0 = |T|$. Außerdem gilt $(S \setminus T) \cup F = S$, also ist diese Menge linear unabhängig.

Sei nun $n \in \mathbb{N}_0$, und setzen wir die Aussage für $n$ voraus. Sei $T \subseteq S \setminus E$ eine $(n+1)$-elementige Teilmenge und $v \in T$ ein beliebiges Element. Setzen wir $T' = T \setminus \{v\}$, dann gilt $|T'| = n$, und nach Induktionsvoraussetzung existiert eine Teilmenge $F' \subseteq E \setminus S$ mit $|F'| = n$ und der Eigenschaft, dass $(S \setminus T') \cup F'$ linear unabhängig ist. Setzen wir $S' = (S \setminus T) \cup F'$, dann gilt $(S \setminus T') \cup F' = S' \cup \{v\}$ und $v \notin S'$, wegen $v \notin S \setminus T$ und $v \notin F' \subseteq E \setminus S$.

Nehmen wir nun zunächst an, dass ein $w \in E \setminus S'$ mit der Eigenschaft existiert, dass $S' \cup \{w\}$ linear unabhängig ist. Dann gilt insbesondere $w \notin F'$ (ansonsten wäre $w$ auch Element von $S'$, was ja ausgeschlossen wurde). Wäre $w \in S$, dann würde wegen $w \notin S'$ folgen, dass $w$ in $T$ liegt. Aber dies ist wegen $T \subseteq S \setminus E$ und $w \in E$ unmöglich. Setzen wir $F = F' \cup \{w\}$, dann gilt somit $F \subseteq E \setminus S$ und $|F| = |F'| + 1 = n + 1 = |T|$. Außerdem ist $(S \setminus T) \cup F = S' \cup \{w\}$ linear unabhängig. Der Induktionsschritt wäre damit abgeschlossen.

Existiert andererseits kein solches $w$, dann ist $S' \cup \{w\}$ für alle $w \in E \setminus S'$ linear abhängig. Nach (6.6) (ii) folgt daraus $w \in \mathrm{lin}(S')$ für alle $w \in E \setminus S'$. Wir erhalten $E \setminus S' \subseteq \mathrm{lin}(S')$, zusammen mit $S' \subseteq \mathrm{lin}(S')$ also $E \subseteq \mathrm{lin}(S')$. Weil $\mathrm{lin}(S')$ ein Untervektorraum von $V$ ist, folgt daraus wiederum $V = \mathrm{lin}(E) \subseteq \mathrm{lin}(S')$, nach Satz (6.3). Insbesondere wäre $v$ in $\mathrm{lin}(S')$ enthalten, wegen $v \notin S'$ die Menge $S' \cup \{v\} = (S \setminus T') \cup F'$ also linear abhängig, wiederum nach (6.6) (i). Aber dies widerspricht unserer Feststellung von oben.

Damit ist der Beweis der Hilfsaussage abgeschlossen. Sei nun $T \subseteq S$ eine endliche Teilmenge. Definieren wir $T' = T \setminus E$, dann besitzt $T$ die disjunkte Zerlegung $T = T' \cup (T \cap E)$, und außerdem gilt $T' \subseteq S \setminus E$. Auf Grund der Hilfsaussage existiert eine Teilmenge $F' \subseteq E \setminus S$ mit $|F'| = |T'|$ und der Eigenschaft, dass $(S \setminus T') \cup F'$ linear unabhängig ist. Sei nun $F = F' \cup (T \cap E)$. Wegen $F' \cap S = \varnothing$ ist auch dies eine disjunkte Zerlegung, und folglich gilt $|F| = |F'| + |T \cap E| = |T'| + |T \cap E| = |T|$. Außerdem gilt $(S \setminus T) \cup F = (S \setminus T) \cup (T \cap E) \cup F' = (S \setminus T') \cup F'$, also ist diese Menge linear unabhängig. □

Wir bezeichnen einen $K$-Vektorraum $V$ als ***endlich erzeugt***, wenn eine endliche Teilmenge $E \subseteq V$ mit $V = \mathrm{lin}(E)$ existiert.

**(7.4) Satz** Sei $V$ ein endlich erzeugter $K$-Vektorraum.

(i) In $V$ existiert eine endliche Basis $B$.

(ii) Für jede Basis $B'$ von $V$ gilt $|B'| = |B|$, insbesondere ist jede Basis endlich.

(iii) Ist $S \subseteq V$ linear unabhängig und $E \subseteq V$ ein Erzeugendensystem von $V$, dann gibt es eine Basis $B'$ mit $S \subseteq B' \subseteq E$.

***Beweis:*** zu (i) Nach Voraussetzung existiert eine endliche Teilmenge $E_0 \subseteq V$ mit $V = \mathrm{lin}(E_0)$. Sei $B$ eine *maximale* linear unabhängige Teilmenge von $E_0$; dies soll bedeuten, dass $B \cup \{v\}$ für alle $v \in E_0 \setminus B$ linear abhängig ist. Dann gilt $v \in \mathrm{lin}(B)$ für alle $v \in E_0 \setminus B$, nach (6.7) (ii). Aus $E_0 \setminus B \subseteq \mathrm{lin}(B)$ und $B \subseteq \mathrm{lin}(B)$ folgt $E_0 \subseteq \mathrm{lin}(B)$ und damit $V = \mathrm{lin}(E_0) \subseteq \mathrm{lin}(B)$ nach Satz (6.3), weil $\mathrm{lin}(B)$ ein Untervektorraum von $V$ ist. Also gilt $V = \mathrm{lin}(B)$. Damit ist $B$ linear unabhängig und zugleich ein Erzeugendensystem von $V$, insgesamt also eine Basis von $V$. Als Teilmenge von $E_0$ ist $B$ außerdem endlich.

zu (ii) Sei $B'$ eine weitere Basis von $V$. Zunächst zeigen wir, dass $|B'| \geq |B|$ gilt. Dazu wenden wir den Austauschsatz (7.3) auf $S = T = B$ und $E = B'$ an. Demzufolge existiert eine Teilmenge $F \subseteq B'$ mit $|F| = |B|$. Insbesondere gilt also $|B'| \geq |F| = |B|$.

Um die Endlichkeit von $B'$ zu beweisen, wählen wir in $B'$ eine beliebige Teilmenge $T'$ mit $|T'| = |B|$. Der Austauschsatz liefert uns eine Teilmenge $F' \subseteq B$ mit der Eigenschaft, dass $|F'| = |T'| = |B|$ gilt und $(B' \setminus T') \cup F'$ linear unabhängig ist. Wegen $F' \subseteq B$ und $|F'| = |B|$ gilt $F' = B$. Folglich ist die Menge $(B' \setminus T') \cup B$ linear unabhängig. Weil aber $B$ als Basis nach Satz (7.2) eine *maximale* linear unabhängige Teilmenge von $V$ ist, muss $B' \setminus T' \subseteq B$ gelten. Es folgt $|B'| \leq |B' \setminus T'| + |T'| = |B' \setminus T'| + |B| \leq |B| + |B| = 2|B|$, insbesondere ist $B'$ endlich.

Dasselbe Argument, dass oben die Ungleichung $|B| \le |B'|$ gezeigt hat, kann nun auch auf $B'$ an Stelle von $B$ angewendet werden, und ergibt damit die Abschätzung $|B'| \le |B|$. Insgesamt ist damit $|B'| = |B|$ gezeigt.

zu (iii) Sei $n = |B|$; wir zeigen zunächst, dass $|B'| \le n$ für jede linear unabhängige Menge mit $S \subseteq B' \subseteq E$ gelten muss. Wenden wir den Austauschsatz auf $B'$, eine beliebige *endliche* Teilmenge $T \subseteq B'$ und das Erzeugendensystem $B$ an, so erhalten wir eine Teilmenge $F \subseteq B$ mit $|F| = |T|$. Daraus folgt $n = |B| \ge |F| = |T|$. Jede endliche Teilmenge von $B'$ hat also eine Mächtigkeit $\le n$; dies ist nur möglich, wenn $B'$ endlich ist und $|B'| \le n$ gilt. Auf Grund der soeben bewiesenen Ungleichung gibt es in $E$ eine *maximale* linear unabhängige (und endliche) Teilmenge $B'$ mit $B' \supseteq S$. Wie in Teil (i) folgt dann, dass $B'$ eine Basis von $V$ ist; die Endlichkeit von $E_0$ spielte dort für den Nachweis von $\mathrm{lin}(B) = V$ keine Rolle. □

Aus Teil (iii) von Satz (7.4) ergibt sich unmittelbar

**(7.5) Folgerung** Sei $V$ ein endlich erzeugter $K$-Vektorraum.

(i) *(Basisergänzungssatz)*
Jede linear unabhängige Teilmenge $S \subseteq V$ kann zu einer Basis von $V$ ergänzt werden.

(ii) *(Basisauswahlsatz)*
Aus jedem Erzeugendensystem $E$ von $V$ kann man eine Basis von $V$ auswählen.

***Beweis:*** Für Aussage (i) genügt es, Satz (7.4) (iii) auf $S$ und $E = V$ anzuwenden. Für Aussage (ii) wendet man den Satz auf $S = \varnothing$ und $E$ an. □

**(7.6) Definition** Sei $V$ ein $K$-Vektorraum. Dann ist die ***Dimension*** von $V$ definiert durch

$$\dim V \quad = \quad \begin{cases} |B| & \text{falls } B \text{ eine endliche Basis von } V \text{ ist,} \\ \infty & \text{falls } V \text{ nicht endlich erzeugt ist.} \end{cases}$$

Man überprüfe anhand der bisherigen Resultate, dass die Dimension eines beliebigen $K$-Vektorraums damit wohldefiniert ist: Nach Satz (7.4) hat $V$ entweder eine endliche Basis, oder $V$ ist nicht endlich erzeugt. Im ersten Fall gilt außerdem für beliebig gewählte endliche Basen $B, B'$ von $V$ jeweils $|B| = |B'|$; es ist somit gleichgültig, welche Basis man für die Definition der Dimension heranzieht.

Wir bestimmen nun die Dimensionen der meisten uns bereits bekannten Vektorräume.

(i) Ist $K$ ein Körper und $n \in \mathbb{N}$, dann gilt $\dim K^n = n$. Denn wie wir bereits festgestellt haben, bilden die Einheitsvektoren $e_1, ..., e_n$ in $K^n$ eine $n$-elementige Basis.

(ii) Seien $K$ ein Körper und $m, n \in \mathbb{N}$. Dann gilt $\dim \mathscr{M}_{m\times n,K} = mn$, weil die Basismatrizen $B_{k\ell}^{(m\times n)}$ mit $1 \le k \le m$ und $1 \le \ell \le n$ eine $mn$-elementige Basis von $\mathscr{M}_{m\times n,K}$ bilden.

(iii) Wir wissen bereits, dass $\mathbb{C}$ auf natürliche Weise als $\mathbb{R}$-Vektorraum angesehen werden kann. Eine Basis dieses Vektorraums ist durch $\{1, i\}$ gegeben. Denn einerseits kann jedes $z \in \mathbb{C}$ auf Grund der Zerlegung in Real- und Imaginärteil in der Form $z = a \cdot 1 + b \cdot i$ mit $a, b \in \mathbb{R}$ dargestellt werden. Dies zeigt, dass $\{1, i\}$ ein Erzeugendensystem ist. Andererseits ist diese Darstellung auch eindeutig, denn aus $z = a \cdot 1 + b \cdot i$ mit $a, b \in \mathbb{R}$ folgt $a = \mathrm{Re}(z)$ und $b = \mathrm{Im}(z)$. Deshalb ist $\{1, i\}$ auch linear unabhängig. Für $\mathbb{C}$ als $\mathbb{R}$-Vektorraum gilt also $\dim \mathbb{C} = 2$; um zu verdeutlichen, dass $\mathbb{C}$ als $\mathbb{R}$-Vektorraum betrachtet wird, schreibt man $\dim_{\mathbb{R}} \mathbb{C} = 2$.

(iv) Fassen wir $\mathbb{C}$ dagegen als $\mathbb{C}$-Vektorraum auf, dann ist $\{1\}$ eine Basis, und es gilt $\dim \mathbb{C} = \dim_{\mathbb{C}} \mathbb{C} = 1$. Allgemein ist es nicht schwer zu zeigen, dass $\dim_{\mathbb{C}} \mathbb{C}^n = n$ und $\dim_{\mathbb{R}} \mathbb{C}^n = 2n$ für alle $n \in \mathbb{N}$ gilt; eventuell erledigen wir das in den Übungen.

(v) Ist $K$ ein Körper und $V$ ein $K$-Vektorraum mit $V = \{0_V\}$, dann gilt $\dim V = 0$, denn in diesem Fall ist die leere Menge $\varnothing$ eine nullelementige Basis von $V$. Tatsächlich ist $\varnothing$ linear unabhängig, und die einzige Linearkombination von $\varnothing$ ist der Nullvektor; es gilt also $\mathrm{lin}(\varnothing) = \{0_V\}$. Man kann sich leicht überlegen, dass umgekehrt aus $\dim V = 0$ jeweils $V = \{0_V\}$ folgt.

(vi) Für jeden Körper $K$ gilt $\dim K[x] = \infty$. Denn wie wir in § 6 festgestellt haben, besitzt $K[x]$ als $K$-Vektorraum kein endliches Erzeugendensystem.

Als weitere Konsequenz aus Satz (7.4) notieren wir noch

**(7.7) Folgerung** Sei $V$ ein endlich erzeugter $K$-Vektorraum und $n = \dim V$.

(i) Für jede linear unabhängige Teilmenge $S \subseteq V$ gilt $|S| \le n$ mit Gleichheit genau dann, wenn $S$ eine Basis von $V$ ist.

(ii) Für jedes Erzeugendensystem $E$ von $V$ gilt $|E| \ge n$ mit Gleichheit genau dann, wenn $E$ eine Basis von $V$ ist.

***Beweis:*** zu (i) Sei $S \subseteq V$ linear unabhängig. Nach dem Basisergänzungssatz gibt es eine Basis $B$ von $V$ mit $B \supseteq S$. Daraus folgt $|S| \le |B| = \dim V = n$. Beweisen wir nun die Äquivalenz. Gilt $|S| = n$, dann folgt aus $S \subseteq B$ und $|S| = n = |B|$ die Gleichheit $S = B$. In diesem Fall ist $S$ also selbst eine Basis. Setzen wir umgekehrt voraus, dass $S$ eine Basis von $V$ ist, dann muss $|S| = \dim V = n$ gelten, denn die Dimension von $V$ kann mit jeder beliebigen Basis bestimmt werden.

zu (ii) Sei $E \subseteq V$ ein Erzeugendensystem. Nach dem Basisauwahlsatz gibt es eine Basis $B \subseteq E$ von $V$. Daraus folgt $n = \dim V = |B| \le |E|$. Gilt $|E| = n$ dann folgt aus $|E| = n = |B|$ und $B \subseteq E$ die Gleichheit $E = B$. Also ist $E$ in diesem Fall selbst eine Basis. Setzen wir andererseits voraus, dass $E$ eine Basis von $V$ ist, dann muss $|E| = \dim V = n$ gelten, mit demselben Argument wie in Teil (i). □

Die soeben bewiesene Aussage kann folgendermaßen praktisch genutzt werden: Wenn man von einem endlich erzeugten $K$-Vektorraum $V$ die (endliche) Dimension $n$ bereits kennt und $T \subseteq V$ eine $n$-elementige Teilmenge ist, dann folgt aus der linearen Unabhängigkeit bereits die Basiseigenschaft von $T$. Ebenso folgt aus $V = \operatorname{lin}(T)$ bereits die Basiseigenschaft.

Beispielsweise ist wegen $\dim \mathbb{R}^3 = 3$ jede dreielementige linear unabhängige Teilmenge des $\mathbb{R}^3$ bereits eine Basis von $\mathbb{R}^3$, und ebenso ist jedes dreielementige Erzeugendensystem eine Basis. Andererseits zeigt die Folgerung auch, dass es in $\mathbb{R}^3$ kein zweielementiges Erzeugendensystem und keine vierelementige linear unabhängige Teilmenge geben kann.

**(7.8) Folgerung** Sei $V$ ein endlich erzeugter $K$-Vektorraum, $n = \dim V$ und $U$ ein Untervektorraum von $V$. Dann gilt $\dim U \le n$ mit Gleichheit genau dann, wenn $U = V$ gilt.

***Beweis:*** Sei $B$ eine Basis von $U$. Dann ist $B$ insbesondere eine linear unabhängige Teilmenge von $V$, und aus (7.7) (i) folgt $\dim U = |B| \le n$. Setzen wir $U = V$ voraus, dann folgt offenbar $\dim U = \dim V = n$. Sei nun umgekehrt $\dim U = n = \dim V$ vorausgesetzt. Dann ist $B$ wegen $|B| = \dim U$ eine $n$-elementige linear unabhängige Teilmenge von $V$. Aus (7.7) (i) folgt, dass $B$ eine Basis von $V$ ist. Somit gilt $U = \operatorname{lin}(B) = V$. □

Zum Abschluss des Kapitels soll eine praktische Umsetzung von Basisauswahlsatz und Basisergänzungssatz diskutiert werden. Konkret beantworten wir die folgenden beiden Fragen. Sei $V$ ein endlich-dimensionaler $K$-Vektorraum.

- Wenn $S = \{v_1, ..., v_r\}$ ein Erzeugendensystem von $V$ ist, wie findet man eine in $S$ enthaltene Basis?
- Wenn $S = \{v_1, ..., v_r\}$ eine $r$-elementige linear unabhängige Teilmenge von $V$ ist, wie lässt sich die Menge $S$ zu einer Basis von $V$ ergänzen?

Der folgende Satz liefert eine Antwort auf die erste Frage im Spezialfall $V = K^m$. Wir werden später sehen, dass jeder endlich-dimensionale $K$-Vektoraum $V \neq \{0_V\}$ isomorph zu $K^m$ für ein $m \in \mathbb{N}$ ist. Das Problem der Basisauswahl lässt sich dann leicht auf solche Vektorräume übertragen.

**(7.9) Satz** Seien $m, n \in \mathbb{N}$, und sei $S = \{v_1, ..., v_n\}$ eine $n$-elementige Teilmenge von $K^m$. Sei $A \in \mathscr{M}_{m\times n,K}$ die Matrix, deren Spalten genau die Vektoren $v_1, ..., v_n$ sind, und sei $A'$ die Matrix in normierter ZSF, die man durch Anwendung des Gauß-Verfahrens auf $A$ erhält. Seien $r, j_1, ..., j_r$ die Kennzahlen der ZSF. Dann ist $\{v_{j_1}, ..., v_{j_r}\}$ eine Basis von $\text{lin}(S)$.

***Beweis:*** Die gemeinsame Lösungsmenge $\mathscr{L} \subseteq K^n$ der homogenen linearen Gleichungssysteme mit den Koeffizientenmatrizen $A$ und $A'$ sind genau die Tupel $(\lambda_1, ..., \lambda_n)$ mit der Eigenschaft $\lambda_1 v_1 + ... + \lambda_n v_n = 0_{K^m}$. Sei $S = \{1, ..., n\} \setminus \{j_1, ..., j_r\}$. Wir betrachten nun das LGS zur Matrix $A'$. Für $1 \le k \le r$ enthält die Gleichung zur $k$-ten Zeile nach Definition der normierten ZSF jeweils den Term $x_{j_k}$ und ansonsten nur die Variablen $x_\ell$ mit $\ell \in S$ und $\ell > j_k$. Setzen wir ein $x_\ell$ mit $\ell \in S$ auf 1 und die die übrigen $x_{\ell'}$ mit $\ell' \in S$ auf Null, so erhalten wir jeweils ein Element $(\lambda_1, ..., \lambda_n)$ in $\mathscr{L}$ mit $\lambda_\ell = 1$ und $\lambda_{\ell'} = 0$ für alle $\ell' \in S$ mit $\ell' \neq \ell$. Die entsprechende Gleichung $\lambda_1 v_1 + ... + \lambda_n v_n = 0_{K^m}$ zeigt, dass $v_\ell$ eine Linearkombination der Vektoren $\{v_{j_1}, ..., v_{j_r}\}$ ist. Somit ist $T = \{v_{j_1}, ..., v_{j_r}\}$ ein in $S$ enthaltenes Erzeugendensystem von $K^m$.

Wäre $T$ keine Basis von $\text{lin}(S)$, also linear abhängig, dann müsste es nach (6.7) möglich sein, ein Element $v_{j_k}$ als Linearkombination der Vektoren $v_{j_s}$ mit $s \neq k$ darzustellen. Es gäbe dann in $\mathscr{L}$ ein Element der Form $(\lambda_1, ..., \lambda_n)$ mit $\lambda_{j_k} = 1$ und $\lambda_\ell = 0$ für alle $\ell \in S$. Setzt man diese Werte aber in die $k$-te Gleichung von $A'$ ein, so erhält man die Gleichung $1 \neq 0$. Dies zeigt, dass es in $\mathscr{L}$ kein derartiges Element und somit auch keine Darstellung von $v_{j_k}$ als Linearkombination der Vektoren $v_{j_s}$ mit $s \neq k$ existiert. □

Wir demonstrieren die Anwendung des Satzes an einem konkreten Beispiel. Unser Ziel ist es, aus der Teilmenge $S = \{v_1, v_2, v_3\}$ des $\mathbb{R}^3$ gegeben durch

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad , \quad v_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \text{und} \quad v_3 = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}$$

eine Basis des Vektorraums $V = \text{lin}(S)$ auszuwählen. Dazu tragen wir die Vektoren als Spalten in eine Matrix ein und formen auf normierte ZSF um.

$$\begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 3 & 1 & 4 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 1 & 2 \\ 0 & -1 & -1 \\ 0 & -2 & -2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & -2 & -2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Die normierte ZSF am Ende hat die Kennzahlen $r = 2$, $j_1 = 1$ und $j_2 = 2$. Nach Satz (7.9) ist somit $\{v_1, v_2\}$ eine Basis von $\text{lin}(S)$. Anhand der Lösungsmenge des zur Matrix gehörenden homogenen LGS lässt sich auch leicht erkennen, dass der Vektor $v_3$ als Linearkombination von $\{v_1, v_2\}$ dargestellt werden kann und somit für eine Basis von $\text{lin}(S)$ nicht benötigt wird. Die Matrix in normierter ZSF entspricht dem LGS bestehend aus den Gleichungen $x_1 + x_3 = 0$, $x_2 + x_3 = 0$. Die Lösungsmenge ist somit gegeben durch

$$\mathscr{L} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; x_1 = -x_3 \;,\; x_2 = -x_3 \right\} = \left\{ \begin{pmatrix} -x_3 \\ -x_3 \\ x_3 \end{pmatrix} \;\middle|\; x_3 \in \mathbb{R} \right\}$$

$$= \left\{ \begin{pmatrix} -x_3 \\ -x_3 \\ x_3 \end{pmatrix} \;\middle|\; x_3 \in \mathbb{R} \right\} = \left\{ x_3 \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \;\middle|\; x_3 \in \mathbb{R} \right\}.$$

Der Lösungsvektor $(-1,-1,1) \in \mathscr{L}$ liefert die Gleichung $(-1)v_1 + (-1)v_2 + v_3 = 0_{\mathbb{R}^3}$, was zu $v_3 = v_1 + v_2$ äquivalent ist.

Kommen wir nun zur Beantwortung der zweiten Frage. Gegeben sei eine linear unabhängige Teilmenge $S = \{v_1, ..., v_n\}$ im $K^m$, die zu einer Basis von $K^m$ ergänzt werden soll. Wir wissen bereits, dass die Menge $\{e_1, ..., e_m\}$ der Einheitsvektoren eine Basis und damit erst recht ein Erzeugendensystem von $K^m$ bildet. Also ist auch die Menge $T = \{v_1, ..., v_n, e_1, ..., e_m\}$ ein Erzeugendensystem von $K^m$. Mit dem in Satz (7.9) formulierten Kriterium kann aus $T$ eine Basis ausgewählt werden. Dabei ist nur zu beachten, dass die Vektoren $v_1, ..., v_n, e_1, ..., e_m$ tatsächlich in dieser Reihenfolge als Spalten in die Matrix $A$ eingetragen werden. Aus dem Beweis des Satzes geht hervor, dass nur solche Vektoren aus dem Erzeugendensystem entfernt werden, die als Linearkombination der Vektoren links davon darstellbar sind. Weil die Menge $S = \{v_1, ..., v_n\}$ linear unabhängig ist, ist kein $v_k$ mit $k \in \{1, ..., n\}$ als Linearkombination von $v_1, ..., v_{k-1}$ darstellbar. Dies bedeutet, dass keiner der Vektoren $v_k$ aus dem Erzeugendensystem entfernt wird. Somit ist gewährleistet, dass wir tatsächlich eine Basis von $K^m$ erhalten, die $S$ als Teilmenge enthält.

Auch die Basisergänzung demonstrieren wir an einem konkreten Beispiel. Wie man mit dem in (6.5) angegebenen Kriterium leicht überprüft, ist die Menge $S = \{v_1, v_2\}$ bestehend aus den Vektoren

$$v_1 = \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix} \quad \text{und} \quad v_2 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

linear unabhängig. Unser Ziel besteht darin, $S$ zu einer Basis von $\mathbb{R}^3$ zu ergänzen. Dazu schreiben wir die Vektoren $v_1, v_2, e_1, e_2, e_3$ als Spalten in eine Matrix und formen diese auf normierte ZSF um.

$$\begin{pmatrix} 2 & 1 & 1 & 0 & 0 \\ 2 & 2 & 0 & 1 & 0 \\ 3 & 3 & 0 & 0 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 2 & 2 & 0 & 1 & 0 \\ 3 & 3 & 0 & 0 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & -1 & 1 & 0 \\ 0 & \frac{3}{2} & -\frac{3}{2} & 0 & 1 \end{pmatrix} \mapsto$$

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & -\frac{3}{2} & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 & -\frac{2}{3} \end{pmatrix}$$

$$\mapsto \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & -1 & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 1 & -\frac{2}{3} \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 1 & 0 & -\frac{1}{3} \\ 0 & 1 & -1 & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 1 & -\frac{2}{3} \end{pmatrix}$$

Die normierte ZSF hat die Kennzahlen $r = 3$, $j_1 = 1$, $j_2 = 2$, $j_3 = 4$. Mit Satz (7.9) folgt daraus, dass $B = \{v_1, v_2, e_2\}$ eine Basis von $\mathbb{R}^3$ ist, die zudem $S$ als Teilmenge enthält. Ähnlich wie im vorherigen Beispiel findet man durch Bestimmung der Lösungsmenge $\mathscr{L} \subseteq \mathbb{R}^5$ des homogenen LGS zur umgeformten Matrix konkrete Darstellungen von $e_1$ und $e_3$ als Linearkombinationen der Basis $B$; es gilt

$$1 \cdot \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad = \quad \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

und

$$(-\tfrac{1}{3}) \cdot \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix} + \tfrac{2}{3} \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + (-\tfrac{2}{3}) \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad = \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

# § 8. Dimensionssätze

***Inhaltsübersicht***

In diesem Abschnitt beweisen wir zwei wichtige Sätze über die Dimension von Vektorräumen. Der *Schnittdimensionssatz* stellt einen Zusammenhang zwischen den Dimensionen von $W \cap W'$ und $W + W'$, falls $W$ und $W'$ Untervektorräume eines $K$-Vektorraums $V$ sind. Der Dimensionssatz für eine lineare Abbildung $\phi : V \to W$ besagt, dass sich die Dimension des Kerns und die Dimension des Bildes von $\phi$ immer zur Dimension von $V$ addieren. Dieser Satz lässt sich auch für Matrizen formulieren und liefert auf diese Weise Aussagen über die Lösungsmengen linearer Gleichungssysteme.

***Wichtige Definitionen und Sätze***

- Schnittdimensionssatz
- Dimensionssatz für lineare Abbildungen
- Zeilen- und Spaltenraum, Zeilen- und Spaltenrang einer Matrix
- Rangsatz: Für jede Matrix stimmen Zeilen- und Spaltenrang überein.
  (Es ist deshalb gerechtfertigt, einfach vom *Rang* einer Matrix zu sprechen.)
- Bezeichnet $\mathscr{L} \subseteq K^n$ den Lösungsraum eines homogenen LGS mit Darstellungsmatrix $A$,
  dann gilt $\dim \mathscr{L} = n - \mathrm{rg}(A)$.

**(8.1) Satz** *(Schnittdimensionssatz)*

Sei $V$ ein endlich erzeugter $K$-Vektorraum, und seien $W, W'$ Untervektorräume von $V$. Dann gilt

$$\dim(W + W') \quad = \quad \dim(W) + \dim(W') - \dim(W \cap W').$$

***Beweis:*** Sei $n = \dim(W \cap W')$ und $\{v_1, ..., v_n\}$ eine Basis von $W \cap W'$. Weil $W \cap W'$ ein Untervektorraum sowohl von $W$ als auch von $W'$ ist, gilt $\dim(W \cap W') \le \dim W$ und $\dim(W \cap W') \le \dim W'$ nach (7.8). Es gibt also $k, \ell \in \mathbb{N}_0$ mit $\dim W = n + k$ und $\dim W' = n + \ell$.

Weil $\{v_1, ..., v_n\}$ eine linear unabhängige Menge in $W$ ist, finden wir nach dem Basisergänzungssatz Vektoren $w_1, ..., w_k$, so dass $B = \{v_1, ..., v_n, w_1, ..., w_k\}$ eine Basis von $W$ ist. Ebenso finden wir Elemente $w'_1, ..., w'_\ell$ mit der Eigenschaft, dass die Familie $B' = \{v_1, ..., v_n, w'_1, ..., w'_\ell\}$ eine Basis von $W'$ ist. Der Satz ist bewiesen, wenn wir zeigen können, dass es sich bei

$$B_0 \quad = \quad B \cup B' \quad = \quad \{v_1, ..., v_n, w_1, ..., w_k, w'_1, ..., w'_\ell\}$$

um eine $n + k + \ell$-elementige Basis von $W + W'$ handelt, denn dann gilt

$$\dim(W + W') \quad = \quad n + k + \ell \quad = \quad (n + k) + (n + \ell) - n \quad = \quad \dim(W) + \dim(W') - \dim(W \cap W').$$

Zunächst zeigen wir, dass $B_0$ ein Erzeugendensystem von $W+W'$ ist. Jedes $v \in W+W'$ lässt sich in der Form $v = w+w'$ mit $w \in W$ und $w' \in W'$ schreiben. Da $\{v_1, ..., v_n, w_1, ..., w_k\}$ eine Basis von $W$ ist, finden wir Koeffizienten $\mu_i, \lambda_i \in K$ mit $w = \sum_{i=1}^n \mu_i v_i + \sum_{i=1}^k \lambda_i w_i$. Ebenso gibt es $\mu_i', \lambda_i' \in K$ mit $w' = \sum_{i=1}^n \mu_i' v_i + \sum_{i=1}^\ell \lambda_i' w_i'$. Insgesamt erhalten wir

$$v \quad = \quad w + w' \quad = \quad \sum_{i=1}^n (\mu_i + \mu_i')v_i + \sum_{i=1}^k \lambda_i w_i + \sum_{i=1}^\ell \lambda_i' w_i' \quad ,$$

also kann jedes $v \in W + W'$ tatsächlich als Linearkombination von $B_0$ dargestellt werden.

Als nächstes überprüfen wir, dass $B_0$ tatsächlich aus $n + k + \ell$ verschiedenen Elementen besteht. Besteht die Menge aus weniger Elementen, dann muss $w_i = w_j'$ für gewissen $i, j$ mit $1 \le i \le k$ und $1 \le j \le \ell$ gelten. Dies würde bedeuteten, dass $w_i$ in $W \cap W'$ enthalten ist. Damit wäre $w_i$ also in $\mathrm{lin}(v_1, ..., v_n)$ enthalten und die Menge $B$ damit linear abhängig, im Widerspruch zur Basis-Eigenschaft von $B$. Also ist $|B_0| = n + k + \ell$ erfüllt.

Nun beweisen wir die lineare Unabhängigkeit. Seien $\mu_i, \lambda_i, \lambda_i' \in K$ Koeffizienten mit

$$\sum_{i=1}^n \mu_i v_i + \sum_{i=1}^k \lambda_i w_i + \sum_{i=1}^\ell \lambda_i' w_i' \quad = \quad 0.$$

Sei $v = \sum_{i=1}^n \mu_i v_i + \sum_{i=1}^k \lambda_i w_i \in W$. Wegen $v = -\sum_{i=1}^\ell \lambda_i' w_i'$ liegt $v$ in $W \cap W'$. Weil $\{v_1, .., v_n\}$ eine Basis von $W \cap W'$ ist, gibt es auch $\alpha_1, ..., \alpha_n \in K$ mit $v = \sum_{i=1}^n \alpha_i v_i$. Es folgt

$$\sum_{i=1}^n (\mu_i - \alpha_i)v_i + \sum_{i=1}^k \lambda_i w_i \quad = \quad \left( \sum_{i=1}^n \mu_i v_i + \sum_{i=1}^k \lambda_i w_i \right) - \sum_{i=1}^n \alpha_i v_i \quad = \quad v - v \quad = \quad 0.$$

Auf Grund der linearen Unabhängigkeit von $B$ erhalten wir $\mu_i = \alpha_i$ für $1 \le i \le n$ und $\lambda_i = 0$ für $1 \le i \le k$. Setzen wir dies oben ein, so erhalten wir $\sum_{i=1}^n \mu_i v_i + \sum_{i=1}^\ell \lambda_i' w_i' = 0$. Wegen der linearen Unabhängigkeit von $B'$ folgt daraus wiederum $\lambda_i' = 0$ für $1 \le i \le \ell$ und $\mu_i = 0$ für $1 \le i \le n$. □

**(8.2) Folgerung** Sei $V$ ein endlich-dimensionaler $K$-Vektorraum, und seien $W, W'$ Untervektorräume von $V$, so dass $V = W \oplus W'$ erfüllt ist. Sei $B$ eine Basis von $W$ und $B'$ eine Basis von $W'$. Dann gilt

(i) $\dim V = \dim W + \dim W'$

(ii) Die Mengen $B$ und $B'$ sind disjunkt.

(iii) Die Vereinigung $B \cup B'$ ist eine Basis von $V$.

***Beweis:*** Sei $m = \dim W$ und $m' = \dim W'$. Nach Voraussetzung gilt $W \cap W' = \{0_V\}$, also $\dim(W \cap W') = 0$. Aus dem Schnittdimensionssatz folgt

$$\dim V \quad = \quad \dim W + \dim W' - \dim(W \cap W') \quad = \quad m + m' - 0 \quad = \quad m + m'.$$

Sei $B = \{w_1, ..., w_m\}$ eine Basis von $W$ und $B' = \{w'_1, ..., w'_{m'}\}$ eine Basis von $W'$. Wir zeigen, dass $E = B \cup B'$ ein Erzeugendensystem von $V$ ist. Sei $v \in V$ vorgegeben. Wegen $V = W + W'$ gibt es $w \in W$ und $w' \in W'$ mit $v = w + w'$. Weil $B$ eine Basis von $W$ und $B'$ eine Basis von $W'$ ist, gibt es $\lambda_1, ..., \lambda_m \in K$ und $\mu_1, ..., \mu_{m'} \in K$ mit

$$w = \sum_{k=1}^{m} \lambda_k w_k \qquad \text{und} \qquad w' = \sum_{k=1}^{m'} \mu_k w'_k.$$

Es folgt

$$v \quad = \quad w + w' \quad = \quad \sum_{k=1}^{m} \lambda_k w_k + \sum_{k=1}^{m'} \mu_k w'_k.$$

Dies zeigt, dass $E$ tatsächlich ein Erzeugendensystem von $V$ ist. Wegen $\dim V = m + m'$ besteht jedes Erzeugendensystem von $V$ aus mindestens $m + m'$ Elementen. Die Mengen $B$ und $B'$ sind also disjunkt, da ansonsten $|E| < m + m'$ gelten würde. Als $(m + m')$-elementiges Erzeugendensystem ist $E$ wegen $\dim V = m + m'$ eine Basis von $V$. □

Durch vollständige Induktion über $r$ erhält man

**(8.3) Folgerung** Sei $V$ ein $K$-Vektorraum, und seien $W_1, ..., W_r$ Untervektorräume von $V$ mit $V = \bigoplus_{k=1}^{r} W_k$. Dann gilt $\dim V = \sum_{k=1}^{r} \dim W_k$. Ist $B_k$ eine Basis von $W_k$ für $1 \le k \le r$, dann ist $B = \bigcup_{k=1}^{r} B_k$ eine Basis von $V$, und es gilt $B_k \cap B_\ell = \varnothing$ für $k \neq \ell$.

Als nächstes untersuchen wir die Vektorraum-Dimension im Zusammenhang mit linearen Abbildungen.

**(8.4) Satz** Seien $V, W$ endlich-dimensionale Vektorräume über einem Körper $K$, und sei $\phi : V \to W$ eine lineare Abbildung. Dann gilt

$$\dim V \quad = \quad \dim \ker(\phi) + \dim \operatorname{im}(\phi).$$

***Beweis:*** Sei $\{u_1, ..., u_m\}$ eine Basis von $\ker(\phi)$ und $\{w_1, ..., w_n\}$ eine Basis von $\operatorname{im}(\phi)$. Wir wählen für jedes $w_i$ einen Vektor $v_i \in V$ mit $\phi(v_i) = w_i$ und zeigen, dass durch

$$B \quad = \quad \{u_1, ..., u_m, v_1, ..., v_n\}$$

eine $(m+n)$-elementige Basis von $V$ gegeben ist. Haben wir dies gezeigt, dann ist damit $\dim V = m+n = \dim \ker(\phi) + \dim \operatorname{im}(\phi)$ bewiesen. Dass $B$ aus weniger als $m + n$ Elementen besteht ist nur möglich, wenn $u_i = v_j$ für gewisse $i, j$ mit $1 \le i \le m$ und $1 \le j \le n$ gilt. Aber dann wäre $w_j = \phi(v_j) = \phi(u_i) = 0_W$ im Widerspruch dazu, dass $w_j$ in einer Basis von $W$ liegt und somit ungleich Null sein muss.

Zunächst weisen wir nun nach, dass es sich bei $B$ um ein Erzeugendensystem von $V$ handelt. Sei dazu $v \in V$ vorgegeben. Da $\{w_1, ..., w_n\}$ eine Basis von $\mathrm{im}(\phi)$ ist, finden wir $\lambda_1, ..., \lambda_n \in K$ mit

$$\phi(v) \quad = \quad \sum_{i=1}^{n} \lambda_i w_i.$$

Aus der Linearität der Abbildung $\phi$ folgt $\phi(v) = \sum_{i=1}^{n} \lambda_i \phi(v_i) = \phi(v')$ mit $v' = \sum_{i=1}^{n} \lambda_i v_i$. Wegen $\phi(v) - \phi(v') = \phi(v-v') = 0_W$ liegt dann der Vektor $v - v'$ in $\ker(\phi)$. Da $\{u_1, ..., u_m\}$ eine Basis dieses Untervektorraums ist, existieren $\mu_1, ..., \mu_m \in K$ mit

$$v - v' = \sum_{j=1}^{m} \mu_j u_j \qquad \Leftrightarrow \qquad v = \sum_{j=1}^{m} \mu_j u_j + v' = \sum_{j=1}^{m} \mu_j u_j + \sum_{i=1}^{n} \lambda_i v_i.$$

Damit haben wir gezeigt, dass $B$ ein Erzeugendensystem von $V$ ist. Nun beweisen wir die lineare Unabhängigkeit. Seien $\mu_i, \lambda_j \in K$ mit

$$\sum_{i=1}^{m} \mu_i u_i + \sum_{j=1}^{n} \lambda_j v_j \quad = \quad 0_V$$

vorgegeben. Wenden wir die lineare Abbildung $\phi$ auf beide Seiten der Gleichung an, dann folgt

$$0_W \quad = \quad \phi(0_V) \quad = \quad \phi\left(\sum_{i=1}^{m} \mu_i u_i + \sum_{j=1}^{n} \lambda_j v_j\right) \quad = \quad 0_W + \sum_{j=1}^{n} \lambda_j \phi(v_j) \quad = \quad \sum_{j=1}^{n} \lambda_j w_j.$$

Dabei haben wir verwendet, dass die Summe $\sum_{i=1}^{m} \mu_i u_i$ in $\ker(\phi)$ enthalten ist. Weil die Menge $\{w_1, ..., w_n\}$ linear unabhängig ist, bedeutet dies $\lambda_1 = ... = \lambda_n = 0$. Setzen wir dies in die Ausgangsgleichung ein, dann erhält man $\sum_{i=1}^{m} \mu_i u_i = 0_V$. Da $\{u_1, ..., u_m\}$ nach Voraussetzung einer Basis von $\ker(\phi)$ und insbesondere linear unabhängig ist, hat dies wiederum $\mu_1 = ... = \mu_m = 0$ zur Folge. Also $B$ tatsächlich linear unabhängig. □

**(8.5) Folgerung** Für isomorphe Vektorräume $V, W$ gilt $\dim V = \dim W$.

***Beweis:*** Sei $\phi : V \to W$ ein Isomorphismus. Dann ist $\ker(\phi) = \{0_V\}$ und $\mathrm{im}(\phi) = W$. Also folgt die Aussage aus dem Dimensionssatz (8.4) für lineare Abbildungen. □

Wir werden den Dimensionssatz für lineare Abbildungen nun verwenden, um die Struktur von Matrizen genauer zu untersuchen.

**(8.6) Proposition** Sei $A = (a_{ij})$ eine $(m \times n)$-Matrix über $K$ und $\phi_A : K^n \to K^m$ die lineare Abbildung gegeben durch $v \mapsto Av$. Dann gilt

(i) Für $1 \le k \le n$ gilt $\phi_A(e_k) = a_{\bullet k}$. Die Bilder der Einheitsvektoren sind also genau die Spalten der Matrix.

(ii) Es gilt $\mathrm{im}(\phi_A) = \mathrm{lin}\{a_{\bullet 1}, ..., a_{\bullet n}\}$.

***Beweis:*** zu (i) Sei $k \in \{1, ..., n\}$ vorgegeben. Nach Definition des Matrix-Vektor-Produkts erhält man für jedes $i \in \{1, ..., m\}$ den $i$-ten Eintrag von $\phi_A(e_k)$ durch die Formel

$$\sum_{j=1}^{n} a_{ij}\delta_{jk} \quad = \quad a_{ik}\delta_{kk} \quad = \quad a_{ik}.$$

Dies ist genau der $i$-te Eintrag des $k$-ten Spaltenvektors $a_{\bullet k}$ der Matrix.

zu (ii) Sei $v \in K^n$ beliebig vorgegeben, $v = (\lambda_1, ..., \lambda_n)$. Da $\phi_A$ eine lineare Abbildung ist, gilt

$$\phi_A\left(\sum_{k=1}^{n} \lambda_k e_k\right) \quad = \quad \sum_{k=1}^{n} \lambda_k \phi_A(e_k) \quad = \quad \sum_{k=1}^{n} \lambda_k a_{\bullet k}.$$

Damit ist $\mathrm{im}(\phi_A) \subseteq \mathrm{lin}\{a_{\bullet 1}, ..., a_{\bullet n}\}$ nachgewiesen. Andererseits ist $\mathrm{im}(\phi_A)$ ein Untervektorraum von $K^m$, der nach Teil (i) die Menge $\{a_{\bullet 1}, ..., a_{\bullet n}\}$ der Spaltenvektoren enthält. Nach (6.3) (ii) folgt $\mathrm{lin}\{a_{\bullet 1}, ..., a_{\bullet n}\} \subseteq \mathrm{im}(\phi_A)$. □

**(8.7) Definition** Sei $A$ eine $(m \times n)$-Matrix über $K$.

(i) Der Untervektorraum $\mathrm{im}(\phi_A) = \mathrm{lin}\{a_{\bullet 1}, ..., a_{\bullet n}\}$ von $K^m$ wird der ***Spaltenraum*** der Matrix $A$ genannt und von uns mit $\mathrm{SR}(A)$ bezeichnet. Die Dimension $\mathrm{sr}(A) = \dim \mathrm{SR}(A)$ nennt man den ***Spaltenrang*** von $A$.

(ii) Ebenso nennt man den Untervektorraum von $K^n$ gegeben durch $\mathrm{lin}\{a_{1\bullet}, ..., a_{m\bullet}\}$ den ***Zeilenraum*** von $A$ und bezeichnet ihn mit $\mathrm{ZR}(A)$. Die Dimension $\mathrm{zr}(A) = \dim \mathrm{ZR}(A)$ wird ***Zeilenrang*** von $A$ genannt.

Zur Illustration dieser neuen Begriffe betrachten wir die Matrix

$$A \quad = \quad \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}.$$

Der Zeilenraum von $A$ wird aufgespannt von den Zeilenvektoren der Matrix $A$, es gilt also

$$\mathrm{ZR}(A) \quad = \quad \mathrm{lin}\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} \right\}.$$

Mit dem Kriterium (6.5) sieht man leicht, dass die Menge $\{(1,2,3),(4,5,6)\}$ linear unabhängig ist. Somit ist diese Menge eine Basis des Zeilenraums von $A$, und es folgt $\mathrm{zr}(A) = \dim \mathrm{ZR}(A) = 2$. Der Spaltenraum von $A$ wird von den Spalten der Matrix aufgespannt, es gilt also

$$\mathrm{ZR}(A) \quad = \quad \mathrm{lin}\left\{ \begin{pmatrix} 1 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \end{pmatrix}, \begin{pmatrix} 3 \\ 6 \end{pmatrix} \right\}.$$

Die Menge $\{(1,4),(2,5),(3,6)\}$ ist linear abhängig, denn es gilt

$$\begin{pmatrix} 3 \\ 6 \end{pmatrix} \quad = \quad (-1)\cdot\begin{pmatrix} 1 \\ 4 \end{pmatrix} + 2\cdot\begin{pmatrix} 2 \\ 5 \end{pmatrix}$$

während $\{(1,4),(2,5)\}$ offenbar linear unabhängig ist. Dies zeigt, dass $\{(1,4),(2,5)\}$ eine Basis des Spaltenraums $\mathrm{SR}(A)$ ist, und es folgt $\mathrm{sr}(A) = \dim \mathrm{SR}(A) = 2$. Zeilen- und Spaltenraum haben also die gleiche Dimension, obwohl sie in unterschiedlichen Vektorräumen enthalten sind; nach Definition ist $\mathrm{ZR}(A) \subseteq \mathbb{R}^3$ und $\mathrm{SR}(A) \subseteq \mathbb{R}^2$. Die weiteren Ausführungen werden zeigen, dass diese Übereinstimmung kein Zufall ist.

Für Matrizen in normierter ZSF hatten wir den Zeilenrang bereits in § 2 definiert, siehe (2.1). Wir zeigen, dass in diesem Spezialfall die neu eingeführte Definition des Zeilenrangs mit der alten Definition übereinstimmt.

**(8.8) Proposition** Sei $A \in \mathscr{M}_{m\times n,K}$ eine Matrix in normierter Zeilenstufenform, mit $r$ und $j_1,...,j_r$ als Kennzahlen. Dann ist $r$ der Zeilenrang von $A$ im Sinne von Definition (8.7).

***Beweis:*** Wir zeigen, dass in der Matrix $A$ die Zeilenvektoren $a_{1\bullet},...,a_{r\bullet} \in K^n$ ungleich Null linear unabhängig sind und somit eine Basis des Zeilenraums $\mathrm{ZR}(A)$ bilden. Seien $\lambda_1,...,\lambda_r \in K$ vorgegeben, mit

$$\sum_{i=1}^{r} \lambda_i a_{i\bullet} \quad = \quad 0_{K^n}. \tag{8.1}$$

Nach Definition der normierten ZSF ist in der $j_k$-ten Spalte der Eintrag $a_{kj_k} = 1_K$ der einzige Eintrag ungleich Null. Insgesamt sind die Einträge der $j_k$-ten Spalte also gegeben durch $a_{ij_k} = \delta_{ik}$ für $1 \le i \le m$. Betrachtet man in der Gleichung (8.1) also jeweils die $j_k$-te Komponenten für $k = 1,...,r$, so erhält man die Gleichungen

$$\sum_{i=1}^{r} \lambda_i a_{ij_k} = 0_K \quad \Longleftrightarrow \quad \sum_{i=1}^{r} \lambda_i \delta_{ik} = 0_K \quad \Longleftrightarrow \quad \lambda_k = 0_K.$$

Damit ist die lineare Unabhängigkeit nachgewiesen. Nach Definition bilden die Zeilen von $A$ ein Erzeugendensystem von $\mathrm{ZR}(A)$, und dasselbe gilt auch für die Zeilen ungleich Null. Somit besitzt der Zeilenraum $\mathrm{ZR}(A)$ eine $r$-elementige Basis, und es folgt $\mathrm{zr}(A) = \dim \mathrm{ZR}(A) = r$. □

**(8.9) Proposition** Sei $A \in \mathscr{M}_{m\times n,K}$ und $i \in \{1,...,m\}$ eine Zeilennummer mit der Eigenschaft, dass die $i$-te Zeile von $A$ eine Linearkombination der übrigen $m-1$ Zeilen ist. Entsteht nun die Matrix $\bar{A} \in \mathscr{M}((m-1)\times n, K)$ aus $A$ durch Streichung der $i$-ten Zeile, dann gilt $\mathrm{ZR}(\bar{A}) = \mathrm{ZR}(A)$ (also insbesondere $\mathrm{zr}(\bar{A}) = \mathrm{zr}(A)$) und ebenso $\mathrm{sr}(\bar{A}) = \mathrm{sr}(A)$.

***Beweis:*** Nach Definition der Untervektorräume $\mathrm{ZR}(A)$ und $\mathrm{ZR}(\bar{A})$ gilt

$$\mathrm{ZR}(A) = \mathrm{lin}\{a_{1\bullet},...,a_{m\bullet}\} \qquad \text{und} \qquad \mathrm{ZR}(\bar{A}) = \mathrm{lin}\{a_{1\bullet},...,a_{i-1\bullet},a_{i+1\bullet},...,a_{m\bullet}\}.$$

Nach Voraussetzung enthält $\mathrm{ZR}(\bar{A})$ neben den Vektoren $a_{k\bullet}$ mit $k \neq i$ auch den $i$-ten Zeilenvektor $a_{i\bullet}$. Aus der Inklusion $\{a_{1\bullet},...,a_{m\bullet}\} \subseteq \mathrm{ZR}(\bar{A})$ und der Untervektorraum-Eigenschaft von $\mathrm{ZR}(\bar{A})$ folgt $\mathrm{ZR}(A) \subseteq \mathrm{ZR}(\bar{A})$. Die umgekehrte Inklusion $\mathrm{ZR}(\bar{A}) \subseteq \mathrm{ZR}(A)$ ist offensichtlich.

Wir betrachten nun die Abbildung $\pi : K^m \to K^{m-1}$, die aus jedem Vektor $c \in K^m$ die $i$-te Komponente entfernt, also $\pi(c) = (c_1, ..., c_{i-1}, c_{i+1}, ..., c_m)$ für $c = (c_1, ..., c_m) \in K^m$. Man überprüft unmittelbar, dass $\pi$ eine lineare Abbildung ist. Die Spalten der Matrix $A$ werden von $\pi$ auf die Spalten von $\bar{A}$ abgebildet. Durch Übergang zur eingeschränkten Abbildung $\phi = \pi|_{SR(A)}$ erhalten wir also eine surjektive lineare Abbildung $\phi : \mathrm{SR}(A) \to \mathrm{SR}(\bar{A})$.

Nun zeigen wir, dass $\ker(\phi) = \{0_{K^m}\}$ gilt. Weil die $i$-te Zeile von $A$ eine Linearkombination der übrigen Zeilen ist, gibt es Koeffizienten $\mu_j \in K$ mit

$$a_{i\bullet} \quad = \quad \sum_{k=1}^{i-1} \mu_k a_{k\bullet} + \sum_{k=i+1}^{m} \mu_k a_{k\bullet}.$$

Die Einträge der Matrix erfüllen also die Gleichungen $a_{ij} = \sum_{k=1}^{i-1} \mu_k a_{kj} + \sum_{k=i+1}^{m} \mu_k a_{kj}$ für $1 \le j \le n$. Die Spalten $w_1, ..., w_n$ von $A$ sind damit im Untervektorraum

$$W \quad = \quad \left\{ c \in K^m \;\middle|\; c_i = \sum_{k=1}^{i-1} \mu_k c_k + \sum_{k=i+1}^{m} \mu_k c_k \right\}$$

von $K^m$ enthalten, es gilt also $\mathrm{SR}(A) \subseteq W$. Sei nun $c \in \ker(\phi)$ vorgegeben. Es gilt $(c_1, ..., c_{i-1}, c_{i+1}, ..., c_m) = \phi(c) = (0, ..., 0)$, also $c_j = 0$ für $j \neq i$. Wegen $c \in W$ ist damit auch

$$c_i \quad = \quad \sum_{k=1}^{i-1} \mu_k c_k + \sum_{k=i+1}^{m} \mu_k c_k \quad = \quad \sum_{k=1}^{i-1} \mu_k \cdot 0 + \sum_{k=i+1}^{m} \mu_k \cdot 0 \quad = \quad 0 \quad , \quad \text{also} \quad c = 0_{K^m}.$$

Damit ist $\ker(\phi) = \{0_{K^m}\}$ bewiesen. Durch Anwendung des Dimensionssatzes für lineare Abbildungen auf $\phi$ erhalten wir $\mathrm{sr}(A) = \dim \mathrm{SR}(A) = \dim \ker(\phi) + \dim \mathrm{im}(\phi) = 0 + \dim SR(\bar{A}) = \mathrm{sr}(\bar{A})$. □

**(8.10) Satz** *(Rangsatz)*

Für jede Matrix $A \in \mathscr{M}_{m\times n,K}$ gilt $\mathrm{zr}(A) = \mathrm{sr}(A)$. Wir bezeichnen die Zahl $\mathrm{zr}(A)$ deshalb einfach als den ***Rang*** $\mathrm{rg}(A)$ der Matrix.

***Beweis:*** Sei $r = \mathrm{zr}(A)$. Nach dem Basisauswahlsatz können wir so lange Zeilen aus $A$ streichen, bis die verbleibenden $r$ Zeilen der Restmatrix $A' \in \mathscr{M}(r \times n, K)$ eine Basis von $\mathrm{ZR}(A)$ bilden. Durch wiederholte Anwendung von (8.9) erhalten wir $\mathrm{zr}(A) = \mathrm{zr}(A') = r$ und $\mathrm{sr}(A) = \mathrm{sr}(A')$. Wegen $\mathrm{SR}(A') \subseteq K^r$ und $\dim K^r = r$ gibt es in $\mathrm{SR}(A')$ keine linear unabhängige Teilmenge mit mehr als $r$ Elementen; es gilt also $\mathrm{sr}(A) = \mathrm{sr}(A') \le r = \mathrm{zr}(A)$. Anwendung derselben Abschätzung auf die transponierte Matrix ${}^{\mathrm{t}}A$ liefert $\mathrm{zr}(A) = \mathrm{sr}({}^{\mathrm{t}}A) \le \mathrm{zr}({}^{\mathrm{t}}A) = \mathrm{sr}(A)$, denn die Zeilen von $A$ sind die Spalten von ${}^{\mathrm{t}}A$ und umgekehrt. Insgesamt gilt also $\mathrm{zr}(A) = \mathrm{sr}(A)$. □

Wie man sich leicht überzeugt, wird der Zeilenrang einer Matrix durch elementare Zeilenumformungen nicht geändert. Denn jede Zeile in der Matrix *nach* einer solchen Umformung ist Linearkombination der Zeilen in der Matrix *vor* der Umformung. Der Zeilenrang kann also durch eine elementare Zeilenumformung nicht größer werden. Weil andererseits jede solche Umformung durch eine weitere elementare Zeilenumformung rückgängig gemacht werden kann, ist es ebenso unmöglich, dass der Zeilenrang kleiner wird.

Der Rang einer Matrix $A$ lässt sich leicht berechnen: Wie wir in § 2 gezeigt haben, lässt sich $A$ durch endlich viele Zeilenumformungen in eine Matrix $A'$ in normierter ZSF überführen. Seien $r$ und $j_1, ..., j_r$ die Kennzahlen dieser ZSF. Nach (8.8) ist $r$ der Zeilenrang von $A'$. Weil sich der Zeilenrang durch elementare Zeilenumformungen nicht ändert, ist $r$ auch der Zeilenrang und somit der Rang der Matrix $A$.

Den Kern der linearen Abbildung $\phi_A : K^n \to K^m$, $v \mapsto Av$ nennt man auch den ***Kern der Matrix*** $A$ und bezeichnet ihn mit $\ker(A)$. Aus dem Rangsatz und dem Dimensionssatz für lineare Abbildungen ergibt sich die folgende Formel für die Dimension von Lösungsmengen linearer Gleichungssysteme.

**(8.11) Satz** Sei $A \in \mathscr{M}_{m\times n,K}$ und $\mathscr{L} \subseteq K^n$ die Lösungsmenge des linearen Gleichungssystems

$$Ax \quad = \quad 0_{K^m}.$$

Dann gilt $\dim \mathscr{L} = n - \mathrm{rg}(A)$.

***Beweis:*** Wie oben sei $\phi_A : K^n \to K^m$ gegeben durch $\phi_A(v) = Av$. Nach Definition ist der Lösungsraum $\mathscr{L}$ gegeben durch $\mathscr{L} = \{x \in K^n \mid Ax = 0_{K^m}\} = \ker(\phi_A)$. Der Dimensionssatz (8.4) liefert $\dim\ker(\phi_A) + \dim\mathrm{im}(\phi_A) = \dim K^n = n$. Wie wir oben bereits festgestellt haben, ist $\mathrm{im}(\phi_A)$ genau der Spaltenraum $\mathrm{SR}(A)$ von $A$. Folglich gilt $\dim\mathrm{im}(\phi_A) = \dim\mathrm{SR}(A) = \mathrm{sr}(A)$ und somit $\dim\ker(\phi_A) + \mathrm{sr}(A) = n$. Auf Grund des Rangsatzes (8.10) dürfen wir den Spaltenrang $\mathrm{sr}(A)$ durch den Rang $\mathrm{rg}(A)$ ersetzen und erhalten somit insgesamt $\dim\mathscr{L} = \dim\ker(\phi_A) = n - \mathrm{sr}(A) = n - \mathrm{rg}(A)$. □

Damit ist nun auch klar, wie man eine Basis des Lösungsraums $\mathscr{L}$ von $Ax = 0_{K^m}$ erhält: Sei $A'$ die umgeformte Matrix in normierter ZSF mit Kennzahlen $r$ und $j_1, ..., j_r$, und sei $S = \{1, ..., n\} \setminus \{j_1, ..., j_r\}$. In § 2 wurde beschrieben, wie man mit Hilfe der Matrix $A'$ jedem $\ell \in S$ einen Vektor $b_\ell \in K^n$ zuordnet, so dass jeder Vektor $v \in \mathscr{L}$ dann als Linearkombination der Vektoren $b_\ell$ darstellbar ist. Damit ist $E = \{b_\ell \mid \ell \in S\}$ ein Erzeugendensystem von $\mathscr{L}$. Weil $\dim\mathscr{L} = n - \mathrm{rg}(A) = n - r = |S|$ mit de Anzahl der Elemente von $E$ übereinstimmt, muss $E$ nach (7.7) (ii) eine Basis von $\mathscr{L}$ sein.

Als weitere Anwendung des Dimensionssatzes zeigen wir noch

**(8.12) Satz** Sei $n \in \mathbb{N}$, und sei $\phi : V \to W$ eine lineare Abbildung zwischen Vektorräumen $V, W$ derselben Dimension $n$. Dann sind äquivalent

(i) Die Abbildung $\phi$ ist injektiv.
(ii) Sie ist surjektiv.
(iii) Sie ist bijektiv.

***Beweis:*** „(i) $\Rightarrow$ (ii)" Ist $\phi$ injektiv, dann gilt $\ker(\phi) = \{0_V\}$ nach (5.10). Es folgt $\dim\ker(\phi) = 0$, und der Dimensionssatz für lineare Abbildungen liefert $\dim\mathrm{im}(\phi) = \dim V - \dim\ker(\phi) = n - 0 = n$. Sei $B$ eine Basis von $\mathrm{im}(\phi)$. Dann ist $B$ eine $n$-elementige linear unabhängige Teilmenge von $W$ und wegen $\dim W = n$ nach (7.7) (i) eine Basis von $W$. Es folgt $\mathrm{im}(\phi) = \mathrm{lin}(B) = W$ und damit die Surjektivität von $\phi$.

„(ii) $\Rightarrow$ (iii)" Ist $\phi$ surjektiv, dann gilt $\mathrm{im}(\phi) = W$ und somit $\dim\mathrm{im}(\phi) = \dim W = n$. Der Dimensionssatz für lineare Abbildungen liefert $\dim\ker(\phi) = \dim V - \dim\mathrm{im}(\phi) = n - n = 0$. Es folgt $\ker(\phi) = \{0_V\}$, also ist $\phi$ injektiv und damit insgesamt bijektiv.

„(iii) $\Rightarrow$ (i)" Als bijektive Abbildung ist $\phi$ insbesondere injektiv. □

Kehren wir noch einmal zum Schnittdimensionssatz zurück, den wir zu Beginn des Kapitels behandelt haben. Auch hier stellt sich wieder die Frage nach einer konkreten Berechnungsmethode. Genauer: Ist $V$ ein endlich-dimensionaler $K$-Vektorraum und sind $U$ und $W$ Untervektorräume gegeben jeweils durch eine Basis, wie findet man Basen der Untervektorräume $U+W$ und $U \cap W$? Bei der Summe $U+W$ ist die Sache einfach: Wie man leicht sieht, bilden die Basen von $U$ und $W$ bilden zusammengenommen ein Erzeugendensystem von $U+W$, und mit dem Basisauswahlverfahren aus § 7 kommt man zu einer Basis von $U+W$. Beim Durchschnitt erhält man ein Rechenverfahren mit Hilfe der folgenden Aussage.

**(8.13) Proposition** Sei $V$ ein endlich-dimensionaler $K$-Vektorraum und seien $U$ und $W$ Untervektorräume mit $r = \dim U$ und $s = \dim W$. Es sei $\{u_1, ..., u_r\}$ eine Basis von $U$ und $\{w_1, ..., w_s\}$ eine Basis von $W$. Weiter definieren wir die Teilmenge $\mathscr{L} \subseteq K^{r+s}$ durch

$$\mathscr{L} \quad = \quad \left\{ (\lambda_1, ..., \lambda_r, \mu_1, ..., \mu_s) \mid \sum_{i=1}^{r} \lambda_i u_i + \sum_{j=1}^{s} \mu_j w_j = 0_V \right\}.$$

Dann gilt $U \cap W = \left\{ \sum_{i=1}^{r} \lambda_i u_i \;\middle|\; \lambda_1, ..., \lambda_r \in K \text{ , } \exists\, \mu_1, ..., \mu_s \text{ mit } (\lambda_1, ..., \lambda_r, \mu_1, ..., \mu_s) \in \mathscr{L} \right\}$.

***Beweis:*** „$\subseteq$" Ist $v \in U \cap W$, dann existieren Koeffizienten $\lambda_1, ..., \lambda_r$ und $\mu'_1, ..., \mu'_s$ in $K$ mit

$$\sum_{i=1}^{r} \lambda_i u_i \quad = \quad v \quad = \quad \sum_{j=1}^{r} \mu'_j w_j.$$

Setzen wir $\mu_j = -\mu'_j$ für $1 \le j \le s$, dann folgt $\sum_{i=1}^{r} \lambda_i u_i + \sum_{j=1}^{r} \mu_j w_j = 0_V$. Also ist $v = \sum_{i=1}^{r} \lambda_i u_i$ ein Element der Menge auf der rechten Seite der Gleichung. „$\supseteq$" Sei $v = \sum_{i=1}^{r} \lambda_i u_i$ ein Element der Menge rechts. Dann gibt es nach Definition Koeffizienten $\mu_1, ..., \mu_s \in K$, so dass $(\lambda_1, ..., \lambda_r, \mu_1, ..., \mu_s)$ in $\mathscr{L}$ liegt. Daraus wiederum folgt $\sum_{i=1}^{r} \lambda_i u_i + \sum_{j=1}^{s} \mu_j w_j = 0_V$ und somit $\sum_{i=1}^{r} \lambda_i u_i = -\sum_{j=1}^{s} \mu_j w_j \in U \cap W$. □

Mit Hilfe der Proposition erhält man das folgende Berechnungsverfahren für den Durchschnitt. Sei $V = K^m$, und seien $U$, $W$, $r$, $s$ sowie die Vektoren $u_i$ und $w_j$ wie in der Proposition definiert.

- Trage die Vektoren $u_1, ..., u_r, w_1, ..., w_s$ als Spalten in eine Matrix $A \in \mathscr{M}_{m\times(r+s),K}$ ein.
- Wende das Gaußsche Eliminationsverfahren an, um $A$ in eine Matrix $A' = (a'_{ij}) \in \mathscr{M}_{m\times(r+s),K}$ in normierter Zeilenstufenform umzuwandeln.
- Seien $b_1, ..., b_\ell \in K^{r+s}$ die Basisvektoren von $\mathscr{L}$, die (wie in §2 beschrieben) an der Matrix $A'$ abgelesen werden können. Definieren $v_k = \sum_{i=1}^r b_{ki}u_i$ für $1 \le k \le \ell$. Dann ist $\{v_1, ..., v_\ell\}$ eine Basis von $U \cap W$.

Wir überprüfen die Korrektheit des Verfahrens. Die Matrix $A$, die im ersten Schritt definiert wurde, entspricht einem linearen Gleichungssystem bestehend aus $n$ Gleichungen in $r+s$ Unbekannten, dessen Lösungsmenge genau mit der Menge $\mathscr{L}$ aus (8.13) übereinstimmt. In der Tat, ein Tupel $(\lambda_1, ..., \lambda_r, \mu_1, ..., \mu_s) \in K^{r+s}$ ist genau dann ein Element der Lösungsmenge, dann gilt $\sum_{i=1}^r \lambda_i u_{ik} + \sum_{j=1}^s \mu_j w_{jk} = 0$ für $1 \le k \le m$. Dabei ist die $k$-te Gleichung jeweils äquivalent dazu, dass die $k$-te Komponenten des Vektors $\sum_{i=1}^r \lambda_i u_i + \sum_{j=1}^s \mu_j w_j$ gleich Null ist. Alle Gleichungen zusammen sind also äquivalent zu $\sum_{i=1}^r \lambda_i u_i + \sum_{j=1}^s \mu_j w_j = 0_V$ und somit zu $(\lambda_1, ..., \lambda_r, \mu_1, ..., \mu_s) \in \mathscr{L}$. Weil die Spalten von $A$ ein Erzeugendensystem von $U + W$ durchlaufen, gilt außerdem $\mathrm{rg}(A) = \dim(U + W)$.

Wegen $\mathrm{rg}(A) = \mathrm{rg}(A') = t$ gilt $\dim \mathscr{L} = r + s - t$ nach Satz (8.11). Aus (8.13) folgt, dass $U \cap W$ genau das Bild von $\pi(\mathscr{L})$ unter der Abbildung $\mathbb{R}^r \to \mathbb{R}^m$, $(\lambda_1, ..., \lambda_r) \mapsto \sum_{i=1}^R \lambda_i u_i$ ist. Daraus, dass die im letzten Schritt definierten Vektoren $v_1, ..., v_\ell$ den Untervektorraum $U \cap W$ aufspannen. Auf Grund des Schnittdimensionssatzes (8.1) gilt

$$\ell \quad = \quad \dim \mathscr{L} \quad = \quad r+s-t \quad = \quad r+s-\mathrm{rg}(A) \quad = \quad \dim U + \dim W - \dim(U+W) \quad = \quad \dim(U \cap W).$$

Deshalb bilden diese Vektoren sogar eine Basis von $U \cap W$.

Wir demonstrieren das Verfahren an einem konkreten Beispiel. Berechnet werden soll der Durchschnitt $U \cap W$ der Untervektorräume

$$U = \mathrm{lin}\left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 2 \\ -3 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\} \qquad \text{und} \qquad W = \mathrm{lin}\left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \\ 5 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ -1 \\ -1 \end{pmatrix} \right\}$$

im $\mathbb{R}^4$. Wir setzen voraus, dass bereits überprüft wurde, dass es sich bei den angegebenen Erzeugendensystemen um Basen von $U$ und $W$ handelt. Der Anleitung von oben folgend, schreiben wir die Basisvektoren in eine Matrix und formen diese auf normierte ZSF um.

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

$$\begin{pmatrix} 1 & 0 & 1 & 1 & 2 \\ 0 & 1 & 1 & 0 & 3 \\ 0 & 0 & -2 & -1 & -9 \\ 0 & 0 & 3 & 5 & 8 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 1 & 1 & 2 \\ 0 & 1 & 1 & 0 & 3 \\ 0 & 0 & 1 & \frac{1}{2} & \frac{9}{2} \\ 0 & 0 & 3 & 5 & 8 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 1 & 1 & 2 \\ 0 & 1 & 1 & 0 & 3 \\ 0 & 0 & 1 & \frac{1}{2} & \frac{9}{2} \\ 0 & 0 & 0 & \frac{7}{2} & -\frac{11}{2} \end{pmatrix} \mapsto$$

$$\begin{pmatrix} 1 & 0 & 1 & 1 & 2 \\ 0 & 1 & 1 & 0 & 3 \\ 0 & 0 & 1 & \frac{1}{2} & \frac{9}{2} \\ 0 & 0 & 0 & 1 & -\frac{11}{7} \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 1 & 0 & \frac{25}{7} \\ 0 & 1 & 1 & 0 & 3 \\ 0 & 0 & 1 & 0 & \frac{37}{7} \\ 0 & 0 & 0 & 1 & -\frac{11}{7} \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 & -\frac{12}{7} \\ 0 & 1 & 0 & 0 & -\frac{16}{7} \\ 0 & 0 & 1 & 0 & \frac{37}{7} \\ 0 & 0 & 0 & 1 & -\frac{11}{7} \end{pmatrix}$$

Wie in § 2 kann an der letzten Matrix abgelesen werden, dass der Lösungsraum des LGS durch

$$\mathscr{L} \quad = \quad \operatorname{lin}\{(\tfrac{12}{7}, \tfrac{16}{7}, -\tfrac{37}{7}, \tfrac{11}{7}, 1)\} \quad = \quad \operatorname{lin}\{(12, 16, -37, 11, 7)\}$$

gegeben ist. Laut Verfahren ist somit auch $U \cap W$ eindimensional, und es gilt $U \cap W = \operatorname{lin}(v)$ mit dem Vektor

$$v \quad = \quad 12 \cdot \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix} + 16 \cdot \begin{pmatrix} 0 \\ -1 \\ 2 \\ -3 \end{pmatrix} + (-37) \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \quad = \quad \begin{pmatrix} -25 \\ -29 \\ 7 \\ -48 \end{pmatrix}.$$

# § 9. Koordinatenabbildungen und Darstellungsmatrizen

***Inhaltsübersicht***

Das Rechnen in endlich-dimensionalen Vektorräumen wird durch die Einführung von Koordinatenabbildungen erheblich erleichtert, weil sich hierdurch jede Rechnung auf den einfach zu handhabenden Vektorraum $K^n$ reduzieren lässt. Von großer praktischer Bedeutung ist die Umrechnung zwischen verschiedenen Koordinatensystemen; man denke zum Beispiel an die Verwendung unterschiedlicher Bezugssysteme in der Physik. Wir werden sehen, dass sich eine solche Umrechnung stets durch eine einfache Matrix-Vektor-Multiplikation bewerkstelligen lässt.

Genauso wie sich jedes Element eines endlich-dimensionalen Vektorraums $V$ nach Wahl einer Basis durch ein Element des $K^n$ beschreiben lässt, so kann eine lineare Abbildung zwischen zwei solchen Vektorräumen $V, W$ durch eine Matrix angegeben werden. Zuvor müssen hierfür allerdings auf $V$ *und* $W$ gewählt werden. Auch hier ist eine wichtige Frage, wie sich die Matrix ändert, wenn man auf $V$ oder auf $W$ zu einer anderen Basis übergeht.

***Wichtige Definitionen und Sätze***

- Koordinatenabbildung $\Phi_{\mathscr{B}}$ zu einer geordneten Basis.
- Jeder $n$-dimensionale $K$-Vektorraum ist isomorph zu $K^n$.
- Existenz- und Eindeutigkeitssatz für lineare Abbildungen
- lineare Abbildung $\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A)$ zu einer Matrix $A$
- Darstellungsmatrix $\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)$ zu einer linearen Abbildung $\phi$
- Rechenregeln für Darstellungsmatrizen (Verträglichkeit mit Addition, Verhalten bezüglich Komposition und Umkehrabbildung)
- Transformationsformel / Satz vom Basiswechsel

**(9.1) Proposition** Sei $V$ ein $n$-dimensionaler $K$-Vektorraum und $\mathscr{B} = (v_1, ..., v_n)$ eine geordnete Basis. Dann gibt es für jeden Vektor $v \in V$ ein eindeutig bestimmtes Tupel $(\lambda_1, ..., \lambda_n) \in K^n$ mit

$$v = \sum_{k=1}^{n} \lambda_k v_k.$$

***Beweis:*** Weil $B = \{v_1, ..., v_n\}$ ein Erzeugendensystem von $V$ ist, gibt es jedenfalls Elemente $\lambda_1, ..., \lambda_n \in K$, so dass $v = \sum_{k=1}^n \lambda_k v_k$ erfüllt ist. Nehmen wir nun an, dass $(\lambda_1, ..., \lambda_n)$ und $(\mu_1, ..., \mu_n)$ beides Tupel mit der Eigenschaft

$$\sum_{k=1}^{n} \lambda_k v_k = v = \sum_{k=1}^{n} \mu_k v_k$$

Dann folgt

$$\sum_{k=1}^{n} (\lambda_k - \mu_k) v_k = \sum_{k=1}^{n} \lambda_k v_k - \sum_{k=1}^{n} \mu_k v_k = v - v = 0_V.$$

Weil $B$ auch linear unabhängig ist, folgt $\lambda_k - \mu_k = 0_K$ und somit $\lambda_k = \mu_k$ für $1 \le k \le n$. Damit ist die Eindeutigkeit bewiesen. $\square$

**(9.2) Satz** Sei $n \in \mathbb{N}$, $V$ ein $n$-dimensionaler $K$-Vektorraum und $\mathscr{B} = (v_1, ..., v_n)$ eine geordnete Basis. Dann ist durch die Zuordnung

$$\Phi_{\mathscr{B}} : V \longrightarrow K^n \quad , \qquad \sum_{i=1}^{n} \lambda_i v_i \mapsto (\lambda_1, ..., \lambda_n)$$

ein Isomorphismus von $K$-Vektorräumen definiert. Wir nennen $\Phi_{\mathscr{B}}$ die ***Koordinatenabbildung*** bezüglich der Basis $\mathscr{B}$. Für jeden Vektor $v \in V$ sind $\Phi_{\mathscr{B}}(v) \in K^n$ die ***Koordinaten*** von $V$ bezüglich der geordneten Basis $\mathscr{B}$.

***Beweis:*** Aus (9.1) folgt unmittelbar, dass $\Phi_{\mathscr{B}}$ eine wohldefinierte bijektive Abbildung ist. Zu zeigen bleibt die Linearität. Seien $v, w \in V$ und $\lambda \in K$ vorgegeben. Sei $\Phi_{\mathscr{B}}(v) = (\lambda_1, ..., \lambda_n)$ und $\Phi_{\mathscr{B}}(w) = (\mu_1, ..., \mu_n)$. Dann gilt nach Definition der Koordinaten

$$v = \sum_{k=1}^{n} \lambda_k v_k \qquad \text{und} \qquad w = \sum_{k=1}^{n} \mu_k v_k.$$

Es folgt $v+w = \sum_{k=1}^{n} (\lambda_k + \mu_k) v_k$ und $\lambda v = \sum_{k=1}^{n} (\lambda \lambda_k) v_k$, also gilt $\Phi_{\mathscr{B}}(v+w) = (\lambda_1 + \mu_1, ..., \lambda_n + \mu_n) = \Phi_{\mathscr{B}}(v) + \Phi_{\mathscr{B}}(w)$ und $\Phi_{\mathscr{B}}(\lambda v) = (\lambda\lambda_1, ..., \lambda\lambda_n) = \lambda \Phi_{\mathscr{B}}(v)$. Damit ist die Linearität nachgewiesen. □

Wir bemerken noch, dass nach Definition der Koordinatenabbildung das Bild $\Phi_{\mathscr{B}}(v_j)$ des $j$-ten Basisvektors gerade der $j$-te Einheitsvektor $e_j$ ist. Es gilt nämlich $v_j = \sum_{k=1}^{n} \delta_{jk} v_k$, und die Koeffizienten $\delta_{jk}$ sind gerade die Komponenten des $j$-ten Einheitsvektors $e_j$.

**(9.3) Folgerung** Zwischen zwei beliebigen $K$-Vektorräumen derselben endlichen Dimension existiert ein Isomorphismus.

***Beweis:*** Seien $V$ und $W$ zwei $n$-dimensionale $K$-Vektorräume, und seien $\mathscr{B}, \mathscr{C}$ geordnete Basen von $V$ bzw. $W$. Dann erhält man durch Komposition der Isomorphismen $\Phi_{\mathscr{B}} : V \to K^n$ und $\Phi_{\mathscr{C}}^{-1} : K^n \to W$ insgesamt einen Isomorphismus $\Phi_{\mathscr{C}}^{-1} \circ \Phi_{\mathscr{B}} : V \to W$ zwischen $V$ und $W$. □

Für jedes $n \in \mathbb{N}$ sei $\mathscr{E}_n = (e_1, ..., e_n)$ die Basis des $K^n$ bestehend aus den Einheitsvektoren. Man nennt $\mathscr{E}_n$ auch die ***kanonische Basis*** von $K^n$. Für jeden Vektor $v = (v_1, ..., v_n)$ gilt $\Phi_{\mathscr{E}_n}(v) = v$, also $\Phi_{\mathscr{E}_n} = \mathrm{id}_{K^n}$. Dies folgt unmittelbar aus der Gleichung $v = \sum_{k=1}^{n} v_k e_k$ und der Definition von $\Phi_{\mathscr{E}_n}(v)$.

Wir geben einige konkrete Beispiele für $\mathscr{B}$-Koordinaten an.

(i) Sei $K = \mathbb{R}$, $V = \mathbb{R}^3$ und $\mathscr{E}_3 = (e_1, e_2, e_3)$ die geordnete Basis bestehend aus den Einheitsvektoren. Gesucht werden die $\mathscr{E}_3$-Koordinaten des Vektors $v = (1, 3, 5)$. Es gilt

$$v \quad = \quad \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} \quad = \quad 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 3 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 5 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Also ist $\Phi_{\mathscr{E}_3}(v) = (1, 3, 5)$.

(ii) Wieder sei $K = \mathbb{R}$ und $V = \mathbb{R}^3$, aber diesmal suchen wir die $\mathscr{B}$-Koordinaten von $v = (1, 3, 5)$ bezüglich der Basis

$$\mathscr{B} \quad = \quad \left( \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} \right).$$

Die Gleichung

$$v \quad = \quad \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} \quad = \quad 3 \cdot \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + (-5) \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \tfrac{5}{3} \cdot \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix}$$

zeigt, dass diese Koordianten durch $\Phi_{\mathscr{B}}(v) = (3, -5, \frac{5}{3})$ gegeben sind.

In Beispiel (ii) war es nicht schwierig, die Koeffizienten 3, $-5$ und $\frac{5}{3}$ durch Vergleich der einzelnen Komponenten direkt zu finden. Im Allgemeinen bestimmt man die Koordinaten durch Lösen eines linearen Gleichungssystems. Dazu macht man in der vorliegenden Situation den Ansatz

$$\begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} \quad = \quad \lambda_1 \cdot \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_3 \cdot \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix}$$

mit $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$. Diese Gleichung ist äquivalent zu

$$\begin{pmatrix} 2\lambda_1 + \lambda_2 \\ \lambda_1 \\ 3\lambda_3 \end{pmatrix} \quad = \quad \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}$$

und wird von genau den Tupeln $(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3$ erfüllt, die das inhomogene LGS bestehend aus den Gleichungen $2x_1 + x_2 = 1$, $x_1 = 3$, $3x_3 = 5$ lösen. Genauer gesagt besitzt dieses LGS genau *eine* Lösung, die wie immer auch mit dem Gauß-Algorithmus bestimmt werden kann.

Betrachten wir noch weitere Beispiele für $\mathscr{B}$-Koordinaten.

(iii) Sei $K = \mathbb{R}$ und $V$ der $\mathbb{R}$-Vektorraum der Polynome in $\mathbb{R}[x]$ vom Grad $\leq 2$. Dieser Raum besitzt $\mathscr{B} = (1, x, x^2)$ als geordnete Basis. Sei $f = x^2 - 2x + 1$. Es gilt

$$f \quad = \quad 1 \cdot 1 + (-2) \cdot x + 1 \cdot x^2$$

und somit $\Phi_{\mathscr{B}}(f) = (1, -2, 1)$.

(iv) Seien $K$, $V$ und $f$ wie im vorherigen Beispiel definiert; diesmal betrachten wir aber die geordnete Basis $\mathscr{C} = (1, x+1, x^2 + x)$. Hier gilt nun

$$f \quad = \quad x^2 - 2x + 1 \quad = \quad 4 \cdot 1 + (-3) \cdot (x+1) + 1 \cdot (x^2 + x)$$

und somit $\Phi_{\mathscr{C}}(f) = (4, -3, 1)$.

(v) Sei $K = \mathbb{C}$, $V = \mathbb{C}^2$ und $v = (3+i, 2-6i)$. Sei $\mathscr{E}_2 = (e_1, e_2)$ die Basis der Einheitsvektoren. Die Gleichung

$$v \quad = \quad \begin{pmatrix} 3+i \\ 2-6i \end{pmatrix} \quad = \quad (3+i) \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + (2-6i) \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

zeigt, dass $\Phi_{\mathscr{B}}(v) = (3+i, 2-6i)$ ist.

(vi) Diesmal betrachten wir $V = \mathbb{C}^2$ als $\mathbb{R}$-Vektorraum, es sei also $K = \mathbb{R}$. Dieser Vektorraum besitzt die geordnete Basis $\mathscr{B} = (e_1, ie_1, e_2, ie_2)$. Wieder sei $v = (3+i, 2-6i)$. Weil $\mathbb{C}^2$ als $\mathbb{R}$-Vektorraum vierdimensional ist, hat $v$ diesmal vier Koordinaten, die aber in $\mathbb{R}$ liegen. Die Gleichung

$$v \quad = \quad \begin{pmatrix} 3+i \\ 2-6i \end{pmatrix} \quad = \quad 3 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} i \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} + (-6) \cdot \begin{pmatrix} 0 \\ i \end{pmatrix}$$

zeigt, dass diese durch $\Phi_{\mathscr{B}}(v) = (3, 1, 2, -6)$ gegeben sind.

Als nächstes untersuchen wir nun den Zusammenhang zwischen Matrizen und linearen Abbildungen.

**(9.4) Satz** *(Existenz und Eindeutigkeit linearer Abbildungen)*

Sei $n \in \mathbb{N}$, $V$ ein $n$-dimensionaler und $W$ ein beliebiger $K$-Vektorraum. Sei $(v_1, ..., v_n)$ eine geordnete Basis von $V$ und $(w_1, ..., w_n)$ ein Tupel bestehend aus Elementen von $W$. Dann gibt es eine eindeutig bestimmte lineare Abbildung $\phi : V \to W$ mit $\phi(v_i) = w_i$ für $1 \leq i \leq n$.

***Beweis:*** Zunächst beweisen wir die Existenz einer solchen linearen Abbildung. Jeder Vektor $v \in V$ besitzt eine eindeutige Darstellung $v = \sum_{i=1}^n \lambda_i v_i$ als Linearkombination der Basis, mit $\lambda_i \in K$ für $1 \leq i \leq n$. Wir definieren das Bild $\phi(v)$ durch $\phi(v) = \sum_{i=1}^n \lambda_i w_i$. Die so definierte Abbildung ist in der Tat linear. Ist nämlich $v' \in V$ ein weiterer

Vektor mit der Darstellung $v' = \sum_{i=1}^n \lambda_i' v_i$ als Linearkombination, dann besitzt der Vektor $v + v'$ die Darstellung $v + v' = \sum_{i=1}^n (\lambda_i + \lambda_i')v_i$. Nach Definition der Abbildung $\phi$ erhalten wir

$$\phi(v+v') \quad = \quad \sum_{i=1}^n (\lambda_i + \lambda_i')w_i \quad = \quad \sum_{i=1}^n \lambda_i w_i + \sum_{i=1}^n \lambda_i' w_i \quad = \quad \phi(v) + \phi(v').$$

Ist $\lambda \in K$ beliebig, dann gilt $\lambda v = \sum_{i=1}^n \lambda\lambda_i v_i$ und folglich $\phi(\lambda v) = \sum_{i=1}^n \lambda\lambda_i w_i = \lambda \sum_{i=1}^n \lambda_i w_i = \lambda\phi(v)$. Damit ist die Linearität von $\phi$ nachgewiesen.

Es bleibt zu zeigen, dass $\phi$ durch die genannten Eigenschaften eindeutig bestimmt ist. Sei $\phi' : V \to W$ eine weitere lineare Abbildung mit $\phi'(v_i) = w_i$ für $1 \le i \le n$. Ist $v \in V$ dann ein beliebiger Vektor und $v = \sum_{i=1}^n \lambda_i v_i$ eine Darstellung als Linearkombination der Basis, dann folgt aus der Linearität beider Abbildungen

$$\phi(v) \quad = \quad \phi\left(\sum_{i=1}^n \lambda_i v_i\right) \quad = \quad \sum_{i=1}^n \lambda_i \phi(v_i) \quad = \quad \sum_{i=1}^n \lambda_i w_i \quad =$$

$$\sum_{i=1}^n \lambda_i \phi'(v_i) \quad = \quad \phi'\left(\sum_{i=1}^n \lambda_i v_i\right) \quad = \quad \phi'(v)$$

Also stimmen $\phi$ und $\phi'$ als Abbildungen überein. □

Wir verwenden diesen Satz, um jeder Matrix $A \in \mathcal{M}_{m\times n,K}$ eine lineare Abbildung zwischen Vektorräumen $V, W$ der Dimension $n$ und $m$ zuzuordnen.

**(9.5) Definition** Seien $V, W$ endlich-dimensionale $K$-Vektorräume und $\mathscr{A} = (v_1, ..., v_n)$, $\mathscr{B} = (w_1, ..., w_m)$ geordnete Basen von $V$ bzw. $W$. Ferner sei $A = (a_{ij})$ eine Matrix aus $\mathcal{M}_{m\times n,K}$, mit $n = \dim V$ und $m = \dim W$. Dann gibt es nach (9.4) eine eindeutig bestimmte lineare Abbildung

$$\phi : V \longrightarrow W \qquad \text{mit} \qquad \phi(v_j) = \sum_{i=1}^m a_{ij} w_i \quad \text{für} \quad 1 \le j \le n.$$

Wir bezeichnen diese Abbildung $\phi$ mit $\mathscr{L}_{\mathscr{B}}^{\mathscr{A}}(A)$ und nennen sie die ***lineare Abbildung zur Matrix*** $A$ bezüglich der Basen $\mathscr{A}$ und $\mathscr{B}$.

Auch diese Definition illustrieren wir durch zwei Beispiele.

(i) Sei $K = \mathbb{R}$, $V = \mathbb{R}^2$, $W = \mathbb{R}^3$, $\mathscr{E}_2 = (e_1, e_2)$ und $\mathscr{E}_3 = (e_1, e_2, e_3)$. Wir suchen die lineare Abbildung $\phi = \mathscr{L}_{\mathscr{E}_3}^{\mathscr{E}_2}(A)$ zur Matrix

$$A \quad = \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}.$$

Nach Definition erfüllt $\phi$ die Gleichungen $\phi(e_1) = 1\cdot e_1 + 3\cdot e_2 + 5\cdot e_3 = (1,3,5)$ und $\phi(e_2) = 2\cdot e_1 + 4\cdot e_2 + 6\cdot e_3 = (2,4,6)$. Damit kann das Bild von $\phi$ auch für jeden beliebigen Vektor angegeben werden, denn auf Grund der

Linearität von $\phi$ gilt

$$\phi\begin{pmatrix}x_1\\x_2\end{pmatrix} \quad = \quad \phi\left(x_1\cdot\begin{pmatrix}1\\0\end{pmatrix}+x_2\cdot\begin{pmatrix}0\\1\end{pmatrix}\right) \quad = \quad x_1\cdot\phi\begin{pmatrix}1\\0\end{pmatrix}+x_2\cdot\phi\begin{pmatrix}0\\1\end{pmatrix}$$

$$= \quad x_1\cdot\begin{pmatrix}1\\3\\5\end{pmatrix}+x_2\cdot\begin{pmatrix}2\\4\\6\end{pmatrix} \quad = \quad \begin{pmatrix}x_1+2x_2\\3x_1+4x_2\\5x_1+6x_2\end{pmatrix}.$$

(ii) Sei $K=\mathbb{R}$, $V=\mathbb{R}^2$, $W=\mathbb{R}^3$ und $A$ wie in Beispiel (i) definiert. Diesmal aber betrachten wir die Basen $\mathscr{A}=(v_1,v_2)$ und $\mathscr{B}=(w_1,w_2,w_3)$ bestehend aus den Vektoren

$$v_1=\begin{pmatrix}1\\1\end{pmatrix}\ ,\ v_2=\begin{pmatrix}1\\-1\end{pmatrix}\ ,\ w_1=\begin{pmatrix}0\\1\\1\end{pmatrix}\ ,\ w_2=\begin{pmatrix}1\\0\\1\end{pmatrix}\ ,\ w_3=\begin{pmatrix}1\\1\\0\end{pmatrix}.$$

Sei $\psi=\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A)$. Wieder können an den beiden Spalten von $A$ die Bilder der Basisvektoren abgelesen werden: Nach Definition gilt

$$\psi(v_1) \quad = \quad 1\cdot w_1+3\cdot w_2+5\cdot w_3 \quad = \quad 1\cdot\begin{pmatrix}0\\1\\1\end{pmatrix}+3\cdot\begin{pmatrix}1\\0\\1\end{pmatrix}+5\cdot\begin{pmatrix}1\\1\\0\end{pmatrix} \quad = \quad \begin{pmatrix}8\\6\\4\end{pmatrix}$$

und

$$\psi(v_2) \quad = \quad 2\cdot w_1+4\cdot w_2+6\cdot w_3 \quad = \quad 2\cdot\begin{pmatrix}0\\1\\1\end{pmatrix}+4\cdot\begin{pmatrix}1\\0\\1\end{pmatrix}+6\cdot\begin{pmatrix}1\\1\\0\end{pmatrix} \quad = \quad \begin{pmatrix}10\\8\\6\end{pmatrix}.$$

Mit diesen Informationen können wir auch die Bilder $\psi(e_1)$ und $\psi(e_2)$ der Einheitsvektoren ausrechnen. Aus $e_1=\frac{1}{2}v_1+\frac{1}{2}v_2$ folgt

$$\psi(e_1) \quad = \quad \psi(\tfrac{1}{2}v_1+\tfrac{1}{2}v_2) \quad = \quad \tfrac{1}{2}\psi(v_1)+\tfrac{1}{2}\psi(v_2) \quad = \quad \tfrac{1}{2}\cdot\begin{pmatrix}8\\6\\4\end{pmatrix}+\tfrac{1}{2}\cdot\begin{pmatrix}10\\8\\6\end{pmatrix} \quad = \quad \begin{pmatrix}9\\7\\5\end{pmatrix}\quad ,$$

und mit $e_2=\frac{1}{2}v_1+(-\frac{1}{2})v_2$ erhalten wir ebenso

$$\psi(e_2) \quad = \quad \psi(\tfrac{1}{2}v_1+(-\tfrac{1}{2})v_2) \quad = \quad \tfrac{1}{2}\psi(v_1)+(-\tfrac{1}{2})\psi(v_2) \quad = \quad \tfrac{1}{2}\cdot\begin{pmatrix}8\\6\\4\end{pmatrix}+(-\tfrac{1}{2})\cdot\begin{pmatrix}10\\8\\6\end{pmatrix} \quad = \quad \begin{pmatrix}-1\\-1\\-1\end{pmatrix}.$$

Damit können wir nun wieder das Bild jedes beliebigen Vektors angeben.

$$\psi\begin{pmatrix}x_1\\x_2\end{pmatrix} \quad = \quad \psi\left(x_1\cdot\begin{pmatrix}1\\0\end{pmatrix}+x_2\cdot\begin{pmatrix}0\\1\end{pmatrix}\right) \quad = \quad x_1\cdot\psi\begin{pmatrix}1\\0\end{pmatrix}+x_2\cdot\psi\begin{pmatrix}0\\1\end{pmatrix}$$

$$= \quad x_1\cdot\begin{pmatrix}9\\7\\5\end{pmatrix}+x_2\cdot\begin{pmatrix}-1\\-1\\-1\end{pmatrix} \quad = \quad \begin{pmatrix}9x_1-x_2\\7x_1-x_2\\5x_1-x_2\end{pmatrix}.$$

Man sieht, dass $\phi\neq\psi$ ist, obwohl unter (i) und (ii) dieselbe Matrix $A$ verwendet wurde.

Sind $V$ und $W$ endlich-dimensionale $K$-Vektorräume mit geordneten Basen $\mathscr{A}$ und $\mathscr{B}$, so kann jeder Matrix $A \in \mathscr{M}_{m\times n,K}$ also eine lineare Abbildung $\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A) : V \to W$ zugeordnet werdne, sofern $n = \dim V$ und $m = \dim W$ ist. Nun sehen wir uns an, wie man umgekehrt jeder linearen Abbildung $V \to W$ eine Matrix zuordnet.

**(9.6) Definition** Seien $V, W$ endlich-dimensionale $K$-Vektorräume und $\mathscr{A} = (v_1, ..., v_n)$, $\mathscr{B} = (w_1, ..., w_m)$ geordnete Basen von $V$ bzw. $W$. Sei $\phi : V \to W$ eine lineare Abbildung. Für jedes $j \in \{1, ..., n\}$ stellen wir $\phi(v_j)$ als Linearkombination von $\mathscr{B}$ dar; es gilt

$$\phi(v_j) \quad = \quad \sum_{i=1}^{m} a_{ij} w_i \qquad 1 \le j \le n$$

mit eindeutig bestimmten Koeffizienten $a_{ij} \in K$. Wir nennen $A = (a_{ij}) \in \mathscr{M}_{m\times n,K}$ die ***Darstellungsmatrix*** von $\phi$ bezüglich der Basen $\mathscr{A}, \mathscr{B}$ und bezeichnen sie mit $\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)$.

Die Darstellungsmatrix $\mathscr{M}_{m\times n,K}$ kann ausgerechnet werden, indem man für jeden Basisvektor $v_j$ aus $\mathscr{A}$ das Bild $\phi(v_j)$ als Linearkombination von $\mathscr{B}$ schreibt und die entsprechenden Koeffizienten als Spalten in die Darstellungsmatrix einträgt. Wieder betrachten wir zwei konkrete Beispiele.

(i) Es seien $K = \mathbb{R}$ und $V = W = \mathscr{M}_{2,\mathbb{R}}$, der $\mathbb{R}$-Vektorraum der reellen $2\times 2$-Matrizen. Wir betrachten die geordneten Basen $\mathscr{A} = \mathscr{B} = (B_{11}, B_{12}, B_{21}, B_{22})$ bestehend aus den Basismatrizen und die Abbildung $\phi : V \to W$ gegeben durch
$$X \mapsto \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} X \qquad \text{für} \quad X \in \mathscr{M}_{2,\mathbb{R}}.$$
Es ist leicht zu überprüfen, dass $\phi$ eine lineare Abbildung ist. Die Bilder der Elemente von $\mathscr{A}$ sind nun gegeben durch
$$\begin{aligned}
\phi(B_{11}) &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 3 & 0 \end{pmatrix} = 1\cdot B_{11} + 0\cdot B_{12} + 3\cdot B_{21} + 0\cdot B_{22} \\
\phi(B_{12}) &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 3 \end{pmatrix} = 0\cdot B_{11} + 1\cdot B_{12} + 0\cdot B_{21} + 3\cdot B_{22} \\
\phi(B_{21}) &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 4 & 0 \end{pmatrix} = 2\cdot B_{11} + 0\cdot B_{12} + 4\cdot B_{21} + 0\cdot B_{22} \\
\phi(B_{22}) &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 0 & 4 \end{pmatrix} = 0\cdot B_{11} + 2\cdot B_{12} + 0\cdot B_{21} + 4\cdot B_{22}.
\end{aligned}$$
Jede Rechnung liefert eine Spalte der Darstellungsmatrix $\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)$. Insgesamt ist die gesuchte Matrix gegeben durch
$$\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi) \quad = \quad \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \\ 3 & 0 & 4 & 0 \\ 0 & 3 & 0 & 4 \end{pmatrix}.$$

(ii) Sei $K = \mathbb{R}$, $V = \mathbb{R}^2$, $W = \mathbb{R}^3$ und $\phi_A : V \to W$, $v \mapsto Av$ die Matrix-Vektor-Multiplikation mit der Matrix

$$A \quad = \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}.$$

Unser Ziel ist die Bestimmung der Darstellungsmatrix $\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)$ bezüglich der geordneten Basen

$$\mathscr{A} = \left( \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right) \qquad \text{und} \qquad \mathscr{B} = \left( \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right).$$

Wir rechnen die Bilder der Elemente von $\mathscr{A}$ aus und stellen sie als Linearkombination von $\mathscr{B}$ dar.

$$\phi_A\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) \quad = \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad = \quad \begin{pmatrix} 3 \\ 7 \\ 11 \end{pmatrix} \quad = \quad (-4)\cdot\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (-4)\cdot\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 11\cdot\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\phi_A\left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right) \quad = \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad = \quad \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} \quad = \quad 0\cdot\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 0\cdot\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + (-1)\cdot\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Wieder liefert jede Rechnung eine Spalte von $\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)$. Insgesamt ist die gesuchte Matrix gegeben durch

$$\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi) \quad = \quad \begin{pmatrix} -4 & 0 \\ -4 & 0 \\ 11 & -1 \end{pmatrix}.$$

Der folgende Satz zeigt, dass man mit Hilfe der Darstellungsmatrix $\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)$ und den $\mathscr{A}$-Koordinaten eines Vektors $v \in V$ die $\mathscr{B}$-Koordinaten des Bildvektors $\phi(v) \in W$ ausrechnen kann.

**(9.7) Satz** Seien die Bezeichnungen wie in der Definition gewählt. Dann gilt

$$\Phi_{\mathscr{B}}(\phi(v)) \quad = \quad \mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)\Phi_{\mathscr{A}}(v) \qquad \text{für alle} \quad v \in V.$$

***Beweis:*** Sei $A = \mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)$ und $\phi_A : K^n \to K^m$ die Abbildung $v \mapsto Av$ gegeben durch das Matrix-Vektor-Produkt. Zum Beweis der Gleichung $\Phi_{\mathscr{B}} \circ \phi = \phi_A \circ \Phi_{\mathscr{A}}$ genügt es auf Grund des Existenz- und Eindeutigkeitssatzes zu zeigen, dass

$$(\Phi_{\mathscr{B}} \circ \phi)(v_j) \quad = \quad (\phi_A \circ \Phi_{\mathscr{A}})(v_j) \qquad \text{für} \quad 1 \le j \le n$$

erfüllt ist. Für die linke Seite der Gleichung gilt nach Definition

$$(\Phi_{\mathscr{B}} \circ \phi)(v_j) \quad = \quad \Phi_{\mathscr{B}}\left(\sum_{i=1}^m a_{ij} w_i\right) \quad = \quad \sum_{i=1}^m a_{ij}\Phi_{\mathscr{B}}(w_i) \quad = \quad \sum_{i=1}^m a_{ij} e_i \quad = \quad (a_{1j}, \ldots, a_{mj}).$$

Für die rechte Seite erhalten wir

$$(\phi_A \circ \Phi_{\mathscr{A}})(v_j) \quad = \quad \phi_A(e_j) \quad = \quad (a_{1j}, ..., a_{mj}) \quad ,$$

denn nach (8.6) sind die Bilder der Einheitsvektoren unter $\phi_A$ genau die Spalten der Matrix $A$. Damit ist gezeigt, dass die beiden Abbildungen übereinstimmen. □

Auch diesen Satz illustrieren wir durch eine kurze Rechnung. Wieder seien $V = \mathbb{R}^2$, $W = \mathbb{R}^3$ mit den geordneten Basen $\mathscr{A}$ und $\mathscr{B}$ aus dem letzten Beispiel vorgegeben, und wir betrachten dieselbe lineare Abbildung $\phi_A : V \to W$. Wir berechnen das Bild des Vektors $v = (5,3)$ auf zwei verschiedene Arten: einerseits durch direktes Einsetzen in die Definition, andererseits mit über den Koordinatenvektor $\Phi_{\mathscr{A}}(v)$ und die Darstellungsmatrix $\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi_A)$. Das direkte Einsetzen ergibt

$$\phi_A(v) \quad = \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 5 \\ 3 \end{pmatrix} \quad = \quad \begin{pmatrix} 11 \\ 27 \\ 43 \end{pmatrix}.$$

Der Koordinatenvektor von $v$ bezüglich $\mathscr{A}$ ist gegeben durch $\Phi_{\mathscr{A}} = (4,1)$, denn es gilt $(5,3) = 4 \cdot (1,1) + 1 \cdot (1,-1)$. Mit Satz (9.7) erhalen wir

$$\Phi_{\mathscr{B}}(\phi_A(v)) \quad = \quad \mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi_A)\Phi_{\mathscr{A}}(v) \quad = \quad \begin{pmatrix} -4 & 0 \\ -4 & 0 \\ 11 & -1 \end{pmatrix} \begin{pmatrix} 4 \\ 1 \end{pmatrix} \quad = \quad \begin{pmatrix} -16 \\ -16 \\ 43 \end{pmatrix}.$$

Aus den $\mathscr{B}$-Koordinaten von $\phi_A(v)$ können wir den Vektor $\phi_A(v)$ zurückgewinnen: Es ist

$$\phi_A(v) \quad = \quad (-16) \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (-16) \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 43 \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad = \quad \begin{pmatrix} 11 \\ 27 \\ 43 \end{pmatrix}$$

Beide Rechenwege führen also zum gleichen Ergebnis.

**(9.8) Proposition** Seien $m, n \in \mathbb{N}$, $A \in \mathscr{M}_{m \times n, K}$ und $\phi_A : K^n \to K^m$ die lineare Abbildung gegeben durch $\phi_A(v) = Av$ für alle $v \in K^n$. Dann gilt

$$\mathscr{M}^{\mathscr{E}_n}_{\mathscr{E}_m}(\phi_A) \quad = \quad A.$$

***Beweis:*** Für $1 \le j \le n$ gilt nach (9.7) jeweils

$$\mathscr{M}^{\mathscr{E}_n}_{\mathscr{E}_m}(\phi_A)e_j \quad = \quad \mathscr{M}^{\mathscr{E}_n}_{\mathscr{E}_m}(\phi_A)\Phi_{\mathscr{E}_n}(e_j) \quad = \quad \Phi_{\mathscr{E}_m}(\phi_A(e_j)) \quad = \quad \phi_A(e_j) \quad = \quad Ae_j.$$

Dies zeigt, dass die $j$-te Spalte von $\mathscr{M}^{\mathscr{E}_n}_{\mathscr{E}_m}(\phi_A)$ mit der $j$-ten Spalte von $A$ übereinstimmt, für $1 \le j \le n$. □

Durch den folgenden Satz wird nun der entscheidende Zusammenhang zwischen Matrizen und linearen Abbildungen hergestellt.

**(9.9) Satz** Seien $V, W$ endlich-dimensionale $K$-Vektorräume und $\mathscr{A}$, $\mathscr{B}$ geordnete Basen von $V$ bzw. $W$. Dann sind durch die beiden Abbildungen

$$\mathscr{M}_{m\times n,K} \to \mathrm{Hom}_K(V,W)\ ,\ A \mapsto \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A) \qquad , \qquad \mathrm{Hom}_K(V,W) \to \mathscr{M}_{m\times n,K}\ ,\ \phi \mapsto \mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)$$

zueinander inverse Isomorphismen von $K$-Vektorräumen definiert.

***Beweis:*** Sei $\mathscr{A} = (v_1,...,v_n)$ und $\mathscr{B} = (w_1,...,w_m)$. Zunächst beweisen wir, dass $\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}$ eine lineare Abbildung ist. Seien dazu $A, B \in \mathscr{M}_{m\times n,K}$ und $\lambda \in K$ beliebig vorgegeben, $A = (a_{ij})$, $B = (b_{ij})$. Für $1 \le j \le n$ gilt

$$\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A+B)(v_j) \quad = \quad \sum_{i=1}^m (a_{ij}+b_{ij})w_i \quad = \quad \sum_{i=1}^m a_{ij}w_i + \sum_{i=1}^m b_{ij}w_i \quad = \quad \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A)(v_j) + \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(B)(v_j)$$

und

$$\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(\lambda A)(v_j) \quad = \quad \sum_{i=1}^m (\lambda a_{ij})w_i \quad = \quad \lambda \sum_{i=1}^m a_{ij}w_i \quad = \quad \lambda \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A)(v_j) \quad = \quad (\lambda \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A))(v_j).$$

Damit sind die Gleichungen $\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A+B) = \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A) + \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(B)$ und $\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(\lambda A) = \lambda \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}$ bewiesen, und $\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}$ ist tatsächlich eine lineare Abbildung. Um zu zeigen, dass die Abbildungen $\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}$ und $\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}$ zueinander invers sind, müssen die Gleichungen

$$\mathscr{M}^{\mathscr{A}}_{\mathscr{B}} \circ \mathscr{L}^{\mathscr{A}}_{\mathscr{B}} = \mathrm{id}_{\mathscr{M}_{m\times n,K}} \qquad \text{und} \qquad \mathscr{L}^{\mathscr{A}}_{\mathscr{B}} \circ \mathscr{M}^{\mathscr{A}}_{\mathscr{B}} = \mathrm{id}_{\mathrm{Hom}_K(V,W)}$$

überprüft werden. Zum Beweis der ersten Gleichung sei $A = (a_{ij}) \in \mathscr{M}_{m\times n,K}$ vorgegeben. Es gilt

$$\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A)(v_j) \quad = \quad \sum_{i=1}^m a_{ij}w_i \qquad \text{für} \quad 1 \le j \le n.$$

Diese Gleichungen zeigen, dass $A$ die Darstellungsmatrix von $\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A)$ ist. Es gilt also

$$(\mathscr{M}^{\mathscr{A}}_{\mathscr{B}} \circ \mathscr{L}^{\mathscr{A}}_{\mathscr{B}})(A) \quad = \quad \mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A)) \quad = \quad A \quad = \quad \mathrm{id}_{\mathscr{M}_{m\times n,K}}(A).$$

Damit ist die Gleichung $\mathscr{M}^{\mathscr{A}}_{\mathscr{B}} \circ \mathscr{L}^{\mathscr{A}}_{\mathscr{B}} = \mathrm{id}_{\mathscr{M}_{m\times n,K}}$ nachgewiesen.

Zum Beweis der zweiten Gleichung sei $\phi \in \mathrm{Hom}_K(V,W)$ vorgegeben und $A = \mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)$ mit $A = (a_{ij})$. Dann gilt

$$\phi(v_j) \quad = \quad \sum_{i=1}^m a_{ij}w_i \quad = \quad \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A)(v_j) \qquad \text{für} \quad 1 \le j \le n.$$

Diese Gleichungen zeigen, dass die linearen Abbildungen $\phi$ und $\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A)$ auf der Basis $\mathscr{A}$ von $V$ übereinstimmen. Nach dem Existenz- und Eindeutigkeitssatz sind sie damit identisch. Es gilt also

$$(\mathscr{L}^{\mathscr{A}}_{\mathscr{B}} \circ \mathscr{M}^{\mathscr{A}}_{\mathscr{B}})(\phi) \quad = \quad \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)) \quad = \quad \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A) \quad = \quad \phi \quad = \quad \mathrm{id}_{\mathrm{Hom}_K(V,W)}(\phi) \quad ,$$

also $\mathscr{L}^{\mathscr{A}}_{\mathscr{B}} \circ \mathscr{M}^{\mathscr{A}}_{\mathscr{B}} = \mathrm{id}_{\mathrm{Hom}_K(V,W)}$. Insgesamt haben wir gezeigt, dass die Abbildungen $\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}$ und $\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}$ zueinander invers sind, damit insbesondere Umkehrabbildungen besitzen und folglich bijektiv sind. Auf Grund von (4.11) ist $\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}$ als Umkehrabbildung einer linearen Abbildung ebenfalls linear. □

**(9.10) Folgerung** Seien $V, W$ endlich-dimensionale $K$-Vektorräume. Dann gilt

$$\dim \operatorname{Hom}_K(V,W) \quad = \quad (\dim V)(\dim W).$$

***Beweis:*** Sei $m = \dim W$ und $n = \dim V$. Nach (9.9) sind die $K$-Vektorräume $\operatorname{Hom}_K(V,W)$ und $\mathscr{M}_{m\times n,K}$ isomorph. Daraus folgt $\dim \operatorname{Hom}_K(V,W) = \dim \mathscr{M}_{m\times n,K} = mn$. □

**(9.11) Lemma** Sei $V$ ein endlich-dimensionaler $K$-Vektorraum und $\mathscr{A}$ eine geordnete Basis von $V$. Dann gilt

$$\mathscr{M}^{\mathscr{A}}_{\mathscr{A}}(\mathrm{id}_V) \quad = \quad E^{(n)} \quad ,$$

d.h. die Darstellungsmatrix der identischen Abbildung ist die Einheitsmatrix.

***Beweis:*** Sei $\mathscr{A} = (v_1, ..., v_n)$. Für $1 \le j \le n$ gilt nach (9.7) dann jeweils

$$\mathscr{M}^{\mathscr{A}}_{\mathscr{A}}(\mathrm{id}_V)e_j \quad = \quad \mathscr{M}^{\mathscr{A}}_{\mathscr{A}}(\mathrm{id}_V)\Phi_{\mathscr{A}}(v_j) \quad = \quad \Phi_{\mathscr{A}}(\mathrm{id}_V(v_j)) \quad = \quad \Phi_{\mathscr{A}}(v_j) \quad = \quad e_j.$$

Die Spalten von $\mathscr{M}^{\mathscr{A}}_{\mathscr{A}}(\mathrm{id}_V)$ sind also gerade die Einheitsvektoren. Daraus folgt $\mathscr{M}^{\mathscr{A}}_{\mathscr{A}}(\mathrm{id}_V) = E^{(n)}$. □

**(9.12) Satz** Seien $U, V, W$ endlich-dimensionale $K$-Vektorräume mit geordneten Basen $\mathscr{A}$, $\mathscr{B}$ und $\mathscr{C}$. Seien $\phi : U \to V$ und $\psi : V \to W$ lineare Abbildungen. Dann gilt

$$\mathscr{M}^{\mathscr{A}}_{\mathscr{C}}(\psi \circ \phi) \quad = \quad \mathscr{M}^{\mathscr{B}}_{\mathscr{C}}(\psi)\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi).$$

Das Matrixprodukt entspricht also der Komposition linearer Abbildungen.

***Beweis:*** Sei $n = \dim U$ und $\mathscr{A} = (v_1, ..., v_n)$. Für $1 \le j \le n$ gilt nach (9.9) dann einerseits

$$\mathscr{M}^{\mathscr{A}}_{\mathscr{C}}(\psi \circ \phi)e_j \quad = \quad \mathscr{M}^{\mathscr{A}}_{\mathscr{C}}(\psi \circ \phi)\Phi_{\mathscr{A}}(v_j) \quad = \quad \Phi_{\mathscr{C}}((\psi \circ \phi)(v_j)) \quad ,$$

andererseits aber auch

$$\mathscr{M}^{\mathscr{B}}_{\mathscr{C}}(\psi)\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)e_j \quad = \quad \mathscr{M}^{\mathscr{B}}_{\mathscr{C}}(\psi)\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)\Phi_{\mathscr{A}}(v_j) \quad = \quad \mathscr{M}^{\mathscr{B}}_{\mathscr{C}}(\psi)\Phi_{\mathscr{B}}(\phi(v_j)) \quad =$$

$$\Phi_{\mathscr{C}}(\psi(\phi(v_j))) \quad = \quad \Phi_{\mathscr{C}}((\psi \circ \phi)(v_j)).$$

Also stimmt die $j$-te Spalte von $\mathscr{M}^{\mathscr{A}}_{\mathscr{C}}(\psi \circ \phi)$ mit der $j$-ten Spalte von $\mathscr{M}^{\mathscr{B}}_{\mathscr{C}}(\psi)\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)$ überein, für $1 \le j \le n$. □

**(9.13) Satz** Seien $V, W$ beides $n$-dimensionale $K$-Vektorräume mit geordneten Basen $\mathscr{A}, \mathscr{B}$. Eine lineare Abbildung $\phi : V \to W$ ist genau dann bijektiv, wenn die Darstellungsmatrix $\mathscr{M}_{\mathscr{B}}^{\mathscr{A}}(\phi)$ invertierbar ist, und in diesem Fall gilt

$$\mathscr{M}_{\mathscr{A}}^{\mathscr{B}}(\phi^{-1}) \quad = \quad \mathscr{M}_{\mathscr{B}}^{\mathscr{A}}(\phi)^{-1}.$$

***Beweis:*** Sei $\mathscr{A} = (v_1, ..., v_n)$ und $\mathscr{B} = (w_1, ..., w_n)$. Setzen wir zunächst voraus, dass die Matrix $A = \mathscr{M}_{\mathscr{B}}^{\mathscr{A}}(\phi)$ invertierbar ist. Sei $B = A^{-1}$ die inverse Matrix und $\psi = \mathscr{L}_{\mathscr{A}}^{\mathscr{B}}(B)$. Weil die Abbildungen $\mathscr{L}_{\mathscr{A}}^{\mathscr{B}}$ und $\mathscr{M}_{\mathscr{A}}^{\mathscr{B}}$ nach (9.9) zueinander invers sind, gilt $B = \mathscr{M}_{\mathscr{A}}^{\mathscr{B}}(\psi)$. Für $1 \le j \le n$ gilt dann einerseits

$$\mathscr{M}_{\mathscr{A}}^{\mathscr{A}}(\psi \circ \phi)e_j \quad = \quad \mathscr{M}_{\mathscr{A}}^{\mathscr{A}}(\psi \circ \phi)\Phi_{\mathscr{A}}(v_j) \quad = \quad \mathscr{M}_{\mathscr{A}}^{\mathscr{B}}(\psi)\mathscr{M}_{\mathscr{B}}^{\mathscr{A}}(\phi)\Phi_{\mathscr{A}}(v_j) \quad = \quad BA\Phi_{\mathscr{A}}(v_j) \quad = \quad \Phi_{\mathscr{A}}(v_j)$$

und andererseits

$$\mathscr{M}_{\mathscr{A}}^{\mathscr{A}}(\mathrm{id}_V)e_j \quad = \quad \mathscr{M}_{\mathscr{A}}^{\mathscr{A}}(\mathrm{id}_V)(\Phi_{\mathscr{A}}(v_j)) \quad = \quad \Phi_{\mathscr{A}}(\mathrm{id}_V(v_j)) \quad = \quad \Phi_{\mathscr{A}}(v_j).$$

Die $j$-te Spalte von $\mathscr{M}_{\mathscr{A}}^{\mathscr{A}}(\psi \circ \phi)$ stimmt also mit der $j$-ten Spalte von $\mathscr{M}_{\mathscr{A}}^{\mathscr{A}}(\mathrm{id}_V)$ überein. Also sind die beiden Matrizen gleich. Weil die Zuordnung $\mathscr{M}_{\mathscr{A}}^{\mathscr{A}}$ bijektiv ist, folgt $\psi \circ \phi = \mathrm{id}_V$. Nach dem gleichen Schema zeigt man $\phi \circ \psi = \mathrm{id}_W$. Damit ist die Bijektivität von $\phi$ bewiesen.

Gehen wir nun umgekehrt davon aus, dass $\phi$ bijektiv ist, und sei $\psi$ die Umkehrabbildung. Mit Hilfe von (9.11) erhalten wir

$$\mathscr{M}_{\mathscr{A}}^{\mathscr{B}}(\psi)\mathscr{M}_{\mathscr{B}}^{\mathscr{A}}(\phi) \quad = \quad \mathscr{M}_{\mathscr{A}}^{\mathscr{A}}(\psi \circ \phi) \quad = \quad \mathscr{M}_{\mathscr{A}}^{\mathscr{A}}(\mathrm{id}_V) \quad = \quad E^{(n)}$$

Dies zeigt die Invertierbarkeit von $\mathscr{M}_{\mathscr{B}}^{\mathscr{A}}(\phi)$ und beweist zugleich auch die Identität $\mathscr{M}_{\mathscr{A}}^{\mathscr{B}}(\phi^{-1}) = \mathscr{M}_{\mathscr{B}}^{\mathscr{A}}(\phi)^{-1}$. □

**(9.14) Definition** Seien $V$ ein endlich-dimensionaler $K$-Vektorraum, und seien $\mathscr{A}, \mathscr{B}$ zwei geordnete Basen von $V$. Dann nennt man $\mathscr{T}_{\mathscr{B}}^{\mathscr{A}} = \mathscr{M}_{\mathscr{B}}^{\mathscr{A}}(\mathrm{id}_V)$ die ***Matrix des Basiswechsels*** von $\mathscr{A}$ nach $\mathscr{B}$ oder auch einfach eine ***Transformationsmatrix***.

Die wesentliche Eigenschaft der Transformationsmatrix $\mathscr{T}_{\mathscr{B}}^{\mathscr{A}}$ besteht darin, dass sie die $\mathscr{A}$-Koordinaten eines Vektors in $\mathscr{B}$-Koordinaten umrechnet.

**(9.15) Proposition** Seien Bezeichnungen wie in der Definition und $n = \dim V$.

(i) Für alle $v \in V$ gilt $\mathscr{T}_{\mathscr{B}}^{\mathscr{A}}\Phi_{\mathscr{A}}(v) = \Phi_{\mathscr{B}}(v)$.

(ii) Es gilt $\mathscr{T}_{\mathscr{A}}^{\mathscr{A}} = E^{(n)}$ und $\mathscr{T}_{\mathscr{A}}^{\mathscr{B}} = (\mathscr{T}_{\mathscr{B}}^{\mathscr{A}})^{-1}$.

***Beweis:*** Für jeden Vektor $v \in V$ gilt nach (9.7) jeweils

$$\mathscr{T}_{\mathscr{B}}^{\mathscr{A}}\Phi_{\mathscr{A}}(v) \quad = \quad \mathscr{M}_{\mathscr{B}}^{\mathscr{A}}(\mathrm{id}_V)\Phi_{\mathscr{A}}(v) \quad = \quad \Phi_{\mathscr{B}}(\mathrm{id}_V(v)) \quad = \quad \Phi_{\mathscr{B}}(v).$$

Die Gleichung $\mathscr{T}^{\mathscr{A}}_{\mathscr{A}} = E^{(n)}$ ist eine direkte Folgerung aus der Gleichung $\mathscr{T}^{\mathscr{A}}_{\mathscr{A}} \Phi_{\mathscr{A}}(v) = \Phi_{\mathscr{A}}(v)$, denn mit $v \in V$ durchläuft $\Phi_{\mathscr{A}}(v)$ alle Vektoren aus $K^n$. Schließlich liefert (9.13) noch

$$(T^{\mathscr{A}}_{\mathscr{B}})^{-1} \quad = \quad (M^{\mathscr{A}}_{\mathscr{B}}(\mathrm{id}_V))^{-1} \quad = \quad M^{\mathscr{B}}_{\mathscr{A}}(\mathrm{id}_V^{-1}) \quad = \quad M^{\mathscr{B}}_{\mathscr{A}}(\mathrm{id}_V) \quad = \quad T^{\mathscr{B}}_{\mathscr{A}}. \qquad \square$$

Wir geben ein konkretes für die Bestimmung einer Transformationsmatrix und deren Anwendung an. Sei $K = \mathbb{R}$ und $V$ der $\mathbb{R}$-Vektorraum der Polynomfunktionen vom Grad $\leq 2$. Wir betrachten die geordneten Basen $\mathscr{A} = (f, g, h)$ und $\mathscr{B} = (u, v, w)$ bestehend aus den Elementen $f = 1$, $g = x$, $h = x^2$ sowie $u = 1$, $v = x + 1$, $w = x^2 + x$. Um die Transformationsmatrix $\mathscr{T}^{\mathscr{B}}_{\mathscr{A}}$ zu bestimmen, stellen wir die Elemente von $\mathscr{B}$ als Linearkombinationen der Elemente von $\mathscr{A}$ dar. Es gilt

$$\begin{aligned} u &= 1 &&= 1\cdot 1 + 0 \cdot x + 0 \cdot x^2 \\ v &= x+1 &&= 1\cdot 1 + 1 \cdot x + 0 \cdot x^2 \\ w &= x^2 + x &&= 0\cdot 1 + 1 \cdot x + 1 \cdot x^2. \end{aligned}$$

Jede Rechnung liefert eine Spalte von $\mathscr{T}^{\mathscr{B}}_{\mathscr{A}}$; insgesamt erhalten wir

$$\mathscr{T}^{\mathscr{B}}_{\mathscr{A}} \quad = \quad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Der Algorithmus aus § 3 zur Invertierung von Matrizen liefert

$$\mathscr{T}^{\mathscr{A}}_{\mathscr{B}} \quad = \quad (\mathscr{T}^{\mathscr{B}}_{\mathscr{A}})^{-1} \quad = \quad \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Wir testen nun an einem Beispiel, dass mit $\mathscr{T}^{\mathscr{A}}_{\mathscr{B}}$ tatsächlich Koordinaten umgerechnet werden können, wie in Satz (9.7) angegeben. Das Element $r = x^2 - 2x + 1 \in V$ hat wegen $r = 1 \cdot 1 + (-2) \cdot x + 1 \cdot x^2$ die $\mathscr{A}$-Koordianten $\Phi_{\mathscr{A}}(r) = (1, -2, 1)$. Mit Satz (9.7) erhalten wir

$$\Phi_{\mathscr{B}}(r) \quad = \quad \mathscr{T}^{\mathscr{A}}_{\mathscr{B}} \Phi_{\mathscr{A}}(r) \quad = \quad \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \quad = \quad \begin{pmatrix} 4 \\ -3 \\ 1 \end{pmatrix}.$$

Dies sind tatsächlich die $\mathscr{B}$-Koordinaten von $r$, denn es gilt $4 \cdot 1 + (-3) \cdot (x+1) + 1 \cdot x^2 = x^2 - 2x + 1 = r$.

Zum Schluss sehen wir uns noch an, wie man Darstellungsmatrizen bezüglich unterschiedlicher Basen ineinander umrechnet.

**(9.16) Satz** *(Transformationsformel / Satz vom Basiswechsel)*

Seien $V, W$ endlich-dimensionale $K$-Vektorräume, $\mathscr{A}, \mathscr{A}'$ zwei Basen von $V$ und $\mathscr{B}, \mathscr{B}'$ zwei Basen von $W$. Für jede lineare Abbildung $\phi : V \to W$ gilt dann

$$\mathscr{M}^{\mathscr{A}'}_{\mathscr{B}'}(\phi) \quad = \quad \mathscr{T}^{\mathscr{B}}_{\mathscr{B}'} \, \mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi) \, \mathscr{T}^{\mathscr{A}'}_{\mathscr{A}}.$$

***Beweis:*** Die Gleichung erhält man direkt durch Anwendung von (9.12). Es gilt

$$\mathscr{T}^{\mathscr{B}}_{\mathscr{B}'}\,\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)\,\mathscr{T}^{\mathscr{A}'}_{\mathscr{A}} \quad = \quad \mathscr{M}^{\mathscr{B}}_{\mathscr{B}'}(\mathrm{id}_W)\,\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)\,\mathscr{M}^{\mathscr{A}'}_{\mathscr{A}}(\mathrm{id}_V) \quad = \quad \mathscr{M}^{\mathscr{A}}_{\mathscr{B}'}(\mathrm{id}_W \circ \phi)\mathscr{M}^{\mathscr{A}'}_{\mathscr{A}}(\mathrm{id}_V) \quad =$$

$$\mathscr{M}^{\mathscr{A}'}_{\mathscr{B}'}(\mathrm{id}_W \circ \phi \circ \mathrm{id}_V) \quad = \quad \mathscr{M}^{\mathscr{A}'}_{\mathscr{B}'}(\phi). \qquad \square$$

# § 10. Determinanten

**Inhaltsübersicht**

Jeder quadratischen Matrix $A$ kann auf natürliche Weise eine Zahl $\det(A)$, die sog. *Determinante*, zugeordnet werden, für die es eine ganze Reihe von Anwendungen gibt. Wie wir später sehen werden, lässt sich an der Determinante zum Beispiel erkennen, ob $A$ invertierbar ist. Auch die Inverse $A^{-1}$ und die Lösungen von linearen Gleichungssystemen können mit Hilfe von Determinanten berechnet werden. Im übernächsten Kapitel werden wir die Determinante mit den *Eigenwerten* einer Matrix in Verbindung bringen. Darüber hinaus spielt die Determinante in der Geometrie eine wichtige Rolle, zum Beispiel kann man mit ihr Flächeninhalte und Volumina berechnen.

In diesem Abschnitt geht es hauptsächlich um die *Definition* der Determinante. Zunächst geben wir eine Charakterisierung der Determinantenfunktion durch drei Eigenschaften an: Sie ist multilinear, alternierend und normiert. Wie wir sehen werden, ist die Funktion durch diese drei Eigenschaften tatsächlich eindeutig festgelegt. Um eine explizite Formel für die Determinante anzugeben, benötigen wir die sog. *symmetrischen Gruppen* $S_n$, deren Elemente als *Permutationen* bezeichnet werden. Jede solche Permutation besitzt ein Vorzeichen, das sog. *Signum*, das in der Determinantenformel eine wichtige Rolle spielt. Ein wichtiger Spezialfall dieser Formel ist die eventuell schon aus der Schule bekannte *Sarrus-Regel* zur Berechnung der Determinante einer $3 \times 3$-Matrix.

**Wichtige Definitionen und Sätze**

- Eigenschaften „multilinear“, „alternierend“ und “normiert“ einer Abbildung $d : \mathscr{M}_{n,K} \to K$
- Es gibt genau eine Abbildung $d$ mit diesen drei Eigenschaften. Man bezeichnet sie mit det und nennt $\det(A)$ für jedes $A \in \mathscr{M}_{n,K}$ jeweils die *Determinante* von $A$.
- Definition der symmetrischen Gruppe $S_n$; die Elemente von $S_n$ werden *Permutationen* genannt
- Definition des Signums einer Permuation; Rechenregeln (10.9) und (10.11) zur Bestimmung des Signums
- Leibniz-Formel zur Berechnung der Determinante
- Ist $A$ eine obere Dreiecksmatrix, dann ist $\det(A)$ einfach das Produkt der Diagnoalelemente $a_{kk}$

In diesem Abschnitt erweist es sich an vielen Stellen als praktisch, Matrizen als Tupel bestehend aus ihren Zeilenvektoren darzustellen. Sei $K$ ein Körper und $n \in \mathbb{N}$. Dann schreiben wir $(a_{1\bullet}, ..., a_{n\bullet})$ für die Matrix $A = (a_{ij}) \in \mathscr{M}_{n,K}$ mit den Zeilenvektoren $a_{k\bullet}$ für $1 \le k \le n$. Ist $v_k \in K^n$ ein beliebiger Vektor, dann schreiben wir $(a_{1\bullet}, ..., v_k, ..., a_{n\bullet})$ für die Matrix, die man erhält, wenn man die $k$-te Zeile $a_{k\bullet}$ durch den Vektor $v_k$ ersetzt.

**(10.1) Definition**

(i) Eine Abbildung $d : \mathscr{M}_{n,K} \to K$ bezeichnet man als ***multilinear***, wenn für $1 \le k \le n$ und alle $a_{1\bullet}, ..., a_{n\bullet}, a'_{k\bullet} \in K^n$ und alle $\lambda \in K$ jeweils
$$d(a_{1\bullet}, ..., a_{k\bullet} + a'_{k\bullet}, ..., a_{n\bullet}) \quad = \quad d(a_{1\bullet}, ..., a_{k\bullet}, ..., a_{n\bullet}) + d(a_{1\bullet}, ..., a'_{k\bullet}, ..., a_{n\bullet})$$
und $d(a_{1\bullet}, ..., \lambda a_{k\bullet}, ..., a_{n\bullet}) = \lambda d(a_{1\bullet}, ..., a_{k\bullet}, ..., a_{n\bullet})$ erfüllt ist.

(ii) Man bezeichnet eine multilineare Abbildung $d : \mathscr{M}_{n,K} \to K$ als ***alternierend***, wenn $d(A) = 0$ gilt, sobald zwei Zeilen von $A \in \mathscr{M}_{n,K}$ übereinstimmen.

(iii) Sie ist ***normiert***, d.h. für die Einheitsmatrix $I^{(n)}$ gilt $d(E^{(n)}) = 1$.

Eine multilineare, alternierende und normierte Abbildung $d$ bezeichnet man als ***Determinantenfunktion***.

Wir zeigen anhand einiger konkreter Beispiele, wie die Eigenschaften einer Determinantenfunktion $d : \mathscr{M}_{3,\mathbb{R}} \to \mathbb{R}$ zu interpretieren sind. Die erste Gleichung unter (i), angewendet auf die zweite Zeile der Matrizen, liefert beispielsweise

$$d\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} + d\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 7 & 8 & 9 \end{pmatrix} \quad = \quad d\begin{pmatrix} 1 & 2 & 3 \\ 4 & 6 & 7 \\ 7 & 8 & 9 \end{pmatrix}.$$

Die zweite Gleichung unter (i) bedeutet unter anderem, das man aus jeder einzelnen Zeile eine Faktor 2 „herausziehen“ kann, zum Beispiel in der Form

$$d\begin{pmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \\ 14 & 16 & 18 \end{pmatrix} \quad = \quad 2 \cdot d\begin{pmatrix} 1 & 2 & 3 \\ 8 & 10 & 12 \\ 14 & 16 & 18 \end{pmatrix} \quad = \quad 2 \cdot d\begin{pmatrix} 2 & 4 & 6 \\ 4 & 5 & 6 \\ 14 & 16 & 18 \end{pmatrix} \quad = \quad 2 \cdot d\begin{pmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \\ 7 & 8 & 9 \end{pmatrix}.$$

Diese Regel lässt sich natürlich auch mehrmals hintereinander anwenden. Man erhält so zum Beispiel

$$d\begin{pmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \\ 14 & 16 & 18 \end{pmatrix} \quad = \quad 2 \cdot d\begin{pmatrix} 1 & 2 & 3 \\ 8 & 10 & 12 \\ 14 & 16 & 18 \end{pmatrix} \quad = \quad 4 \cdot d\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 14 & 16 & 18 \end{pmatrix} \quad = \quad 8 \cdot d\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

Allgemein gilt für eine beliebige quadratische Matrix $A \in \mathscr{M}_{n,K}$ und ein beliebiges $\lambda \in K$ jeweils die Gleichung $d(\lambda A) = \lambda^n d(A)$. Konkrete Beispiele für die Anwendung der Regeln (ii) und (iii) sind

$$d\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad = \quad 0 \qquad \text{und} \qquad d\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad = \quad 1.$$

Unser Hauptziel in diesem Abschnitt ist der Nachweis, dass für jeden Körper $K$ und jedes $n \in \mathbb{N}$ genau eine Determinantenfunktion auf $\mathscr{M}_{n,K}$ existiert. Für den Nachweis der Existenz benötigen wir als algebraisches Hilfsmittel die symmetrischen Gruppen.

**(10.2) Proposition** Für jedes $n \in \mathbb{N}$ sei $M_n = \{1, ..., n\}$ die Menge der Zahlen von 1 bis $n$. Dann bilden die bijektiven Abbildungen $\sigma : M_n \to M_n$ mit der Komposition von Abbildungen als Verknüpfung eine Gruppe. Wir nennen sie die ***symmetrische Gruppe*** in $n$ Elementen und bezeichnen sie mit $S_n$.

***Beweis:*** Zunächst beweisen wir das Assoziativgesetz. Für alle $\rho, \sigma, \tau \in S_n$ und alle $x \in M_n$ gilt

$$((\rho \circ \sigma) \circ \tau)(x) \quad = \quad (\rho \circ \sigma)(\tau(x)) \quad = \quad \rho(\sigma(\tau(x))) \quad = \quad \rho((\sigma \circ \tau)(x)) \quad = \quad (\rho \circ (\sigma \circ \tau))(x)$$

und somit $(\rho \circ \sigma) \circ \tau = \rho \circ (\sigma \circ \tau)$. Damit ist die Assoziativität nachwiesen. Die identische Abbildung $\mathrm{id} \in S_n$ gegeben durch $\mathrm{id}(x) = x$ für alle $x \in M_n$ ist in $(S_n, \circ)$ das Neutralelement, denn für alle $\sigma \in S_n$ und alle $x \in M_n$ gilt

$$(\sigma \circ \mathrm{id})(x) \quad = \quad \sigma(\mathrm{id}(x)) \quad = \quad \sigma(x) \quad = \quad \mathrm{id}(\sigma(x)) \quad = \quad (\mathrm{id} \circ \sigma)(x)$$

also $\sigma \circ \mathrm{id} = \mathrm{id} \circ \sigma$. Weil jedes $\sigma \in S_n$ bijektiv ist, existiert jeweils die Umkehrabbildung $\sigma^{-1}$. Diese ist ebenfalls bijektiv, also ein Element in $S_n$. Für alle $x \in M_n$ gilt nach Definition der Umkehrabbildung

$$(\sigma^{-1} \circ \sigma)(x) \quad = \quad \sigma^{-1}(\sigma(x)) \quad = \quad x \quad = \quad \mathrm{id}(x)$$

und somit $\sigma^{-1} \circ \sigma = \mathrm{id}$. Ebenso zeigt man $\sigma \circ \sigma^{-1} = \mathrm{id}$. Damit ist nachgewiesen, dass $\sigma^{-1}$ in $(S_n, \circ)$ das zu $\sigma$ inverse Element ist. Jedes Element in $S_n$ hat also ein Inverses, damit ist $(S_n, \circ)$ eine Gruppe. $\square$

Elemente in $S_n$ können durch Wertetabellen dargestellt werden. Beispielsweise schreibt man

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

für das Element $\sigma \in S_4$, das durch $\sigma(1) = 1$, $\sigma(2) = 4$, $\sigma(3) = 2$ und $\sigma(4) = 3$ gegeben ist. Aus der Analysis einer Variablen ist bekannt, dass für jedes $n \in \mathbb{N}$ und beliebige Mengen $A, B$ mit $|A| = |B| = n$ jeweils genau $n!$ bijektive Abbildungen $A \to B$ existieren. Also gilt auch $|S_n| = n!$ für alle $n \in \mathbb{N}$.

Die Verknüpfung $\sigma \circ \tau$ zweier Elemente $\sigma, \tau \in S_n$ kommt dadurch zu Stande, dass auf jedes $k \in M_n$ erst die Abbildung $\tau$ und dann die Abbildung $\sigma$ angewendet wird. Ist beispielsweise

$$\sigma \quad = \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix} \qquad \text{und} \qquad \tau \quad = \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} \quad ,$$

dann gilt $(\sigma \circ \tau)(1) = \sigma(\tau(1)) = \sigma(2) = 4$ und $(\sigma \circ \tau)(2) = \sigma(\tau(2)) = \sigma(3) = 2$. Ebenso erhält man $(\sigma \circ \tau)(3) = 3$ und $(\sigma \circ \tau)(4) = 1$. Insgesamt gilt also

$$\sigma \circ \tau \quad = \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}$$

Wir werden nun sehen, wie die symmetrische Gruppe $S_n$ mit Determinantenfunktionen auf $\mathcal{M}_{n,K}$ zusammenhängt. Dazu bezeichnen wir mit $\mathrm{Abb}(M_n)$ die Menge aller (nicht notwendig bijektiven) Abbildungen $M_n \to M_n$. Aus dem ersten Semester wissen wir, dass für eine Abbildung $\sigma \in \mathrm{Abb}(M_n)$ die Eigenschaften injektiv, surjektiv und bijektiv äquivalent sind. Ein Element aus $\mathrm{Abb}(M_n)$, das nicht in $S_n$ liegt, ist also weder injektiv noch surjektiv.

Sei nun $d : \mathcal{M}_{n,K} \to K$ eine Determinantenfunktion und $A = (a_{ij}) \in \mathcal{M}_{n,K}$. Mit $e_1, ..., e_n$ bezeichnen wir wie immer die Einheitsvektoren in $K^n$. Die Darstellung der Zeilenvektoren als Linearkombination der Einheitsvektoren liefert $a_{k\bullet} = \sum_{i=1}^n a_{ki} e_i$ für $1 \le k \le n$. Auf Grund der Multilinearität von $d$ gilt

$$d(A) \quad = \quad d(a_{1\bullet}, ..., a_{n\bullet}) \quad = \quad d\left(\sum_{i_1=1}^n a_{1i_1} e_{i_1}, a_{2\bullet}, ..., a_{n\bullet}\right) \quad =$$

$$\sum_{i_1=1}^n a_{1i_1} d(e_{i_1}, a_{2\bullet}, ..., a_{n\bullet}) \quad = \quad \sum_{i_1=1}^n \sum_{i_2=1}^n a_{1i_1} a_{2i_2} d(e_{i_1}, e_{i_2}, a_{3\bullet}, ..., a_{n\bullet}) \quad = \quad ... \quad =$$

$$\sum_{i_1=1}^n ... \sum_{i_n=1}^n a_{1i_1} ... a_{ni_n} d(e_{i_1}, ..., e_{i_n}) \quad = \quad \sum_{\sigma \in \mathrm{Abb}(M_n)} a_{1\sigma(1)} ... a_{n\sigma(n)} d(e_{\sigma(1)}, ..., e_{\sigma(n)})$$

Jedes Element $\sigma \in \mathrm{Abb}(M_n)$ mit $\sigma \notin S_n$ ist auf Grund unserer Vorbemerkung insbesondere nicht injektiv, es gibt also $i, j \in M_n$ mit $i \neq n$ und $\sigma(i) = \sigma(j)$. Weil die Funktion $d$ alternierend ist, gilt dann $d(e_{\sigma(1)}, ..., e_{\sigma(n)}) = 0$. Somit verschwinden in der Summe sämtliche Summanden, die zu Abbildungen $\sigma \in \mathrm{Abb}(M_n) \setminus S_n$ gehören. Wir erhalten

$$d(A) \quad = \quad \sum_{\sigma \in S_n} a_{1\sigma(1)} ... a_{n\sigma(n)} d(e_{\sigma(1)}, ..., e_{\sigma(n)}).$$

Eine Matrix der Form $P_\sigma = (e_{\sigma(1)}, ..., e_{\sigma(n)})$ mit $\sigma \in S_n$ bezeichnet man als ***Permutationsmatrix***. Insbesondere ist $P_{\mathrm{id}} = E^{(n)}$ die Einheitsmatrix. Unserer Rechnung hat ergeben

**(10.3) Proposition** Sei $d : \mathscr{M}_{n,K} \to K$ eine Determinantenfunktion und $A = (a_{ij}) \in \mathscr{M}_{n,K}$. Dann gilt

$$d(A) \quad = \quad \sum_{\sigma \in S_n} a_{1\sigma(1)} ... a_{n\sigma(n)} d(P_\sigma).$$

Nun zeigen wir noch, dass auch die Werte $d(P_\sigma)$ durch die Eigenschaften der Determinantenfunktion eindeutig festgelegt sind. Sei $n \in \mathbb{N}$, und seien $k, \ell \in M_n$ zwei verschiedene Zahlen. Dann ist die Abbildung $\sigma : M_n \to M_n$ gegeben durch

$$\sigma(x) \quad = \quad \begin{cases} \ell & \text{falls } x = k \\ k & \text{falls } x = \ell \\ x & \text{sonst} \end{cases}$$

bijektiv, liegt also in $S_n$. Man verwendet für dieses spezielle Element von $S_n$ die Notation $(k \;\; \ell)$. Allgemein werden Elemente in $S_n$ in dieser Form als ***Transpositionen*** bezeichnet. Jede Transposition $\tau$ hat die Eigenschaft $\tau \circ \tau = \mathrm{id}$, denn die Vertauschung von je zwei Elementen wird durch Wiederholungs des Vorgangs wieder rückgängig gemacht. Es gilt also

$$\tau \quad = \quad \tau^{-1} \qquad \text{für jede Transposition } \tau \in S_n.$$

Wir bestimmen nun das Bild $d(P_\sigma)$ unter einer Determinantenfunktion $d$ zunächst für den Fall, dass $\sigma$ eine Transposition ist. Allgemein gilt

**(10.4) Lemma** Sei $d : \mathscr{M}_{n,K} \to K$ eine Determinantenfunktion, und seien $A, B \in \mathscr{M}_{n,K}$. Entsteht $B$ aus $A$ durch Vertauschung zweier Zeilen, dann gilt $d(B) = -d(A)$.

***Beweis:*** Seien $k, \ell \in M_n$ die beiden Zeilenindizes mit der Eigenschaft, dass $B$ aus $A$ durch Vertauschung der $k$-ten mit der $\ell$-ten Zeile entsteht, wobei $k < \ell$ ist. Weil die Determinantenfunktion multilinear und alternierend ist, gilt

$$d(A) + d(B) \quad = \quad d(..., a_{k\bullet}, ..., a_{\ell\bullet}, ...) + d(..., a_{\ell\bullet}, ..., a_{k\bullet}, ...) \quad =$$

$$d(..., a_{k\bullet}, ..., a_{k\bullet}, ...) + d(..., a_{k\bullet}, ..., a_{\ell\bullet}, ...) + d(..., a_{\ell\bullet}, ..., a_{k\bullet}, ...) + d(..., a_{\ell\bullet}, ..., a_{\ell\bullet}, ...) \quad =$$

$$d(..., a_{k\bullet}, ..., a_{k\bullet} + a_{\ell\bullet}, ...) + d(..., a_{\ell\bullet}, ..., a_{k\bullet} + a_{\ell\bullet}, ...) \quad = \quad d(..., a_{k\bullet} + a_{\ell\bullet}, ..., a_{k\bullet} + a_{\ell\bullet}, ...) \quad = \quad 0.$$

Daraus folgt $d(B) = -d(A)$ □

Die Permutationsmatrix $P_{(k\ \ell)}$ entsteht aus der Einheitsmatrix $E^{(n)}$ durch Vertauschung der $k$-ten und $\ell$-ten Zeile. Nach Eigenschaft (iii) der Determinatenfunktionen gilt also $d(P_{(k\ \ell)}) = -d(E^{(n)}) = -1$. Als nächstes bestimmen wir $d(P_\sigma)$ für beliebige Elemente $\sigma \in S_n$. Dazu bemerken wir zunächst, dass jede Permutation aus Transpositionen zusammengesetzt werden kann.

**(10.5) Proposition** Jedes Element aus $S_n$ ist darstellbar als Produkt von Transpositionen.

***Beweis:*** Sei $\sigma \in S_n$ vorgegeben. Wir beweisen durch vollständige Induktion über $k \in \{0, ..., n\}$: Es gibt ein Produkt $\tau$ von Transpositionen, so dass $(\tau \circ \sigma)(i) = i$ für $1 \le i \le k$ erfüllt ist. Die Aussage für $k = n$ liefert dann $\tau \circ \sigma = \mathrm{id} \Leftrightarrow \sigma = \tau^{-1}$. Mit $\tau$ ist dann auch das Element $\sigma = \tau^{-1}$ ein Produkt von Transpositionen. Ist nämlich $\tau = \tau_1 \circ ... \circ \tau_\ell$ mit Transpositionen $\tau_i$, dann erhalten wir für $\tau^{-1}$ die Produktdarstellung

$$\tau^{-1} \quad = \quad \tau_\ell^{-1} \circ ... \circ \tau_1^{-1} \quad = \quad \tau_\ell \circ ... \circ \tau_1.$$

Kommen wir nun zum Induktionsbeweis. Für $k = 0$ ist nichts zu zeigen. Sei nun $k \in \{0, ..., n-1\}$, und setzen wir die Aussage für $k$ voraus. Dann gibt es ein Produkt $\tau$ von Transpositionen, so dass die Permutation $\tilde{\sigma} = \tau \circ \sigma$ für $1 \le i \le k$ die Gleichung $\tilde{\sigma}(i) = i$ erfüllt. Gilt nun $\tilde{\sigma}(k+1) = k+1$, dann erfüllt $\tau$ die gewünschte Aussage auch für $k+1$. Ansonsten setzen wir $\ell = \tilde{\sigma}(k+1)$; auf Grund der Injektivität von $\tilde{\sigma}$ und wegen $\tilde{\sigma}(i) = i$ für $1 \le i \le k$ und $\tilde{\sigma}(k+1) \ne k+1$ ist $\ell > k+1$. Für das Produkt $\tau' = (k+1\ \ \ell) \circ \tau$ von Transpositionen gilt dann

$$(\tau' \circ \sigma)(k+1) \quad = \quad ((k+1\ \ \ell) \circ \tau \circ \sigma)(k+1) \quad = \quad ((k+1\ \ \ell) \circ \tilde{\sigma})(k+1) \quad = \quad (k+1\ \ \ell)(\ell) \quad = \quad k+1 \quad ,$$

wodurch der Induktionsschritt abgeschlossen ist. □

Es ist nicht schwierig, für ein gegebenes Element $\sigma \in S_n$ eine Darstellungs als Produkt von Transpostionen explizit zu bestimmen. Die Vorgehensweise ist folgende: Zunächst sucht man ein $k \in M_n$ mit $\ell = \sigma(k) \ne k$ und ersetzt den Eintrag $\ell$ an der $k$-ten Stelle durch $k$. Dann sucht man die Stelle $m$ mit $\sigma(m) = k$ und ersetzt den Eintrag $k$ dort durch $\ell$. Bezeichnet man die so modifizierte Permutation, dann gilt $\sigma = (k\ \ \ell) \circ \sigma'$. Denn für jedes $x \notin \{k, \ell, m\}$ gilt offenbar $\sigma(x) = \sigma'(x)$, und somit stimmen $\sigma$ und $(k\ \ \ell) \circ \sigma'$ in $x$ ebenfalls überein. Für $x \in \{k, \ell, m\}$ überprüft man die Gleichung unmittelbar durch Einsetzen: Es gilt

$$((k\ \ \ell) \circ \sigma')(k) = (k\ \ \ell)(k) = \ell = \sigma(k) \qquad \text{und} \qquad ((k\ \ \ell) \circ \sigma')(m) = (k\ \ \ell)(\ell) = k = \sigma(m).$$

Ist $\ell \ne m$, dann folgt $\sigma(\ell) = \sigma'(\ell)$ nach Definition von $\sigma'$, außerdem $\sigma(\ell) \notin \{k, \ell\}$ und somit $((k\ \ \ell) \circ \sigma')(\ell) = (k\ \ \ell)(\sigma(\ell)) = \sigma(\ell)$. Im Fall $\ell = m$ ist die Gleichung $((k\ \ \ell) \circ \sigma')(\ell) = \sigma(\ell)$ bereits überprüft. Die Permutation $\sigma'$ besitzt nun wegen $\sigma'(k) = k$ ein Element mehr als $\sigma$, das auf sich selbst abgebildet wird. Die Berechnung wird nun mit $\sigma'$ an Stelle von $\sigma$ fortgesetzt. Nach endlich vielen Schritten wird $\sigma$ auf diese Weise in die identische Abbildung id umgewandelt, und man erhält eine Darstellung von $\sigma$ als Produkt von Transpositionen.

Wir demonstrieren die Vorgehensweise an einem konkreten Beispiel und betrachten das Element $\sigma \in S_7$ gegeben durch

$$\sigma \quad = \quad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 4 & 7 & 1 & 2 & 3 & 6 \end{pmatrix}.$$

Es gilt dann

$$\begin{aligned}
\sigma \quad &= \quad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 4 & 7 & 1 & 2 & 3 & 6 \end{pmatrix} \quad = \quad (1\ \ 5) \circ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 4 & 7 & 5 & 2 & 3 & 6 \end{pmatrix} \\
&= \quad (1\ \ 5) \circ (2\ \ 4) \circ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 2 & 7 & 5 & 4 & 3 & 6 \end{pmatrix} \quad = \quad (1\ \ 5) \circ (2\ \ 4) \circ (3\ \ 7) \circ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 2 & 3 & 5 & 4 & 7 & 6 \end{pmatrix} \quad = \\
&(1\ \ 5) \circ (2\ \ 4) \circ (3\ \ 7) \circ (4\ \ 5) \circ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 2 & 3 & 4 & 5 & 7 & 6 \end{pmatrix} \quad = \\
&(1\ \ 5) \circ (2\ \ 4) \circ (3\ \ 7) \circ (4\ \ 5) \circ (6\ \ 7) \circ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{pmatrix} \quad = \\
&(1\ \ 5) \circ (2\ \ 4) \circ (3\ \ 7) \circ (4\ \ 5) \circ (6\ \ 7) \circ \mathrm{id} \quad = \quad (1\ \ 5) \circ (2\ \ 4) \circ (3\ \ 7) \circ (4\ \ 5) \circ (6\ \ 7).
\end{aligned}$$

**(10.6) Definition** Sei $\sigma \in S_n$ ein beliebiges Element. Eine zweielementige Teilmenge $\{i,j\}$ von $M_n$ wird ***Fehlstand*** von $\sigma$ genannt, wenn $i < j$, aber $\sigma(i) > \sigma(j)$ gilt. Ist $k \in \mathbb{N}_0$ die Anzahl der Fehlstände von $\sigma$, dann nennt man $\operatorname{sgn}(\sigma) = (-1)^k$ das ***Signum*** oder ***Vorzeichen*** der Permutation.

Wir bestimmen das Signum des Elements $\sigma \in S_4$ gegeben durch

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

durch Abzählen der Fehlstände. Die Menge $M_4 = \{1,2,3,4\}$ besitzt genau sechs zweielementige Teilmengen, nämlich

$$\{1,2\} \quad , \quad \{1,3\} \quad , \quad \{1,4\} \quad , \quad \{2,3\} \quad , \quad \{2,4\} \quad \text{und} \quad \{3,4\}.$$

Die Menge $\{1,2\}$ ist ein Fehlstand von $\sigma$, denn es gilt $1 < 2$, aber $\sigma(1) = 3 > 2 = \sigma(2)$. Dagegen ist $\{1,3\}$ kein Fehlstand, denn es ist $1 < 3$ und $\sigma(1) = 3 < 4 = \sigma(3)$. Geht man alle zweielementigen Teilmengen auf diese Weise der Reihe nach durch, so kommt man zu dem Ergebnis, dass insgesamt vier Fehlstände existieren, nämlich $\{1,4\}$, $\{1,4\}$, $\{2,4\}$ und $\{3,4\}$. Somit gilt $\operatorname{sgn}(\sigma) = (-1)^4 = 1$.

Das Beispiel zeigt, dass die Berechnung des Signums direkt anhand der Definition für großes $n$ sehr mühsam wird. Wir leiten deshalb einige Rechenregeln her, mit denen sich das Signum schneller bestimmen lässt.

**(10.7) Lemma** Für jede Transposition $\tau$ gibt es ein Element $\sigma \in S_n$ mit $\tau = \sigma \circ (1\ \ 2) \circ \sigma^{-1}$.

***Beweis:*** Sei $\tau = (k\ \ell)$ mit $k,\ell \in \{1,...,n\}$ und $\sigma \in S_n$ ein beliebiges Element mit $\sigma \in S_n$ mit $\sigma(1) = k$ und $\sigma(2) = \ell$. Setzen wir $\tau' = \sigma \circ (1\ \ 2) \circ \sigma^{-1}$, dann ist $\tau = \tau'$ zu überprüfen. Zunächst gilt

$$\tau'(k) \quad = \quad (\sigma \circ (1\ \ 2) \circ \sigma^{-1})(k) \quad = \quad (\sigma \circ (1\ \ 2))(1) \quad = \quad \sigma(2) \quad = \quad \ell$$

und $\tau'(\ell) = (\sigma \circ (1\ \ 2) \circ \sigma^{-1})(\ell) = (\sigma \circ (1\ \ 2))(2) = \sigma(1) = k$. Ist $i \notin \{k,\ell\}$, dann ist $\sigma^{-1}(i) \notin \{1,2\}$, denn die Elemente 1 und 2 werden in die Menge $\{k,\ell\}$ abgebildet. Wir erhalten $(1\ \ 2)(\sigma^{-1}(i)) = \sigma^{-1}(i)$ und somit

$$\tau'(i) \quad = \quad (\sigma \circ (1\ \ 2) \circ \sigma^{-1})(i) \quad = \quad (\sigma \circ (1\ \ 2))(\sigma^{-1}(i)) \quad = \quad \sigma(\sigma^{-1}(i)) \quad = \quad i.$$

Insgesamt gilt $\tau(i) = \tau'(i)$ für $1 \le i \le n$, also $\tau = \tau'$. □

**(10.8) Lemma** Für jedes $\sigma \in S_n$ gilt die Produktformel $\operatorname{sgn}(\sigma) = \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i}$.

***Beweis:*** Sei $\sigma \in S_n$ und $m$ die Anzahl der Fehlstände von $\sigma$. Dann gilt

$$\prod_{i<j} (\sigma(j) - \sigma(i)) \quad = \quad \prod_{\substack{i<j \\ \sigma(j)>\sigma(i)}} (\sigma(j) - \sigma(i)) \cdot \prod_{\substack{i<j \\ \sigma(j)<\sigma(i)}} (\sigma(j) - \sigma(i)) \quad =$$

$$\prod_{\substack{i<j \\ \sigma(j)>\sigma(i)}} (\sigma(j) - \sigma(i)) \cdot (-1)^m \prod_{\substack{i<j \\ \sigma(j)<\sigma(i)}} |\sigma(j) - \sigma(i)| \quad = \quad (-1)^m \prod_{i<j} |\sigma(j) - \sigma(i)|$$

Sei $\mathscr{T}$ die Menge der zweielementigen Teilmengen von $\{1,...,n\}$. Dann entsprechen die Paare $(i,j)$ mit $i < j$ und $i,j \in \{1,...,n\}$ bijektiv den Mengen in $\mathscr{T}$. Weil $\sigma$ bijektiv ist, ist mit $\{i,j\} \in \mathscr{T}$ auch $\{\sigma(i),\sigma(j)\}$ eine zweielementige Menge, also in $\mathscr{T}$ enthalten. Die Zuordnung $\mathscr{T} \to \mathscr{T}$, $\{i,j\} \mapsto \{\sigma(i),\sigma(j)\}$ ist bijektiv, da durch $\{i,j\} \mapsto \{\sigma^{-1}(i),\sigma^{-1}(j)\}$ eine Umkehrabbildung gegeben ist; wir bezeichnen diese ebenfalls mit $\sigma$. Für jedes $S = \{i,j\} \in \mathscr{T}$ sei außerdem $r_S = |i - j| \in \mathbb{N}$. Wir erhalten nun

$$(-1)^m \prod_{i<j} |\sigma(j) - \sigma(i)| \quad = \quad (-1)^m \prod_{\{i,j\} \in \mathscr{T}} |\sigma(j) - \sigma(i)| \quad = \quad (-1)^m \prod_{S \in \mathscr{T}} r_{\sigma(S)} \quad =$$

$$(-1)^m \prod_{S \in \sigma(\mathscr{T})} r_S \quad = \quad (-1)^m \prod_{S \in \mathscr{T}} r_S \quad = \quad (-1)^m \prod_{i<j} |j - i| \quad = \quad (-1)^m \prod_{i<j} (j - i).$$

Insgesamt haben wir also gezeigt, dass

$$\prod_{i<j} (\sigma(j) - \sigma(i)) \quad = \quad (-1)^m \prod_{i<j} (j - i) \quad = \quad \operatorname{sgn}(\sigma) \prod_{i<j} (j - i) \qquad \text{erfüllt ist.} \qquad \square$$

**(10.9) Proposition** Für beliebige Elemente $\sigma,\tau \in S_n$ gilt $\operatorname{sgn}(\tau \circ \sigma) = \operatorname{sgn}(\tau)\operatorname{sgn}(\sigma)$.

***Beweis:*** Auf Grund des vorhergehenden Lemmas ist $\operatorname{sgn}(\tau \circ \sigma)$ gegeben durch

$$\prod_{i<j} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{j - i} \quad = \quad \prod_{i<j} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j) - \sigma(i)} \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Das zweite Produkt stimmt mit $\mathrm{sgn}(\sigma)$ überein. Es genügt also zu zeigen, dass das erste Produkt gleich $\mathrm{sgn}(\tau)$ ist. Es gilt

$$\prod_{i<j} \frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)} \quad = \quad \prod_{\substack{i<j \\ \sigma(i)<\sigma(j)}} \frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)} \cdot \prod_{\substack{i<j \\ \sigma(i)>\sigma(j)}} \frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)} \quad =$$

$$\prod_{\substack{i<j \\ \sigma(i)<\sigma(j)}} \frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)} \cdot \prod_{\substack{i>j \\ \sigma(i)<\sigma(j)}} \frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)} \quad = \quad \prod_{\sigma(i)<\sigma(j)} \frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)} \quad ,$$

wobei wir beim zweiten „=“ lediglich die Rollen von $i$ und $j$ im zweiten Faktor vertauscht haben. Um das Produkt weiter zu vereinfachen, beweisen wir die Gleichung

$$\{(k,\ell) \mid 1 \le k < \ell \le n\} \quad = \quad \{(\sigma(i),\sigma(j)) \mid 1 \le i,j \le n \, , \, \sigma(i) < \sigma(j)\}.$$

Der Beweis der Inklusion „$\supseteq$“ ist offensichtlich, denn das Paar $(k,\ell)$ mit $k = \sigma(i)$, $j = \sigma(\ell)$ erfüllt nach Definition die Bedingungen an die Elemente in der Menge links. Zum Nachweis von „$\subseteq$“ sei ein Paar $(k,\ell)$ in der Menge links vorgegeben. Sei $i = \sigma^{-1}(k)$ und $j = \sigma^{-1}(\ell)$. Dann gilt $\sigma(i) = k < \ell = \sigma(j)$, also ist $(k,\ell) = (\sigma(i),\sigma(j))$ ein Element der Menge rechts. Es folgt nun

$$\prod_{\sigma(i)<\sigma(j)} \frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)} \quad = \quad \prod_{k<\ell} \frac{\tau(\ell)-\tau(k)}{\ell-k} \quad = \quad \mathrm{sgn}(\tau). \qquad \square$$

**(10.10) Folgerung** Für jede Permutation $\sigma \in S_n$ gilt $\mathrm{sgn}(\sigma) = \mathrm{sgn}(\sigma^{-1})$.

***Beweis:*** Da das Neutralelement id der Gruppe keine Fehlstände besitzt, gilt $\mathrm{sgn}(\mathrm{id}) = 1$. Aus

$$\mathrm{sgn}(\sigma)\mathrm{sgn}(\sigma^{-1}) \quad = \quad \mathrm{sgn}(\sigma \circ \sigma^{-1}) \quad = \quad \mathrm{sgn}(\mathrm{id}) = 1$$

folgt dann die behauptete Gleichung. $\square$

**(10.11) Satz**

(i) Ist $\tau$ eine Transposition, dann gilt $\mathrm{sgn}(\tau) = -1$.

(ii) Ist $\sigma$ als Produkt von $r$ Transpositionen darstellbar, dann gilt $\mathrm{sgn}(\sigma) = (-1)^r$.

***Beweis:*** zu (i) Die Transposition $(1\ \ 2)$ hat offenbar $\{1,2\}$ als einzigen Fehlstand, also gilt $\mathrm{sgn}((1\ \ 2)) = -1$. Für eine beliebige Transposition $\tau$ finden wir nach (10.7) immer ein $\sigma \in S_n$ mit $\tau = \sigma \circ (1\ \ 2) \circ \sigma^{-1}$. Es folgt $\mathrm{sgn}(\tau) = \mathrm{sgn}(\sigma)\mathrm{sgn}((1\ \ 2))\mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}(\sigma)^2\mathrm{sgn}((1\ \ 2)) = -1$.

zu (ii) Ist $\sigma = \tau_1 \circ \ldots \circ \tau_r$ bestehend aus Transpositionen $\tau_1, \ldots, \tau_r$, dann folgt aus Teil (i) die Gleichung $\mathrm{sgn}(\sigma) = \prod_{i=1}^r \mathrm{sgn}(\tau_i) = \prod_{i=1}^r (-1) = (-1)^r$. $\square$

Mit Hilfe dieses Satzes haben wir nun eine effiziente Methode für die Berechnung des Signums zur Verfügung. Wir betrachten noch einmal das Beispiel im Anschluss an Definition (10.6) Für das dort betrachtete Element $\sigma \in S_4$ gilt

$$\sigma \quad = \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix} \quad = \quad (1 \ \ 3) \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 3 \end{pmatrix} \quad = \quad (1 \ \ 3) \circ (3 \ \ 4).$$

Also ist $\sigma$ als Produkt von zwei Transpositionen darstellbar. Mit Satz (10.11) (ii) erhalten wir $\mathrm{sgn}(\sigma) = (-1)^2 = 1$.

Nun sind wir auch in der Lage, die Determinante der Permutationsmatrizen direkt anzugeben.

**(10.12) Folgerung** Sei $d : \mathcal{M}_{n,K} \to K$ eine Determinantenfunktion. Dann gilt

$$d(P_\sigma) \quad = \quad \mathrm{sgn}(\sigma) \qquad \text{für alle} \quad \sigma \in S_n.$$

***Beweis:*** Ist $\rho \in S_n$ und $\tau = (k \ \ \ell)$ eine Transposition, dann gilt $d(P_{\rho\circ\tau}) = -d(P_\rho)$ nach (10.4), denn die Matrix $P_{\rho\circ\tau}$ entsteht aus $P_\rho$ durch Vertauschung $k$-te und $\ell$-ten Zeile: Für $i \neq k, \ell$ ist die $i$-te Zeile von $P_{\rho\circ\tau}$ gegeben durch $e_{(\rho\circ\tau)(i)} = e_{\rho(i)}$, und die $k$-te und $\ell$-te Zeile sind $e_{(\rho\circ\tau)(k)} = e_{\rho(\ell)}$ bzw. $e_{(\rho\circ\tau)(\ell)} = e_{\rho(k)}$.

Sei nun $\sigma \in S_n$ beliebig vorgegeben. Nach (10.5) gibt es ein $r \in \mathbb{N}_0$ und Transpositionen $\tau_1, ..., \tau_r \in S_n$, so dass $\sigma = \tau_1 \circ ... \circ \tau_r$ erfüllt ist. Aus (10.8) folgt daraus $\mathrm{sgn}(\sigma) = (-1)^r$. Wir beweisen die behauptete Gleichung nun durch vollständige Induktion über $r$. Im Fall $r = 0$ gilt $\sigma = \mathrm{id}$ und $d(P_\sigma) = d(E^{(n)}) = 1$ auf Grund der Bedingung (iii) für Determinantenfunktionen.

Sei nun $r > 1$, und setzen wir die Gleichung für Werte $< r$ voraus. Definieren wir das Element $\sigma' \in S_n$ durch $\sigma' = \tau_1 \circ ... \circ \tau_{r-1}$, dann gilt $\sigma = \sigma' \circ \tau_r$, außerdem $\mathrm{sgn}(\sigma') = (-1)^{r-1}$ und auf Grund unserer Vorüberlegung

$$d(P_\sigma) \quad = \quad d(P_{\sigma'\circ\tau_r}) \quad = \quad -d(P_{\sigma'}) \quad = \quad -(-1)^{r-1} \quad = \quad (-1)^r \quad = \quad \mathrm{sgn}(\sigma). \qquad \square$$

Setzen wir dieses Ergebnis in (10.3) ein, so erhalten wir

**(10.13) Folgerung** Sei $d : \mathcal{M}_{n,K} \to K$ eine Determinantenfunktion und $A = (a_{ij}) \in \mathcal{M}_{n,K}$. Dann gilt

$$d(A) \quad = \quad \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{1\sigma(1)} \dots a_{n\sigma(n)}.$$

Nach diesen Vorbereitungen können wir nun endlich definieren

> **(10.14) Definition** Sei $A = (a_{ij})_{n\in\mathbb{N}} \in \mathscr{M}_{n,K}$. Dann nennen wir
>
> $$\det(A) \quad = \quad \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} ... a_{n\sigma(n)} \qquad \text{die } \textbf{\textit{Determinante}} \text{ der Matrix } A.$$

Der Summenausdruck in der Definition von $\det(A)$ wird auch die ***Leibniz-Formel*** für die Determinante genannt. Wir geben die Formel für die Werte $n = 1, 2, 3$ noch einmal explizit an. Für $n = 1$ ist $S_n = \{\mathrm{id}\}$. In diesem Fall besteht die Summe also nur aus einem einzigen Term. Es gilt $\det(A) = \operatorname{sgn}(\mathrm{id}) a_{1\mathrm{id}(1)} = a_{11}$, also zum Beispiel

$$\det\begin{pmatrix} 3 \end{pmatrix} \quad = \quad 3.$$

Im Fall $n = 2$ gilt $S_2 = \{\ \mathrm{id}, (1\ 2)\}$, und die Transposition $\tau = (1\ 2)$ hat nach (10.11) ein negatives Signum. Damit erhalten wir $\det(A) = \operatorname{sgn}(\mathrm{id}) a_{1\mathrm{id}(1)} a_{2\mathrm{id}(2)} + \operatorname{sgn}(\tau) a_{1\tau(1)} a_{2\tau(2)} = a_{11}a_{22} - a_{21}a_{12}$. Beispielsweise ist

$$\det\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad = \quad 1\cdot 4 - 2\cdot 3 \quad = \quad 4 - 6 \quad = \quad -2.$$

Für $n = 3$ besteht $S_3$ bereits aus $3! = 6$ Elementen, es gilt $S_3 = \{\mathrm{id}, \sigma_1, \sigma_2, \tau_1, \tau_2, \tau_3\}$ mit den Elementen

$$\sigma_1 = (2\ 3)\circ(1\ 3) \ \ , \ \ \sigma_2 = (1\ 3)\circ(2\ 3) \ \ , \ \ \tau_1 = (1\ 3) \ \ , \ \ \tau_2 = (2\ 3) \quad \text{und} \quad \tau_3 = (1\ 2).$$

Zum Beweis genügt es zu überprüfen, dass diese sechs Elemente tatsächlich *verschiedene* Elemente von $S_3$ sind. Nach (10.11) gilt $\operatorname{sgn}(\mathrm{id}) = \operatorname{sgn}(\sigma_1) = \operatorname{sgn}(\sigma_2) = 1$ und $\operatorname{sgn}(\tau_1) = \operatorname{sgn}(\tau_2) = \operatorname{sgn}(\tau_3) = -1$.

Mit den sechs Elemente der Gruppe $S_3$ können wir nun die Formel für die Determinante im Fall $n = 3$ aufstellen.

$$\begin{aligned} \det(A) \quad = \quad & a_{1\mathrm{id}(1)} a_{2\mathrm{id}(2)} a_{3\mathrm{id}(3)} + a_{1\sigma_1(1)} a_{2\sigma_1(2)} a_{3\sigma_1(3)} + a_{1\sigma_2(1)} a_{2\sigma_2(2)} a_{3\sigma_2(3)} \\ & - a_{1\tau_1(1)} a_{2\tau_1(2)} a_{3\tau_1(3)} - a_{1\tau_2(1)} a_{2\tau_2(2)} a_{3\tau_2(3)} - a_{1\tau_3(1)} a_{2\tau_3(2)} a_{3\tau_3(3)} \quad = \\ & a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}. \end{aligned}$$

Man bezeichnet diese Formel auch als ***Sarrus-Regel*** zur Berechnung der Determinante. Die drei positiven Summanden entsprechen im Zahlenschema

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

den drei nebeneinanderliegenden Diagonalen von links oben nach rechts unten. Die drei negativen Summanden entsprechen den drei Diagonalen, die von links unten nach rechts oben verlaufen. Mit der Sarrus-Regel erhält man zum Beispiel

$$\begin{aligned} \det\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \quad & = \quad 1\cdot5\cdot9 + 2\cdot6\cdot7 + 3\cdot4\cdot8 - 7\cdot5\cdot3 - 8\cdot6\cdot1 - 9\cdot4\cdot2 \\ & = \quad 45 + 84 + 96 - 105 - 48 - 72 \quad = \quad 225 - 255 \quad = \quad 0. \end{aligned}$$

Zu beachten ist, dass ein Analogon der Sarrus-Regel für $n = 4$ ***falsch*** ist. Die Leibniz-Formel für eine Matrix $A \in \mathscr{M}_{4,K}$ besteht aus $4! = 24$ Summanden, wärend in der Sarrus-Regel nur acht Terme vorkommen würden.

Nun beweisen wir noch, dass durch die Leibniz-Formel tatsächlich eine Determinantenfunktion definiert ist. Dazu benötigen wir weitere Grundlagen über die symmetrische Gruppe $S_n$. Für jedes $n \in \mathbb{N}$ bilden die Permutationen mit positivem Signum die sogenannte ***alternierende Gruppe***

$$A_n \quad = \quad \{\sigma \in S_n \mid \operatorname{sgn}(\sigma) = 1\}.$$

Die Gruppe $S_n$ setzt sich zu gleichen Teilen aus Elementen mit positivem und negativem Signum zusammen. Genauer gilt

**(10.15) Proposition** Sei $A_n = \{\sigma \in S_n \mid \operatorname{sgn}(\sigma) = 1\}$ und $\tau \in S_n$ ein beliebiges Element mit $\operatorname{sgn}(\tau) = -1$. Dann ist durch $S_n = A_n \cup (A_n \circ \tau)$ mit $A_n \circ \tau = \{\sigma \circ \tau \mid \sigma \in A_n\}$ eine Darstellung von $S_n$ als disjunkte Vereinigung gegeben. Zwischen $A_n$ und $A_n \circ \tau$ ist durch $\sigma \mapsto \sigma \circ \tau$ eine Bijektion definiert.

***Beweis:*** Zunächst beweisen wir die Gleichung $S_n = A_n \cup A_n \circ \tau$. Die Inklusion „$\supseteq$" ist offensichtlich, denn beide Mengen rechts bestehen aus Elementen von $S_n$. Zum Nachweis von „$\subseteq$" sei $\sigma \in S_n$ vorgegeben. Liegt $\sigma$ in $A_n$, dann ist nichts zu zeigen. Im Fall $\operatorname{sgn}(\sigma) = -1$ liegt $\sigma \circ \tau^{-1}$ in $A_n$, denn es gilt $\operatorname{sgn}(\sigma \circ \tau^{-1}) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)^{-1} = (-1)(-1)^{-1} = 1$. Also ist $\sigma = (\sigma \circ \tau^{-1}) \circ \tau$ in $A_n \circ \tau$ enthalten.

Die Elemente in $A_n$ haben positives Signum. Jedes Element in $A_n \circ \tau$ der Form $\sigma \circ \tau$ mit $\sigma \in A_n$ hat wegen $\operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau) = 1 \cdot (-1) = -1$ negatives Signum. Dies zeigt, dass die Mengen $A_n$ und $A_n \circ \sigma$ disjunkt sind.

Die Abbildung $\sigma \mapsto \sigma \circ \tau$ zwischen $A_n$ und $A_n \circ \sigma$ ist surjektiv, denn jedes Element in $A_n \circ \tau$ kann in der Form $\sigma \circ \tau$ mit $\sigma \in A_n$ geschrieben werden. Sind $\sigma_1, \sigma_2 \in A_n$ mit $\sigma_1 \circ \tau = \sigma_2 \circ \tau$, dann folgt $\sigma_1 \circ \tau \circ \tau^{-1} = \sigma_2 \circ \tau \circ \tau^{-1}$ und damit $\sigma_1 = \sigma_2$. Damit ist auch die Injektivität der Abbildung nachgewiesen. □

**(10.16) Satz** Für jedes $n \in \mathbb{N}$ ist $\det : \mathscr{M}_{n,K} \to K$ eine Determinantenfunktion.

***Beweis:*** Wir müssen überprüfen, dass die Abbildung det die drei Bedingungen aus (10.1) erfüllt. Sei $k \in \{1, ..., n\}$, und seien $A, B \in \mathscr{M}_{n,K}$ zwei Matrizen, die in allen Zeilen mit Ausnahme der $k$-ten übereinstimmen. Sei außerdem $C \in \mathscr{M}_{n,K}$ die Matrix mit $c_{k\bullet} = a_{k\bullet} + b_{k\bullet}$ und $c_{\ell\bullet} = a_{\ell\bullet} = b_{\ell\bullet}$ für $\ell \neq k$. Zu zeigen ist $\det(C) = \det(A) + \det(B)$.

Tatsächlich gilt

$$\det(C) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\prod_{\ell=1}^{n} c_{\ell\sigma(\ell)} = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) c_{k\sigma(k)} \prod_{\ell\neq k} c_{\ell\sigma(\ell)} =$$

$$\sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)(a_{k\sigma(k)} + b_{k\sigma(k)}) \prod_{\ell\neq k} c_{\ell\sigma(\ell)} = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{k\sigma(k)} \prod_{\ell\neq k} c_{\ell\sigma(\ell)} + \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) b_{k\sigma(k)} \prod_{\ell\neq k} c_{\ell\sigma(\ell)} =$$

$$\sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{k\sigma(k)} \prod_{\ell\neq k} a_{\ell\sigma(\ell)} + \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) b_{k\sigma(k)} \prod_{\ell\neq k} b_{\ell\sigma(\ell)} = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{\ell=1}^{n} a_{\ell\sigma(\ell)} + \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{\ell=1}^{n} b_{\ell\sigma(\ell)}$$

$$= \det(A) + \det(B).$$

Sei nun $\lambda \in K$ und $D \in \mathscr{M}_{n,K}$ die Matrix gegeben durch $d_{k\bullet} = \lambda a_{k\bullet}$ und $d_{\ell\bullet} = a_{\ell\bullet}$ für $\ell \neq k$. Dann gilt

$$\det(D) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{\ell=1}^{n} d_{\ell\sigma(\ell)} = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) d_{k\sigma(k)} \prod_{\ell\neq k} d_{\ell\sigma(\ell)} = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)(\lambda a_{k\sigma(k)}) \prod_{\ell\neq k} a_{\ell\sigma(\ell)}$$

$$= \lambda \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{k\sigma(k)} \prod_{\ell\neq k} a_{\ell\sigma(\ell)} = \lambda \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{\ell=1}^{n} a_{\ell\sigma(\ell)} = \lambda \det(A).$$

Damit ist die Eigenschaft (i) aus (10.1) verifiziert. Zum Nachweis von (ii) sei $A \in \mathscr{M}_{n,K}$ eine Matrix, in der die $k$-te und $\ell$-te Zeile übereinstimmen, wobei $k \neq \ell$ sei. Für die Transposition $\tau = (k\ \ell)$ gilt $S_n = A_n \cup A_n \circ \tau$ nach (10.15). Weil die Elemente in $A_n$ positives und die Elemente in $A_n \circ \tau$ negatives Signum haben, erhalten wir für $\det(A)$ den Ausdruck

$$\sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} = \sum_{\sigma\in A_n} \prod_{i=1}^{n} a_{i\sigma(i)} - \sum_{\sigma\in A_n\circ\tau} \prod_{i=1}^{n} a_{i\sigma(i)} = \sum_{\sigma\in A_n} \prod_{i=1}^{n} a_{i\sigma(i)} - \sum_{\sigma\in A_n} \prod_{i=1}^{n} a_{i\sigma(\tau(i))}$$

Weil die $k$-te und die $\ell$-te Zeile von $A$ übereinstimmen, gilt für die einzelnen Terme der rechten Summe

$$\prod_{i=1}^{n} a_{i\sigma(\tau(i))} = a_{k\sigma(\tau(k))} a_{\ell\sigma(\tau(\ell))} \prod_{i\neq k,\ell} a_{i\sigma(\tau(i))} = a_{k\sigma(\ell)} a_{\ell\sigma(k)} \prod_{i\neq k,\ell} a_{i\sigma(\tau(i))}$$

$$= a_{\ell\sigma(\ell)} a_{k\sigma(k)} \prod_{i\neq k,\ell} a_{i\sigma(\tau(i))} = \prod_{i=1}^{n} a_{i\sigma(i)}.$$

Also heben sich die beiden Summen auf, und es folgt $\det(A) = 0$. Nun beweisen wir noch die Eigenschaft (iii). Die Determinante der Einheitsmatrix ist nach Definition gegeben durch

$$\det(E^{(n)}) = \sum_{\sigma\in S_n} s_\sigma \quad , \qquad s_\sigma = \operatorname{sgn}(\sigma)\delta_{1\sigma(1)}\cdots\delta_{n\sigma(n)}.$$

Gilt $\sigma(i) \neq i$ für ein $i$, dann folgt $\delta_{i\sigma(i)} = 0$ und somit $s_\sigma = 0$. Also ist $s_{\mathrm{id}}$ der einzige nicht-verschwindende Summand in der Leibniz-Formel, und dieser ist gleich 1. □

In der folgenden Situation lässt sich die Determinante einer Matrix leicht ausrechnen.

**(10.17) Satz** Man bezeichnet eine Matrix $A = (a_{ij}) \in \mathscr{M}_{n,K}$ als ***obere Dreiecksmatrix***, wenn $a_{ij} = 0$ für alle $i, j \in \{1, ..., n\}$ mit $i > j$ erfüllt ist. Für jede Matrix dieser Form gilt

$$\det(A) \quad = \quad a_{11} \cdot a_{22} \cdot ... \cdot a_{nn}.$$

***Beweis:*** Zunächst zeigen wir, dass für jede Permutation $\sigma \in S_n \setminus \{\mathrm{id}\}$ ein $k \in M_n$ mit $\sigma(k) < k$ existiert. Anderfalls müsste $\sigma(\ell) \geq \ell$ für $1 \leq \ell \leq n$ gelten. Wegen $\sigma \neq \mathrm{id}$ gibt es andererseits auch ein maximales $k \in M_n$ mit $\sigma(k) > k$. Aber auf Grund der Maximalität müsste dann $\sigma(\sigma(k)) = \sigma(k)$ gelten, im Widerspruch zur Injektivität von $\sigma$.

Sei nun $\sigma \in S_n \setminus \{\mathrm{id}\}$ und $k \in M_n$ mit $\sigma(k) < k$. Dann folgt $a_{k\sigma(k)} = 0$ nach Definition der oberen Dreiecksmatrix, und zu $\sigma$ gehörende Summand $\mathrm{sgn}(\sigma) \prod_{\ell=1}^{n} a_{\ell\sigma(\ell)}$ ist ebenfalls gleich null. Der einzige eventuell nicht verschwindende Summand in der Leibniz-Formel ist also der zum Element id Summand; wegen $\mathrm{sgn}(\mathrm{id}) = 1$ ist dieser Summand gleich dem Produkt $\prod_{\ell=1}^{n} a_{\ell\ell}$. □

Es gilt also beispielsweise

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix} \quad = \quad 1 \cdot 4 \cdot 6 \quad = \quad 24.$$

# § 11. Rechenregeln für Determinanten

***Inhaltsübersicht***

In diesem Abschnitt behandeln wir Verfahren zur Berechnung von Determinanten. Ein besonders effizientens Verfahren ist die Überführung in Zeilenstufenform. Als Nebenresultat erhalten wir die wichtige Gleichung $\det(AB) = \det(A)\det(B)$ für Matrixprodukte und das hinreichende und notwendige Kriterium $\det(A) \neq 0$ für die Invertierbarkeit einer Matrix. Auch die Blockgestalt von Matrizen kann zur Berechnung der Determinante genutzt werden. Mit dem Laplaceschen Entwicklungssatz kann die Berechnung der Determinante einer $n \times n$-Matrix auf kleinere Matrizen zurückgeführt werden.

***Wichtige Definitionen und Sätze***

- Das Gaußverfahren kann zur Berechnung von Determinanten eingesetzt werden.
- Eine Matrix $A \in \mathscr{M}_{n,K}$ ist genau dann invertierbar, wenn $\det(A) \neq 0$ ist.
- *Multiplikationssatz:* Es gilt $\det(AB) = \det(A)\det(B)$ für alle $A, B \in \mathscr{M}_{n,K}$.
- Definition der zu $A \in \mathscr{M}_{n,K}$ komplementären Matrix
- Laplacescher Entwicklungssatz

Zunächst untersuchen wir in diesem Abschnitt, wie sich Zeilenumformungen auf die Determinante einer Matrix auswirken. In Kapitel §2 haben wir die Elementarmatrizen

$$\mathrm{M}_{k,\lambda} = E^{(m)} + (\lambda - 1)\mathrm{B}_{kk}^{(m\times m)} \qquad \text{und} \qquad \mathrm{A}_{k,\ell,\lambda} = E^{(m)} + \lambda \mathrm{B}_{\ell k}^{(m\times m)}$$

eingeführt. Die Matrix $\mathrm{M}_{k,\lambda}$ entsteht aus der Einheitsmatrix $E^{(n)}$ durch Multiplikation der $k$-ten Zeilen mit dem Wert $\lambda$. Auf Grund der Multilinearität der Determinantenfunktion gilt deshalb $\det(\mathrm{M}_{k,\lambda}) = \lambda \det E^{(n)} = \lambda \cdot 1_K = \lambda$. Ist $A \in \mathscr{M}_{n,K}$ eine beliebige Matrix, dann entsteht $\mathrm{M}_{k,\lambda}A$ aus $A$ durch Multiplikation der $k$-ten Zeile mit dem Wert $\lambda$. Wir erhalten somit die Rechenregel

$$\det(\mathrm{M}_{k,\lambda}A) \quad = \quad \lambda \det(A) \quad = \quad \det(\mathrm{M}_{k,\lambda})\det(A) \qquad \text{für alle} \quad A \in \mathscr{M}_{n,K}.$$

Allgemein gilt: Entsteht eine Matrix $B = (b_{ij})$ aus $A = (a_{ij}) \in \mathscr{M}_{n,K}$ durch Addition des $\lambda$-fachen der $k$-ten Zeile zur $\ell$-ten, dann gilt $\det(A) = \det(A')$, denn aus den Eigenschaften der Determinantenfunktion folgt

$$\det B \quad = \quad \det(\ldots, b_{k\bullet}, \ldots, b_{\ell\bullet}, \ldots) \quad = \quad \det(\ldots, a_{k\bullet}, \ldots, \lambda a_{k\bullet} + a_{\ell\bullet}, \ldots) \quad =$$

$$\lambda \det(\ldots, a_{k\bullet}, \ldots, a_{k\bullet}, \ldots) + \det(\ldots, a_{k\bullet}, \ldots, a_{\ell\bullet}, \ldots) \quad = \quad \lambda \cdot 0 + \det(A) \quad = \quad \det(A).$$

Insbesondere entsteht die Matrix $\mathrm{A}_{k,\ell,\lambda}$ aus $E^{(n)}$ durch Addition des $\lambda$-fachen der $k$-ten Zeile zur $\ell$-ten. Somit gilt $\det \mathrm{A}_{k,\ell,\lambda} = E^{(n)} = 1$ und allgemein

$$\det(\mathrm{A}_{k,\ell,\lambda}A) \quad = \quad \det(\mathrm{A}_{k,\ell,\lambda})\det(A) \qquad \text{für alle} \quad A \in \mathscr{M}_{n,K}.$$

Zusammenfassend kann also formuliert werden

**(11.1) Lemma** Ist $T \in \mathscr{M}_{n,K}$ eine Elementarmatrix und $A \in \mathscr{M}_{n,K}$ beliebig, dann gilt

$$\det(TA) \quad = \quad \det(T)\det(A).$$

Eine quadratische Matrix in Zeilenstufenform ist insbesondere eine obere Dreiecksmatrix, und deren Determinante erhält man nach (10.17) aus dem Produkt der Diagonalelemente. Wir haben gesehen, wie sich elementare Zeilenumformungen vom Typ $(\mathrm{M}_{k,\lambda})$ und $(\mathrm{A}_{k,\ell,\lambda})$ auf die Determinante auswirken. Außerdem führt die Vertauschung zweier Zeilen nach (10.4) lediglich zu einem Vorzeichenwechsel bei der Determinante. Dies zusammen liefert uns folgende Strategie für die Berechnung von $\det(A)$ für eine beliebige Matrix $A \in \mathscr{M}_{n,K}$.

(i) Forme $A$ mit Hilfe des Gaußverfahrens in eine Matrix $B$ in Zeilenstufenform um.
(ii) Bestimme anhand der durchgeführten Zeilenumformungen den Faktor $\mu \in K^\times$ mit $\det(B) = \mu \det(A)$.
(iii) Berechne $\det(B)$ durch Multiplikation der Diagonalelemente $b_{11}, ..., b_{nn}$ von $B$.
Das Endergebnis der Rechnung ist dann $\det(A) = \mu^{-1} \det(B)$.

Wir demonstrieren die Funktionsweise des Verfahrens anhand der Matrix

$$A \quad = \quad \begin{pmatrix} 1 & 0 & 3 & 7 \\ 0 & 2 & 1 & 4 \\ 1 & 0 & -1 & 1 \\ -2 & 3 & 0 & 2 \end{pmatrix}.$$

Mit dem Gaußverfahren erhalten wir

$$\begin{pmatrix} 1 & 0 & 3 & 7 \\ 0 & 2 & 1 & 4 \\ 1 & 0 & -1 & 1 \\ -2 & 3 & 0 & 2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 3 & 7 \\ 0 & 2 & 1 & 4 \\ 0 & 0 & -4 & -6 \\ 0 & 3 & 6 & 16 \end{pmatrix} \overset{(*)}{\mapsto} \begin{pmatrix} 1 & 0 & 3 & 7 \\ 0 & 3 & 6 & 16 \\ 0 & 0 & -4 & -6 \\ 0 & 2 & 1 & 4 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 3 & 7 \\ 0 & 1 & 5 & 12 \\ 0 & 0 & -4 & -6 \\ 0 & 2 & 1 & 4 \end{pmatrix} \mapsto$$

$$\begin{pmatrix} 1 & 0 & 3 & 7 \\ 0 & 1 & 5 & 12 \\ 0 & 0 & -4 & -6 \\ 0 & 0 & -9 & -20 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 3 & 7 \\ 0 & 1 & 5 & 12 \\ 0 & 0 & -4 & -6 \\ 0 & 0 & -1 & -8 \end{pmatrix} \overset{(*)}{\mapsto} \begin{pmatrix} 1 & 0 & 3 & 7 \\ 0 & 1 & 5 & 12 \\ 0 & 0 & -1 & -8 \\ 0 & 0 & -4 & -6 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 3 & 7 \\ 0 & 1 & 5 & 12 \\ 0 & 0 & -1 & -8 \\ 0 & 0 & 0 & 26 \end{pmatrix}$$

Bezeichnen wir die Matrix am Ende der Rechnung mit $B$, dann gilt $\det(B) = 1 \cdot 1 \cdot (-1) \cdot 26 = -26$. In der Rechnung kommen an den beiden mit (*) markierten Stellen Zeilenvertauschungen vor, die jeweils einen Vorzeichenwechsel bewirken. Insgesamt gilt also $\det(B) = (-1)^2 \det(A) = \det(A)$ und damit $\det(A) = \det(B) = -26$.

Neben diesem praktischen Rechenverfahren führen unsere Überlegungen auch zu neuen theoretischen Resultaten.

**(11.2) Satz** *(Charakterisierung invertierbarer Matrizen)*

Für eine Matrix $A \in \mathscr{M}_{n,K}$ sind folgende Aussagen äquivalent.

(i) $A \in \mathrm{GL}_n(K)$ (d.h. die Matrix $A$ ist invertierbar)

(ii) $\mathrm{rg}(A) = n$

(iii) $\det(A) \neq 0$

***Beweis:*** Zunächst beweisen wir die Äquivalenz „(i) $\Leftrightarrow$ (ii)“. Sei $\phi_A : K^n \to K^n$ gegeben durch $v \mapsto Av$. Nach dem Dimensionssatz für lineare Abbildungen gilt $n = \dim\ker(\phi_A) + \dim\mathrm{im}(\phi_A)$. Aus §8 wissen wir auch, dass $\mathrm{rg}(A) = \dim\mathrm{im}(\phi_A)$ gilt. Also ist $\mathrm{rg}(A) = n$ genau dann erfüllt, wenn

$$\dim\ker(\phi_A) = 0 \quad \text{und} \quad \dim\mathrm{im}(\phi_A) = n$$

erfüllt gilt. Dies wiederum ist genau dann der Fall, wenn $\phi_A$ bijektiv ist. Nach (9.9), angewendet auf die kanonische Basis $\mathscr{E}$ von $K^n$, ist die Bijektivität von $\phi_A$ wiederum äquivalent zur Invertierbarkeit von $A$.

An Stelle der Äquivalenz „(ii) $\Leftrightarrow$ (iii)“ beweisen wir, dass genau dann $\det(A) = 0$ ist, wenn $\mathrm{rg}(A) < n$ gilt. Weil sich die Determinante der Matrix bei Zeilenumformungen höchstens um einen Faktor $\lambda \in K^\times$ ändert, können wir voraussetzen, dass sich die Matrix $A$ in Zeilenstufenform befindet. Der Zeilenrang von $A$ ändert sich durch diese Umformungen nicht. Seien $j_1 < \ldots < j_r$ die Kennzahlen der Zeilenstufenform mit $r = \mathrm{rg}(A)$. Ist $r < n$, dann ist $a_{nn} = 0$, und somit gilt auch $\det(A) = \prod_{i=1}^n a_{ii} = 0$. Ist dagegen $r = n$, dann muss $j_i = i$ für $1 \leq i \leq n$ gelten, und die Einträge $a_{ij_i} = a_{ii}$ sind nach Definition der Zeilenstufenform ungleich Null. Also ist auch $\det(A)$ als Produkt der Diagonaleinträge ungleich Null. □

**(11.3) Satz** *(Multiplikationssatz für Determinanten)*

Seien $A, B \in \mathscr{M}_{n,K}$. Dann gilt $\det(AB) = \det(A)\det(B)$.

*Beweis:* Wie immer bezeichnen wir mit $\phi_A$ bzw. $\phi_B$ die linearen Abbildungen $K^n \to K^n$ gegeben durch $v \mapsto Av$ bzw. $v \mapsto Bv$. Nehmen wir zunächst an, dass $\det(A) = 0$ ist. Dann folgt $\dim\mathrm{im}(\phi_A) = \mathrm{rg}(A) < n$ und somit auch

$$\mathrm{rg}(AB) \quad = \quad \dim\mathrm{im}(\phi_A \circ \phi_B) \quad \leq \quad \dim\phi_A(K^n) \quad < \quad n.$$

Dies wiederum bedeutet $\det(AB) = 0$, d.h. in diesem Fall ist die Gleichung $\det(AB) = \det(A)\det(B)$ erfüllt. Nun betrachten wir den Fall, dass $A$ invertierbar (also $\det(A) \neq 0$) ist. Für den Fall $A = E^{(n)}$ ist die Aussage trivial, und im Fall, dass es sich bei $A$ um eine Elementarmatrix handelt, ist sie nach (11.1) erfüllt. Im allgemeinen Fall ist aus

§3 bekannt, dass $A$ als Produkt $T_1 \cdot \ldots \cdot T_r$ von Elementarmatrizen darstellbar ist. Wir beweisen nun die Gleichung $\det(AB) = \det(A)\det(B)$ durch vollständige Induktion über $r$. Der Fall $r = 1$ ist bereits erledigt, denn in diesem Fall ist $A$ selbst eine Elementarmatrix. Ist nun $r > 1$ und setzen wir die Gleichung für Werte $< r$ voraus, dann können wir $A' = T_1 \cdot \ldots \cdot T_{r-1}$ setzen und erhalten

$$\det(AB) \quad = \quad \det(A'T_rB) \quad = \quad \det(A')\det(T_rB) \quad =$$

$$\det(A')\det(T_r)\det(B) \quad = \quad \det(A'T_r)(\det B) \quad = \quad \det(A)\det(B). \qquad \square$$

Gelegentlich ist auch die folgende Rechenregel nützlich.

**(11.4) Satz** Sei $M \in \mathscr{M}_{n,K}$ eine Blockmatrix der Form

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

mit $A \in \mathscr{M}_{r,K}$, $B \in \mathscr{M}_{r\times(n-r),K}$ und $C \in \mathscr{M}_{n-r,K}$. Dann gilt $\det(M) = \det(A)\det(C)$.

***Beweis:*** Aus §3 ist bekannt, dass es Elementarmatrizen $T_1, ..., T_k \in \mathrm{GL}_r(K)$ und $U_1, ..., U_\ell \in \mathrm{GL}_n(K)$ gibt, so dass $A' = T_k \cdots T_1 A$ und $C' = U_\ell \cdots U_1 C$ in Zeilenstufenform vorliegen, also insbesondere obere Dreiecksmatrizen sind. Für jede Matrix $B' \in \mathscr{M}_{r\times(n-r),K}$ ist dann auch die Blockmatrix

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

eine obere Dreiecksmatrix, und es gilt

$$\det M' \quad = \quad \left(\prod_{i=1}^{r} a'_{ii}\right)\left(\prod_{j=1}^{n} c'_{jj}\right) \quad = \quad \det(A')\det(C').$$

Man überprüft unmittelbar, dass mit $T_i$ und $U_j$ auch

$$\hat{T}_i = \begin{pmatrix} T_i & 0 \\ 0 & E^{(n-r)} \end{pmatrix} \qquad \text{und} \qquad \hat{U}_j = \begin{pmatrix} E^{(r)} & 0 \\ 0 & U_j \end{pmatrix}$$

Elementarmatrizen sind, für die $\det T_i = \det \hat{T}_i$ sowie $\det U_j = \det \hat{U}_j$ gilt. Auf Grund der Rechenregel für Produkte von Blockmatrizen (Abschnitt §2, Seite 15) gilt

$$\hat{T}_k \cdots \hat{T}_1 M \quad = \quad \begin{pmatrix} T_k & 0 \\ 0 & E^{(n-r)} \end{pmatrix} \cdots \begin{pmatrix} T_1 & 0 \\ 0 & E^{(n-r)} \end{pmatrix} \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \quad = \quad \begin{pmatrix} T_k \cdots T_1 A & T_k \cdots T_1 B \\ 0 & C \end{pmatrix} \quad = \quad \begin{pmatrix} A' & B' \\ 0 & C \end{pmatrix}$$

mit $B' = T_k \cdots T_1 B$ und ebenso

$$\hat{U}_\ell \cdots \hat{U}_1 \begin{pmatrix} A' & B' \\ 0 & C \end{pmatrix} \quad = \quad \begin{pmatrix} E^{(r)} & 0 \\ 0 & U_\ell \end{pmatrix} \cdots \begin{pmatrix} E^{(r)} & 0 \\ 0 & U_1 \end{pmatrix} \begin{pmatrix} A' & B' \\ 0 & C \end{pmatrix} \quad = \quad \begin{pmatrix} A' & B' \\ 0 & U_\ell \cdots U_1 C \end{pmatrix} \quad = \quad \begin{pmatrix} A' & B' \\ 0 & C' \end{pmatrix}$$

Damit erhalten wir insgesamt

$$\left(\prod_{i=1}^{k} \det T_i\right)\left(\prod_{j=1}^{\ell} \det U_j\right)\det(A)\det(C) \quad = \quad \det(A')\det(C') \quad = \quad \det M' \quad =$$

$$\left(\prod_{i=1}^{k} \det \hat{T}_i\right)\left(\prod_{j=1}^{\ell} \det \hat{U}_j\right)\det(M) \quad = \quad \left(\prod_{i=1}^{k} \det T_i\right)\left(\prod_{j=1}^{\ell} \det U_j\right)\det(M)$$

und somit $\det(A)\det(C) = \det(M)$. □

Bevor wir als letzte wichtige Rechenregel den Laplaceschen Entwicklungssatz in Angriff nehmen, bemerken wir zunächst, dass sich die Determinate beim Übergang zur transponierten Matrix nicht ändert.

**(11.5) Satz** Für alle $A \in \mathscr{M}_{n,K}$ gilt $\det(A) = \det({}^{\mathrm{t}}A)$.

***Beweis:*** Hier verwenden wir zum Beweis die Leibnizformel. Setzen wir $B = {}^{\mathrm{t}}A$, dann sind die Einträge $b_{ij}$ von $B$ gegeben durch $b_{ij} = a_{ji}$ für $1 \le i,j \le n$. Es gilt

$$\det({}^{\mathrm{t}}A) \quad = \quad \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\prod_{i=1}^{n} b_{i\sigma(i)} \quad = \quad \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\prod_{i=1}^{n} a_{\sigma(i)i} \quad = \quad \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\prod_{i=1}^{n} a_{\sigma(i)\sigma^{-1}(\sigma(i))}$$

Weil jedes $\sigma$ bijektiv ist, durchläuft mit $i$ auch $\sigma(i)$ alle Zahlen in $\{1,\ldots,n\}$. Wir können im Produkt also $\sigma(i)$ durch $i$ ersetzen und erhalten

$$\prod_{i=1}^{n} a_{\sigma(i)\sigma^{-1}(\sigma(i))} \quad = \quad \prod_{i=1}^{n} a_{i\sigma^{-1}(i)}.$$

Nach (10.10) gilt $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$ für alle $\sigma \in S_n$. Somit gilt insgesamt

$$\sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\prod_{i=1}^{n} a_{\sigma(i)\sigma^{-1}(\sigma(i))} \quad = \quad \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma^{-1})\prod_{i=1}^{n} a_{i\sigma^{-1}(i)} \quad =$$

$$\sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\prod_{i=1}^{n} a_{i\sigma(i)} \quad = \quad \det(A).$$

Beim vorletzten „=“ haben wir verwendet, dass mit $\sigma$ auch $\sigma^{-1}$ die gesamte Gruppe $S_n$ durchläuft, so dass wir in jedem Summanden jeweils $\sigma^{-1}$ durch $\sigma$ ersetzen können. □

Bevor wird den Laplaceschen Entwicklungssatz formulieren und beweisen können, benötigen wir ein wenig zusätzliche Notation. Sei $A = (a_{ij}) \in \mathscr{M}_{n,K}$. Für beliebige $i,j \in \{1,\ldots,n\}$ sei dann $A'_{ij} = (a'_{k\ell})$ die Matrix in $\mathscr{M}_{n,K}$ mit den Einträgen

$$a'_{k\ell} \quad = \quad \begin{cases} a_{k\ell} & \text{für } k \neq i, \ell \neq j \\ 0 & \text{für } k = i, \ell \neq j \\ 0 & \text{für } k \neq i, \ell = j \\ 1 & \text{für } k = i, \ell = j \end{cases}$$

Die Matrix $A'_{ij}$ hat also die Form

$$A'_{ij} \quad = \quad \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & 0 & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & 0 & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}$$

Mit $A_{ij} \in \mathscr{M}_{n-1,K}$ bezeichnen wir die Matrix, die aus $A$ durch Streichung der $i$-ten Zeile und $j$-ten Spalten zu Stande kommt, also

$$A_{ij} \quad = \quad \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}$$

**(11.6) Lemma** Es gilt $\det(A'_{ij}) = (-1)^{i+j}\det(A_{ij})$.

***Beweis:*** Durch $i-1$ Zeilenvertauschungen bewegt man die $i$-te Zeile von $A'_{ij}$ nach oben in die erste Zeile. Anschließend führt man $j-1$ Spaltenvertauschungen durch, um die $j$-te Spalte ganz nach links zu bewegen. Insgesamt ändert sich das Vorzeichen dadurch um den Faktor $(-1)^{(i-1)+(j-1)} = (-1)^{i+j}$, und man erhält eine Matrix der Form

$$\begin{pmatrix} 1 & 0 \\ 0 & A_{ij} \end{pmatrix}$$

Auf Grund der Formel in (11.4) über die Determinante von Blockmatrizen stimmt die Determinante dieser Matrix mit $\det(A_{ij})$ überein. □

**(11.7) Definition** Sei $A \in \mathscr{M}_{n,K}$. Die Matrix $\tilde{A} \in \mathscr{M}_{n,K}$ mit den Einträgen

$$\tilde{a}_{ij} \quad = \quad \det(A'_{ji}) \quad = \quad (-1)^{i+j}\det(A_{ji})$$

wird die zu $A$ ***komplementäre*** oder ***adjunkte*** Matrix genannt. (Man beachte die Vertauschung von Zeilen- und Spaltenindex, dies ist kein Tippfehler!)

**(11.8) Lemma** Sei $A \in \mathscr{M}_{n,K}$. Dann gilt für $1 \le i, j \le n$ jeweils

$$\det(A'_{ij}) \quad = \quad \det(a_{1\bullet}, ..., a_{i-1\bullet}, e_j, a_{i+1\bullet}, ..., a_{n\bullet})$$

***Beweis:*** Sei $B$ die Matrix auf der rechten Seite der Gleichung. Dann kann $B$ in $A'_{ij}$ umgeformt werden, indem man für alle $k$ mit $1 \le k \le n$ und $i \ne k$ jeweils das $a_{kj}$-fache der $i$-ten Zeile von der $k$-ten Zeile von $B$ subtrahiert. Dies zeigt, dass die beiden Determinanten übereinstimmen. □

**(11.9) Proposition** Sei $\tilde{A}$ die zu $A \in \mathscr{M}_{n,K}$ komplementäre Matrix. Dann gilt

$$\tilde{A}A \quad = \quad A\tilde{A} \quad = \quad \det(A)E^{(n)}.$$

***Beweis:*** Zunächst zeigen wir die Gleichung $A\tilde{A} = \det(A)E^{(n)}$. Für den Eintrag von $A\tilde{A}$ an der Stelle $(i,k)$ gilt nach (11.8) und der Rechenregeln für die Determinante jeweils

$$\sum_{j=1}^{n} a_{ij}\tilde{a}_{jk} \quad = \quad \sum_{j=1}^{n} a_{ij}\det(A'_{kj}) \quad = \quad \sum_{j=1}^{n} a_{ij}\det(a_{1\bullet}, ..., a_{k-1\bullet}, e_j, a_{k+1\bullet}, ..., a_{n\bullet}) \quad =$$

$$\det\left(a_{1\bullet}, ..., a_{k-1\bullet}, \sum_{j=1}^{n} a_{ij}e_j, a_{k+1\bullet}, ..., a_{n\bullet}\right) \quad = \quad \det(a_{1\bullet}, ..., a_{k-1\bullet}, a_{i\bullet}, a_{k+1\bullet}, ..., a_{n\bullet}) \quad = \quad \delta_{ik}\det(A).$$

Der Eintrag von $A\tilde{A}$ an der Stelle $(i,k)$ stimt also mit dem entsprechenden Eintrag von $(\det A)E^{(n)}$ überein. Für den Beweis der zweiten Gleichung bemerken wir zunächst, dass ${}^{\mathrm{t}}\tilde{A}$ nach Definition die zu ${}^{\mathrm{t}}A$ komplementäre Matrix ist. Auf Grund der bereits bewiesenen Gleichung gilt also

$$\tilde{A}A \quad = \quad {}^{\mathrm{t}}({}^{\mathrm{t}}A\,{}^{\mathrm{t}}\tilde{A}) \quad = \quad {}^{\mathrm{t}}(\det(A)E^{(n)}) \quad = \quad \det(A)E^{(n)}. \qquad \square$$

Wir demonstrieren die Berechnung der komplentären Matrix am Beispiel von

$$A \quad = \quad \begin{pmatrix} 6 & 0 & 4 \\ -2 & 2 & -2 \\ 14 & 0 & 10 \end{pmatrix}.$$

Die Einträge der komplementären Matrix $\tilde{A}$ sind gegeben durch

$$\tilde{a}_{11} = (-1)^{1+1}\det(A_{11}) = \det\begin{pmatrix} 2 & -2 \\ 0 & 10 \end{pmatrix} = 20 \quad , \quad \tilde{a}_{12} = (-1)^{1+2}\det(A_{21}) = -\det\begin{pmatrix} 0 & 4 \\ 0 & 10 \end{pmatrix} = 0 \quad ,$$

$$\tilde{a}_{13} = (-1)^{1+3}\det(A_{31}) = \det\begin{pmatrix} 0 & 4 \\ 2 & -2 \end{pmatrix} = -8 \quad , \quad \tilde{a}_{21} = (-1)^{2+1}\det(A_{12}) = -\det\begin{pmatrix} -2 & -2 \\ 14 & 10 \end{pmatrix} = -8 \quad ,$$

$$\tilde{a}_{22} = (-1)^{2+2}\det(A_{22}) = \det\begin{pmatrix} 6 & 4 \\ 14 & 10 \end{pmatrix} = 4 \quad , \quad \tilde{a}_{23} = (-1)^{2+3}\det(A_{32}) = -\det\begin{pmatrix} 6 & 4 \\ -2 & -2 \end{pmatrix} = -(-4) = 4 \quad ,$$

$$\tilde{a}_{31} = (-1)^{3+1}\det(A_{13}) = \det\begin{pmatrix} -2 & 2 \\ 14 & 0 \end{pmatrix} = -28 \quad , \quad \tilde{a}_{32} = (-1)^{3+2}\det(A_{23}) = -\det\begin{pmatrix} 6 & 0 \\ 14 & 0 \end{pmatrix} = 0 \quad ,$$

$$\tilde{a}_{33} = (-1)^{3+3}\det(A_{33}) = \det\begin{pmatrix} 6 & 0 \\ -2 & 2 \end{pmatrix} = 12$$

also gilt

$$\tilde{A} \quad = \quad \begin{pmatrix} 20 & 0 & -8 \\ -8 & 4 & 4 \\ -28 & 0 & 12 \end{pmatrix}.$$

Wie nach (11.9) zu erwarten, gilt

$$A \cdot \tilde{A} \quad = \quad \begin{pmatrix} 6 & 0 & 4 \\ -2 & 2 & -2 \\ 14 & 0 & 10 \end{pmatrix}\begin{pmatrix} 20 & 0 & -8 \\ -8 & 4 & 4 \\ -28 & 0 & 12 \end{pmatrix} \quad = \quad \begin{pmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{pmatrix} \quad = \quad 8 \cdot E^{(3)}.$$

Daraus folgt $A \cdot (\frac{1}{8}\tilde{A}) = E^{(3)}$. Die zu $A$ inverse Matrix ist also gegeben durch

$$A^{-1} \quad = \quad \tfrac{1}{8}\tilde{A} \quad = \quad \begin{pmatrix} \frac{5}{2} & 0 & -1 \\ -1 & \frac{1}{2} & \frac{1}{2} \\ -\frac{7}{2} & 0 & \frac{3}{2} \end{pmatrix}.$$

**(11.10) Satz** *(Laplacescher Entwicklungssatz)*
Sei $A \in \mathscr{M}_{n,K}$.

(i) Für alle $i \in \{1,...,n\}$ gilt $\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$.

(ii) Für alle $j \in \{1,...,n\}$ gilt $\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$.

Wird die Determinante von $A$ mittels (i) berechnet, spricht man von einer ***Entwicklung zur $i$-ten Zeile***. Die Berechnung mittels (ii) bezeichnet man als Entwicklung zur $j$-ten Spalte.

***Beweis:*** Auf Grund der Proposition ist der Eintrag von $A\tilde{A}$ an der Stelle $(i,i)$ gleich $\det(A)$. Es gilt also

$$\det(A) \quad = \quad \sum_{j=1}^{n} a_{ij}\tilde{a}_{ji} \quad = \quad \sum_{j=1}^{n} a_{ij}\det(A'_{ij}) \quad = \quad \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}).$$

Auch der Eintrag von $\tilde{A}A$ an der Stelle $(j,j)$ ist gleich $\det(A)$. Folglich gilt

$$\det(A) \quad = \quad \sum_{i=1}^{n} \tilde{a}_{ji} a_{ij} \quad = \quad \sum_{i=1}^{n} a_{ij}\det(A'_{ij}) \quad = \quad \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}). \qquad \square$$

Wir demonstrieren die Anwendung des Laplaceschen Entwicklungssatzes anhand der Matrix

$$A = \begin{pmatrix} 1 & 0 & 3 & 7 \\ 0 & 2 & 1 & 4 \\ 1 & 0 & -1 & 1 \\ -2 & 3 & 0 & 2 \end{pmatrix}.$$

Bei jeder Entwicklung müssen die Determinanten der Teilmatrizen nach folgendem Schema mit einem Vorzeichen versehen werden

$$\begin{pmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{pmatrix}$$

denn es ist $(-1)^{1+1} = +1$, $(-1)^{1+2} = -1$, $(-1)^{1+3} = +1$ usw. Die Entwicklung von $\det(A)$ nach der ersten Zeile ergibt nun

$$\begin{aligned} \det(A) &= 1 \cdot \det \begin{pmatrix} 2 & 1 & 4 \\ 0 & -1 & 1 \\ 3 & 0 & 2 \end{pmatrix} - 0 \cdot \det \begin{pmatrix} 0 & 1 & 4 \\ 1 & -1 & 1 \\ -2 & 0 & 2 \end{pmatrix} + 3 \cdot \det \begin{pmatrix} 0 & 2 & 4 \\ 1 & 0 & 1 \\ -2 & 3 & 2 \end{pmatrix} - 7 \cdot \det \begin{pmatrix} 0 & 2 & 1 \\ 1 & 0 & -1 \\ -2 & 3 & 0 \end{pmatrix} \\ &= 1 \cdot 11 + 3 \cdot 4 - 7 \cdot 7 \quad = \quad -26. \end{aligned}$$

Durch Entwicklung zur zweiten Zeile erhalten wir ebenso

$$\begin{aligned} \det(A) &= 1 \cdot \det \begin{pmatrix} 0 & 3 & 7 \\ 2 & 1 & 4 \\ 3 & 0 & 2 \end{pmatrix} - 0 \cdot \det \begin{pmatrix} 1 & 3 & 7 \\ 0 & 1 & 4 \\ -2 & 0 & 2 \end{pmatrix} + (-1) \cdot \det \begin{pmatrix} 1 & 0 & 7 \\ 0 & 2 & 4 \\ -2 & 3 & 2 \end{pmatrix} - 1 \cdot \det \begin{pmatrix} 1 & 0 & 3 \\ 0 & 2 & 1 \\ -2 & 3 & 0 \end{pmatrix} \\ &= 1 \cdot 3 + (-1) \cdot 20 - 1 \cdot 9 \quad = \quad -26. \end{aligned}$$

Eine weitere Möglichkeit wäre die Entwicklung zur zweiten Spalte.

$$\begin{aligned} \det(A) &= -0 \cdot \det \begin{pmatrix} 0 & 1 & 4 \\ 1 & -1 & 1 \\ -2 & 0 & 2 \end{pmatrix} + 2 \cdot \det \begin{pmatrix} 1 & 3 & 7 \\ 1 & -1 & 1 \\ -2 & 0 & 2 \end{pmatrix} - 0 \cdot \det \begin{pmatrix} 1 & 3 & 7 \\ 0 & 1 & 4 \\ -2 & 0 & 2 \end{pmatrix} + 3 \cdot \det \begin{pmatrix} 1 & 3 & 7 \\ 0 & 1 & 4 \\ 1 & -1 & 1 \end{pmatrix} \\ &= 2 \cdot (-28) + 3 \cdot 10 \quad = \quad -26. \end{aligned}$$

Insgesamt gibt es $4 + 4 = 8$ verschiedene Möglichkeiten, $\det(A)$ durch Entwicklung zu einer Zeile oder Spalte zu entwickeln, also noch fünf weitere. Besonders günstig ist natürlich, eine Zeile oder Spalte mit möglichst vielen Nulleinträgen zu wählen, weil dann nur wenige Determinanten von $3 \times 3$-Matrizen ausgerechnet werden müssen.

# § 12. Eigenwerte und Eigenvektoren

***Inhaltsübersicht***

Ein *Eigenvektor* einer linearen Abbildung $\phi : V \to V$ ist ein Vektor $v \neq 0_V$, der von $\phi$ auf ein Vielfaches $\lambda v$ von sich selbst abgebildet wird. In diesem Fall bezeichnet man $\lambda$ dann als *Eigenwert* von $\phi$. Zunächst werden wir sehen, wie man Eigenvektoren und -werte einer vorgegebenen linearen Abbildung $\phi$ ausrechnet; für die Eigenwerte benötigt man das *charakteristische Polynom* einer Matrix.

In einigen Fällen ist es mit Hilfe der Eigenwerte und -vektoren möglich, eine Basis $\mathscr{A}$ zu finden, bezüglich der die Darstellungsmatrix $\mathscr{M}_{\mathscr{A}}^{\mathscr{A}}(\phi)$ eine besonders einfache Form, nämlich Diagonalgestalt, annimmt. Man bezeichnet $\phi$ in diesem Fall als *diagonalisierbar*. Hauptergebnis dieses Abschnitts wird ein einfaches Kriterium sein, mit dem sich die Diagonalisierbarkeit von $\phi$ testen lässt. Wie wir im vierten Semester sehen werden, spielt die hier entwickelte Theorie beim Lösen von Differentialgleichungen eine wichtige Rolle. Auch in der Physik, zum Beispiel in der klassischen Mechanik und in der Quantenmechanik, wird mit diagonalisierbaren linearen Abbildungen gearbeitet.

***Wichtige Definitionen und Sätze***

- Eigenwert und Eigenvektor eines Endomorphismus
- Eigenwert und Eigenvektor einer Matrix
- Ähnlichkeit von Matrizen
- Vielfachheit der Nullstelle eines Polynoms
- charakteristisches Polynom eines Endomorphismus, einer Matrix
- Die Eigenwerte eines Endomorphismus (bzw. einer Matrix) sind die Nullstellen des charakteristischen Polynoms.
- Diagonalmatrix, Diagonalisierbarkeit von Endomorphismen und Matrizen
- algebraische und geometrische Vielfachheit eines Eigenwerts
- Diagonalisierbarkeitskriterium für Endomorphismen und Matrizen

Im gesamten Text bezeichnet $K$ stets einen beliebigen Körper.

**(12.1) Definition** Sei $V$ ein $K$-Vektorraum und $\phi$ ein Endomorphismus von $V$, also eine lineare Abbildung $V \to V$. Man nennt

(i) $\lambda \in K$ einen ***Eigenwert*** von $\phi$, wenn es ein $v \in V$ mit $v \neq 0_V$ und $\phi(v) = \lambda v$, und

(ii) $v \in V$ einen ***Eigenvektor*** von $\phi$, wenn $v \neq 0_V$ ist und ein $\lambda \in K$ mit $\phi(v) = \lambda v$ existiert.

Seien nun $v \in V$ und $\lambda \in K$ vorgegeben. Man nennt $v$ einen Eigenvektor zum Eigenwert $\lambda$, wenn $v \neq 0_V$ und die Gleichung $\phi(v) = \lambda v$ erfüllt ist.

Betrachten wir zum Beispiel über $K = \mathbb{Q}$ die beiden Matrizen

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \qquad \text{und} \qquad B = \begin{pmatrix} -2 & -10 & 0 & -40 \\ -24 & -13 & -8 & -68 \\ 15 & 30 & 3 & 120 \\ 6 & 4 & 2 & 20 \end{pmatrix}.$$

Die Zahlen $1, 2, 3$ sind Eigenwerte von $A$, und die Einheitsvektoren $e_1, e_2, e_3$ sind zugehörige Eigenvektoren, denn es gilt $Ae_1 = 1 \cdot e_1$, $Ae_2 = 2 \cdot e_2$ und $Ae_3 = 3 \cdot e_3$.

Die Matrix $B$ hat die Zahlen $-2$, 3 und 4 als Eigenwerte. Die Vektoren $u = (1, 0, -3, 0)$, $v = (16, 0, -31, -2)$ und $w = (0, 4, 0, -1)$ sind zugehörige Eigenvektoren, denn es gilt

$$\begin{pmatrix} -2 & -10 & 0 & -40 \\ -24 & -13 & -8 & -68 \\ 15 & 30 & 3 & 120 \\ 6 & 4 & 2 & 20 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -3 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 6 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \\ -3 \\ 0 \end{pmatrix} \quad , \quad \begin{pmatrix} -2 & -10 & 0 & -40 \\ -24 & -13 & -8 & -68 \\ 15 & 30 & 3 & 120 \\ 6 & 4 & 2 & 20 \end{pmatrix} \begin{pmatrix} 16 \\ 0 \\ -31 \\ -2 \end{pmatrix} = \begin{pmatrix} 48 \\ 0 \\ -93 \\ -6 \end{pmatrix} = 3 \begin{pmatrix} 16 \\ 0 \\ -31 \\ -2 \end{pmatrix}$$

$$\text{und} \quad \begin{pmatrix} -2 & -10 & 0 & -40 \\ -24 & -13 & -8 & -68 \\ 15 & 30 & 3 & 120 \\ 6 & 4 & 2 & 20 \end{pmatrix} \begin{pmatrix} 0 \\ 4 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 16 \\ 0 \\ -4 \end{pmatrix} = 4 \begin{pmatrix} 0 \\ 4 \\ 0 \\ -1 \end{pmatrix}.$$

**(12.2) Definition** Sei $V$ ein $K$-Vektorraum und $\phi \in \mathrm{End}_K(V)$. Für jedes $\lambda \in K$ bezeichnet man die Menge $\mathrm{Eig}(\phi, \lambda) = \{v \in V \mid \phi(v) = \lambda v\}$ als den ***Eigenraum*** von $\phi$ zum Wert $\lambda \in K$. Er besteht aus dem Nullvektor $0_V$ und den Eigenvektoren zum Eigenwert $\lambda$.

**(12.3) Proposition** Sei $V$ ein $K$-Vektorraum und $\phi \in \mathrm{End}_K(V)$. Für jedes $\lambda \in K$ ist der Eigenraum gegeben durch $\mathrm{Eig}(\phi, \lambda) = \ker(\phi - \lambda \mathrm{id}_V)$. Das Element $\lambda$ ist ein Eigenwert von $\phi$ genau dann, wenn $\mathrm{Eig}(\phi, \lambda) \neq \{0_V\}$ gilt.

***Beweis:*** Für jeden Vektor $v \in V$ gilt die Äquivalenz

$$v \in \mathrm{Eig}(\phi, \lambda) \quad \Leftrightarrow \quad \phi(v) = \lambda v \quad \Leftrightarrow \quad \phi(v) - \lambda v = 0_V \quad \Leftrightarrow \quad \phi(v) - \lambda \mathrm{id}_V(v) = 0_V$$
$$\Leftrightarrow \quad (\phi - \lambda \mathrm{id}_V)(v) = 0_V \quad \Leftrightarrow \quad v \in \ker(\phi - \lambda \mathrm{id}_V).$$

Daraus folgt $\mathrm{Eig}(\phi, \lambda) = \ker(\phi - \lambda \mathrm{id}_V)$. Ein Element $\lambda \in K$ ist nach Definition Eigenwert genau dann, wenn ein $v \in V$ mit $v \neq 0_V$ und $\phi(v) = \lambda v$ existiert, also genau dann, wenn es ein Element ungleich $0_V$ in $\mathrm{Eig}(\phi, \lambda)$ gibt. Weil $0_V$ auf jeden Fall in $\mathrm{Eig}(\phi, \lambda)$ liegt, ist dies wiederum äquivalent zu $\mathrm{Eig}(\phi, \lambda) \neq \{0_V\}$. □

Aus §5 ist bekannt, dass Kerne von linearen Abbildungen Untervektorräume sind. Die Proposition zeigt also, dass $\mathrm{Eig}(\phi, \lambda)$ für jedes $\lambda \in K$ und jedes $\phi \in \mathrm{End}_K(V)$ ein Untervektorraum von $V$ ist. Natürlich kann diese Eigenschaft auch direkt nachgerechnet werden.

Als nächstes sehen wir uns an, wie das Matrixkalkül zur Untersuchung von Eigenwerten und Eigenvektoren eingesetzt werden kann. Sei nun $A \in \mathscr{M}_{n,K}$ eine quadratische Matrix. Wir bezeichnen $v \in K^n$ als einen ***Eigenvektor von*** $A$, wenn $v$ ein Eigenvektor der Abbildung $\phi_A : K^n \to K^n$, $v \mapsto Av$ ist.

Ebenso sind die ***Eigenwerte von*** $A$ nach Definition die Eigenwerte des Endomorphismus $\phi_A$. Für jedes $\lambda \in K$ definieren wir

$$\mathrm{Eig}(A,\lambda) \quad = \quad \mathrm{Eig}(\phi_A,\lambda) \quad = \quad \{v \in K^n \mid Av = \lambda v\}.$$

Wiederum besteht $\mathrm{Eig}(A,\lambda)$ aus den Eigenvektoren von $A$ zum Eigenwert $\lambda$ und dem Nullvektor $0_{K^n}$, und darüber hinaus gilt

$$\mathrm{Eig}(A,\lambda) \quad = \quad \ker(A - \lambda E^{(n)}).$$

Der Kern einer Matrix kann mit dem Gaußverfahren berechnet werden, also erhalten wir auch die Eigenräume einer Matrix mit dem Gaußverfahren. Betrachten wir zum Beispiel den Eigenraum $\mathrm{Eig}(A,-2)$ der Matrix

$$A \quad = \quad \begin{pmatrix} -2 & -5 & 10 \\ -20 & -17 & 40 \\ -10 & -10 & 23 \end{pmatrix}$$

zum Eigenwert $\lambda = -2$. Es gilt

$$\mathrm{Eig}(A,-2) \quad = \quad \ker(A + 2E^{(3)}) \quad = \quad \ker \begin{pmatrix} 0 & -5 & 10 \\ -20 & -15 & 40 \\ -10 & -10 & 25 \end{pmatrix}.$$

Wir wenden auf diese Matrix den Gauß-Algorithmus an.

$$\begin{pmatrix} 0 & -5 & 10 \\ -20 & -15 & 40 \\ -10 & -10 & 25 \end{pmatrix} \mapsto \begin{pmatrix} 0 & 1 & -2 \\ 4 & 3 & -8 \\ 2 & 2 & -5 \end{pmatrix} \mapsto \begin{pmatrix} 2 & 2 & -5 \\ 4 & 3 & -8 \\ 0 & 1 & -2 \end{pmatrix} \mapsto \begin{pmatrix} 2 & 2 & -5 \\ 0 & -1 & 2 \\ 0 & 1 & -2 \end{pmatrix} \mapsto$$

$$\begin{pmatrix} 2 & 2 & -5 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

Die letzte Matrix entspricht dem LGS $x_1 - \frac{1}{2}x_3 = 0$, $x_2 - 2x_3 = 0$, und wir erhalten

$$\mathrm{Eig}(A,-2) \quad = \quad \ker(A + 2E^{(3)}) \quad = \quad \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{Q}^3 \;\middle|\; x_1 - \tfrac{1}{2}x_3 = 0, x_2 - 2x_3 = 0 \right\} \quad = \quad \left\{ \begin{pmatrix} \frac{1}{2}x_3 \\ 2x_3 \\ x_3 \end{pmatrix} \;\middle|\; x_3 \in \mathbb{R} \right\}$$

$$= \quad \mathrm{lin}\left\{ \begin{pmatrix} \frac{1}{2} \\ 2 \\ 1 \end{pmatrix} \right\} \quad = \quad \mathrm{lin}\left\{ \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} \right\}.$$

Tatsächlich ist $(1,4,2)$ ein Eigenvektor von $A$ zum Eigenwert $-2$, denn es gilt

$$A \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} \quad = \quad \begin{pmatrix} -2 & -5 & 10 \\ -20 & -17 & 40 \\ -10 & -10 & 23 \end{pmatrix} \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} \quad = \quad \begin{pmatrix} -2 \\ -8 \\ -4 \end{pmatrix} \quad = \quad (-2) \cdot \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix}.$$

Aus § 9 wissen wir, dass jede lineare Abbildung $\phi : V \to W$ zwischen endlich-dimensionalen $K$-Vektorräumen $V, W$ auf eindeutige Weise durch eine Matrix beschrieben werden kann, sobald man für $V$ und $W$ Basen festgelegt hat. Ist $\mathscr{A}$ eine Basis von $V$ und $\mathscr{B}$ eine Basis von $W$, dann haben wir die Bezeichnung

$$\mathscr{M}_{\mathscr{B}}^{\mathscr{A}}(\phi)$$

für die Darstellungsmatrix von $\phi$ bezüglich der Basen $\mathscr{A}$ und $\mathscr{B}$ eingeführt. Wir erinnern an den wichtigen Zusammenhang

$$\mathscr{M}_{\mathscr{B}}^{\mathscr{A}}(\phi)\Phi_{\mathscr{A}}(v) \quad = \quad \Phi_{\mathscr{B}}(\phi(v)).$$

Die Darstellungsmatrix $\mathscr{M}_{\mathscr{B}}^{\mathscr{A}}(\phi)$ ist also dadurch gekennzeichnet, dass sie den Vektor $v$ in $\mathscr{A}$-Koordinaten entgegennimmt und den Vektor $\phi(v)$ in $\mathscr{B}$-Koordinaten als Ergebnis liefert.

Ist nun $V = W$, die Abbildung $\phi$ also ein ***Endomorphismus*** des Vektorraums $V$, dann braucht man nur noch ***eine*** Basis von $V$, um $\phi$ zu beschreiben. Wir setzen $\mathscr{M}_{\mathscr{A}}(\phi) = \mathscr{M}_{\mathscr{A}}^{\mathscr{A}}(\phi)$ und nennen diese quadratische Matrix die ***Darstellungsmatrix*** von $\phi$ bezüglich der Basis $\mathscr{A}$.

**(12.4) Definition** Zwei Matrizen $A, B \in \mathscr{M}_{n,K}$ werden ***ähnlich*** genannt, wenn eine invertierbare Matrix $T \in \mathrm{GL}_n(K)$ mit $B = TAT^{-1}$ existiert. Zwei Matrizen $A, B \in \mathscr{M}_{m\times n,K}$ bezeichnet man als ***äquivalent***, wenn Elemente $S \in \mathrm{GL}_m(K)$ und $T \in \mathrm{GL}_n(K)$ mit $B = SAT$ existieren.

Ähnliche Matrizen sind also stets äquivalent zueinander. Die Umkehrung ist im Allgemeinen falsch.

**(12.5) Proposition** Sei $V$ ein $n$-dimensionaler $K$-Vektorraum, und sei $\phi \in \mathrm{End}_K(V)$. Sind $A, B \in \mathscr{M}_{n,K}$ Darstellungsmatrizen von $\phi$ bezüglich unterschiedlicher Basen von $V$, dann sind $A$ und $B$ ähnlich.

***Beweis:*** Seien $\mathscr{A}$ und $\mathscr{B}$ Basen von $V$, so dass $A = \mathscr{M}_{\mathscr{A}}(\phi)$ und $B = \mathscr{M}_{\mathscr{B}}(\phi)$ erfüllt ist. Sei außerdem $T = \mathscr{T}_{\mathscr{B}}^{\mathscr{A}}$ die Matrix des Basiswechsels von $A$ nach $B$, also die eindeutig bestimmte Matrix $T \in \mathrm{GL}_n(K)$ mit $T\Phi_{\mathscr{A}}(v) = \Phi_{\mathscr{B}}(v)$ für alle $v \in V$. Auf Grund der Transformationsformel (9.16) gilt

$$B \quad = \quad \mathscr{M}_{\mathscr{B}}(\phi) \quad = \quad \mathscr{M}_{\mathscr{B}}^{\mathscr{B}}(\phi) \quad = \quad \mathscr{T}_{\mathscr{B}}^{\mathscr{A}} \mathscr{M}_{\mathscr{A}}^{\mathscr{A}}(\phi) \mathscr{T}_{\mathscr{A}}^{\mathscr{B}} \quad = \quad \mathscr{T}_{\mathscr{B}}^{\mathscr{A}} \mathscr{M}_{\mathscr{A}}(\phi) \left(\mathscr{T}_{\mathscr{B}}^{\mathscr{A}}\right)^{-1} \quad = \quad TAT^{-1}. \qquad \square$$

**(12.6) Proposition** Sei $V$ ein endlich-dimensionaler $K$-Vektorraum, $\phi \in \mathrm{End}_K(V)$ und $A \in \mathscr{M}_{n,K}$ die Darstellungsmatrix von $\phi$ bezüglich einer beliebigen Basis $\mathscr{A}$ von $V$. Genau dann ist $v \in V$ ein Eigenvektor von $\phi$ zu einem Eigenwert $\lambda \in K$, wenn der Koordinatenvektor $\Phi_{\mathscr{A}}(v)$ ein Eigenvektor von $A$ zum Eigenwert $\lambda$ ist.

***Beweis:*** Nach Definition der Darstellungsmatrix gilt $A\Phi_{\mathscr{A}}(v) = \mathscr{M}_{\mathscr{A}}(\phi)\Phi_{\mathscr{A}}(v) = \Phi_{\mathscr{A}}(\phi(v))$ für jedes $v \in V$. Sei nun $\lambda \in K$ vorgegeben. Ein Vektor $v$ ist genau dann Eigenvektor von $\phi$ zum Eigenwert $\lambda$, wenn $v \neq 0_V$ und $\phi(v) = \lambda v$ erfüllt ist. Auf Grund der Bijektivitiät und der Linearität der Koordinatenabbildung $\Phi_{\mathscr{A}}$ ist $v \neq 0_V$ äquivalent zu $\Phi_{\mathscr{A}}(v) \neq 0_{K^n}$, außerdem gilt

$$\phi(v) = \lambda v \quad \Leftrightarrow \quad \Phi_{\mathscr{A}}(\phi(v)) = \Phi_{\mathscr{A}}(\lambda v) \quad \Leftrightarrow \quad \Phi_{\mathscr{A}}(\phi(v)) = \lambda\Phi_{\mathscr{A}}(v) \quad \Leftrightarrow \quad A\Phi_{\mathscr{A}}(v) = \lambda\Phi_{\mathscr{A}}(v). \qquad \square$$

Die Proposition zeigt insbesondere, dass die Eigenwerte von $\phi$ genau die Eigenwerte der Darstellungsmatrix $\mathscr{M}_{\mathscr{A}}(\phi)$ sind.

Im folgenden beschäftigen wir uns mit der Frage, wie man die Eigenwerte eines Endomorphismus findet. Dafür benötigen wir einige Grundbegriffe und elementare Aussagen über Polynome. Wir nennen ein Polynom $f \in K[x]$ genau dann ***konstant***, wenn $f = 0_K$ oder $\text{grad}(f) = 0$ gilt, wenn $f$ also in $K$ liegt.

**(12.7) Satz** *(Division mit Rest)*

Seien $f, g \in K[x]$, wobei $g$ nicht-konstant ist. Dann gibt es $q, r \in K[x]$ mit $f = qg + r$, wobei $r = 0_K$ ist oder zumindest $\text{grad}(r) < \text{grad}(g)$ gilt.

Wir verzichten an dieser Stelle auf einen Beweis, weil dieser eher in die Algebra-Vorlesung gehört. Aus dem Schulunterricht ist zumindest für $K = \mathbb{R}$ bekannt, dass die Polynome $q$ und $r$ durch Polynomdivision bestimmt werden können.

Jedem Polynom $f \in K[x]$ kann durch $a \mapsto f(a)$ eine Abbildung $K \to K$ zugeordnet werden, die dadurch zu Stande kommt, dass die Elemente $a \in K$ in die Unbestimmte $x$ eingesetzt werden. Man bezeichnet diese Abbildung auch als die dem Polynom $f$ zugeordnete ***Polynomfunktion***.

Für unendliche Körper gilt allgemein, dass verschiedene Polynome auch verschiedene Polynomfunktionen definieren. Für endliche Körper ist das aber nicht mehr richtig: Beispielsweise definieren die Polynome $f, g \in \mathbb{F}_2[x]$ gegeben durch $f = x$ und $g = x^2$ dieselbe Polynomfunktion, denn es gilt

$$f(\bar{0}) = g(\bar{0}) = \bar{0} \qquad \text{und} \qquad f(\bar{1}) = g(\bar{1}) = \bar{1}.$$

Ein Element $a \in K$ wird ***Nullstelle*** von $f \in K[x]$ genannt, wenn $f(a) = 0_K$ gilt. Man nennt ein Polynom $g \in K[x]$ einen ***Teiler*** von $f$, wenn ein $h \in K[x]$ mit $f = gh$ existiert. Ist $\text{grad}(g) = 1$, dann nennt man $g$ auch einen ***Linearfaktor*** des Polynoms $f$. Auch die folgende Aussage ist im Grunde schon aus der Schulmathematik bekannt.

**(12.8) Satz** Sei $f \in K[x]$ und $a \in K$. Genau dann gilt $f(a) = 0_K$, wenn $x - a$ ein Linearfaktor von $f$ ist.

***Beweis:*** Nach (12.7) gibt es Polynome $g, r \in K[x]$ mit $f = (x-a)g + r$, wobei das Polynom $r$ wegen $r = 0_K$ oder $\mathrm{grad}(r) < \mathrm{grad}(x-a) = 1$ konstant ist. Ist nun $a$ eine Nullstelle von $f$, dann gilt $r = r(a) = f(a) - (a-a)g(a) = 0_K - 0_K = 0_K$ und somit $f = (x-a)g$. Ist umgekehrt $x - a$ ein Linearfaktor von $f$, dann gibt es ein $g \in K[x]$ mit $f = (x-a)g$, und es folgt $f(a) = (a-a)g(a) = 0_K$. □

**(12.9) Definition** Sei $f \in K[x]$ mit $f \neq 0_K$ und $a \in K$ eine Nullstelle von $f$. Das maximale $r \in \mathbb{N}$ mit der Eigenschaft, dass $(x-a)^r$ ein Teiler von $f$ ist, wird die ***Vielfachheit*** $\mu(f,a)$ der Nullstelle $a$ genannt.

Nach (12.7) gilt also $\mu(f,a) \geq 1$ für jede Nullstelle $a$ von $f$. Ist $f(a) \neq 0_K$, dann setzen wir $\mu(f,a) = 0$. Das folgende Kriterium ist für die Bestimmung der Vielfachheit einer Nullstelle hilfreich.

**(12.10) Proposition** Sei $f \in K[x]$ ein Polynom mit einer Zerlegung $f = (x-a)^r g$, wobei $r \in \mathbb{N}_0$ und $g(a) \neq 0_K$ ist. Dann gilt $r = \mu(f,a)$.

***Beweis:*** Die Gleichung $f = (x-a)^r g$ zeigt jedenfalls, dass $\mu(f,a) \geq r$ gilt. Nehmen wir nun an, dass sogar $\mu(f,a) > r$ erfüllt ist. Dann gibt es ein $h \in K[x]$ mit $f = (x-a)^{r+1}h$. Teilt man die Polynomgleichung

$$(x-a)^r g \quad = \quad (x-a)^{r+1}h$$

durch $(x-a)^r$, dann folgt $g = (x-a)h$ und $g(a) = (a-a)h(a) = 0_K$, im Widerspruch zur Voraussetzung $g(a) \neq 0_K$. □

Sei $f \in K[x]$ ein Polynom vom Grad $\geq 1$. Man sagt, $f$ ***zerfällt in Linearfaktoren***, wenn es als Produkt von Linearfaktoren geschrieben werden kann. In ausgeschriebener Form bedeutet dies, dass Elemente $c, \lambda_1, ..., \lambda_r \in K$ existieren, so dass

$$f \quad = \quad c\prod_{k=1}^{r}(x-\lambda_k) \qquad \text{gilt.}$$

Ein Körper $K$ wird ***algebraisch abgeschlossen*** genannt, wenn jedes Polynom vom Grad $\geq 1$ in $K[x]$ in Linearfaktoren zerfällt. In der Funktionentheorie zeigt man, dass zum Beispiel der Körper $\mathbb{C}$ der komplexen Zahlen diese Eigenschaft besitzt. Dagegen ist $\mathbb{R}$ nicht algebraisch abgeschlossen, denn das Polynom $x^2+1$ hat keine Nullstellen in $\mathbb{R}$ und kann deshalb nach (12.8) nicht in Linearfaktoren zerlegt werden. In der Algebra-Vorlesung wird aber gezeigt, dass zu einem Körper $K$ ein algebraisch abgeschlossener Erweiterungskörper existiert. Im Fall $K = \mathbb{R}$ ist dies gerade der Körper $\mathbb{C}$.

Nun werden wir sehen, inwiefern Polynome bei der Bestimmung der Eigenwerte einer Matrix weiterhelfen.

**(12.11) Definition** Für jede Matrix $A \in \mathscr{M}_{n,K}$ nennt man

$$\chi_A \quad = \quad (-1)^n \det(A - xE^{(n)}) \quad = \quad \det(xE^{(n)} - A) \quad \in K[x]$$

das ***charakteristische Polynom*** von $A$.

**(12.12) Satz** Die Eigenwerte einer Matrix $A \in \mathscr{M}_{n,K}$ sind genau die Nullstellen des charakteristischen Polynoms $\chi_A$.

***Beweis:*** Für jedes $\lambda \in K$ gilt $\mathrm{Eig}(A,\lambda) = \ker(A - \lambda E^{(n)})$. Genau dann ist $\lambda$ ein Eigenwert von $A$, wenn die Ungleichung $\ker(A - \lambda E^{(n)}) \neq \{0_V\}$ gilt (vgl. (12.3)). Nach dem Dimensionssatz für lineare Abbildungen gilt weiter

$$\dim\ker(A - \lambda E^{(n)}) > 0 \quad \Leftrightarrow \quad \mathrm{rg}(A - \lambda E^{(n)}) < n \quad \Leftrightarrow \quad \det(A - \lambda E^{(n)}) = 0_K \quad \Leftrightarrow \quad \chi_A(\lambda) = 0_K \qquad \square$$

**(12.13) Definition** Ist $V$ ein endlich-dimensionaler $K$-Vektorraum, $\phi \in \mathrm{End}_K(V)$ und $A \in \mathscr{M}_{n,K}$ die Darstellungsmatrix von $V$ bezüglich einer beliebig gewählen Basis, dann bezeichnen wir $\chi_\phi = \chi_A$ als ***charakteristisches Polynom*** von $\phi$.

**(12.14) Proposition** Das charakteristische Polynom $\chi_\phi$ ist unabhängig von der gewählten Basis des Vektorraums $V$.

***Beweis:*** Sind $A, B \in \mathscr{M}_{n,K}$ die Darstellungsmatrizen von $\phi$ bezüglich verschiedener Basen, dann sind $A$ und $B$ nach (12.5) ähnlich. Es gibt also ein $T \in \mathrm{GL}_n(K)$ mit $B = TAT^{-1}$. Auf Grund der Multiplikativität der Determinantenfunktion folgt

$$\begin{aligned} \chi_B \quad &= \quad \det(xE^{(n)} - B) \quad = \quad \det(T(xE^{(n)})T^{-1} - TAT^{-1}) \quad = \quad \det(T(xE^{(n)} - A)T^{-1}) \\ &= \quad \det(T)\det(xE^{(n)} - A)\det(T)^{-1} \quad = \quad \det(xE^{(n)} - A) \quad = \quad \chi_A. \qquad \square \end{aligned}$$

**(12.15) Folgerung** Auch für jeden Endomorphismus $\phi \in \mathrm{End}_K(V)$ eines endlich-dimensionalen $K$-Vektorraums $V$ gilt: Die Eigenwerte von $\phi$ sind genau die Nullstellen des Polynoms $\chi_\phi$.

***Beweis:*** Sei $A$ die Darstellungsmatrix von $\phi$ bezüglich einer beliebigen Basis von $V$. Dann gilt $\chi_\phi = \chi_A$ nach Definition. Auf Grund von (12.6) sind darüber hinaus die Eigenwerte von $\phi$ genau die Eigenwerte von $A$. Also sind die Eigenwerte von $\phi$ nach (12.12) genau die Nullstellen von $\chi_A$ und damit auch genau die Nullstellen von $\chi_\phi$. □

Als Anwendungsbeispiel bestimmen wir die Eigenwerte der Matrix

$$A \quad = \quad \begin{pmatrix} -2 & -5 & 10 \\ -20 & -17 & 40 \\ -10 & -10 & 23 \end{pmatrix} \quad \in \quad \mathscr{M}_{3,\mathbb{R}}$$

mit Hilfe des charakteristischen Polynoms $\chi_A$. Zunächst ermitteln wir dieses Polynom mit Hilfe der Sarrus-Regel.

$$\chi_A \quad = \quad \det(xE^{(3)} - A) \quad = \quad \det\left( \begin{pmatrix} x & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x \end{pmatrix} - \begin{pmatrix} -2 & -5 & 10 \\ -20 & -17 & 40 \\ -10 & -10 & 23 \end{pmatrix} \right) \quad = \quad \det \begin{pmatrix} x+2 & 5 & -10 \\ 20 & x+17 & -40 \\ 10 & 10 & x-23 \end{pmatrix}$$

$$= \quad (x+2)(x+17)(x-23) + 5\cdot(-40)\cdot 10 + (-10)\cdot 20\cdot 10 - 10\cdot(x+17)\cdot(-10) - 10\cdot(-40)\cdot(x+2)$$

$$-(x-23)\cdot 20\cdot 5 \quad = \quad (x^2+2x+17x+34)(x-23) - 2000 - 2000 + 100x + 1700 + 400x + 800 - 100x + 2300$$

$$= \quad x^3 + 19x^2 + 34x - 23x^2 - 437x - 782 + 400x + 800 \quad = \quad x^3 - 4x^2 - 3x + 18.$$

Durch probeweises Einsetzen von $0, \pm 1, \pm 2, \ldots$ in das Polynom $\chi_A$ finden wir die Nullstelle $-2$. Nach Satz (12.8) ist $x+2$ ein Linearfaktor von $\chi_A$. Durch Polynomdivision erhält man die Zerlegung $\chi_A = (x+2)(x^2-6x+9)$. Die Nullstellen des quadratischen Faktors bestimmt man nun mit der $p$-$q$-Formel aus der Schulmathematik: Die Diskriminante des Polynoms $g = x^2 - 6x + 9$ mit $p = -6$ und $q = 9$ ist $d = p^2 - 4q = (-6)^2 - 4\cdot 9 = 36 - 36 = 0$. Die beiden Nullstellen von $g$ sind mit Vielfachheiten gegeben durch $-\frac{1}{2}p \pm \frac{1}{2}\sqrt{d}$; wegen $d = 0$ ist $-\frac{1}{2}p = -\frac{1}{2}\cdot(-6) = 3$ eine doppelte Nullstelle von $g$. Es gilt also $g = (x-3)^2$ und somit $\chi_A = (x+2)(x-3)^2$.

Also sind $-2$ und $3$ die beiden Eigenwerte von $A$. Die beiden zugehörigen Eigenräume lassen sich wiederum mit dem Gaußverfahren ermitteln. Man erhält

$$\mathrm{Eig}(A,-2) = \mathrm{lin}\left\{ \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} \right\} \quad \text{und} \quad \mathrm{Eig}(A,3) = \mathrm{lin}\left\{ \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \right\}.$$

**(12.16) Definition** Sind $\lambda_1, \ldots, \lambda_n \in K$, dann bezeichnen wir mit $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ die Matrix $D = (d_{ij})$ mit den Einträgen

$$d_{ij} \quad = \quad \begin{cases} \lambda_k & \text{falls } i = j = k \\ 0 & \text{falls } i \neq j. \end{cases}$$

Eine Matrix dieser Form wird ***Diagonalmatrix*** genannt. Man bezeichnet eine Matrix $A \in \mathscr{M}_{n,K}$ als ***diagonalisierbar***, wenn sie ähnlich zu einer Diagonalmatrix ist.

**(12.17) Definition** Einen Endomorphismus $\phi$ eines endlich-dimensionalen $K$-Vektorraums $V$ heißt ***diagonalisierbar***, wenn eine Basis von $V$ existiert, so dass die Darstellungsmatrix von $\phi$ bezüglich dieser Basis eine Diagonalmatrix ist.

**(12.18) Proposition** Sei $V$ ein endlich-dimensionaler $K$-Vektorraum, $\phi \in \mathrm{End}_K(V)$ und $A \in \mathscr{M}_{n,K}$ die Darstellungsmatrix von $\phi$ bezüglich einer geordneten Basis von $V$. Genau dann ist $A$ diagonalisierbar, wenn $\phi$ diagonalisierbar ist.

***Beweis:*** Sei $\mathscr{A}$ eine Basis von $V$, so dass $A = \mathscr{M}_{\mathscr{A}}(\phi)$ erfüllt ist.

„$\Leftarrow$" Weil $\phi$ nach Voraussetzung diagonalisierbar ist, gibt es eine Basis $\mathscr{B}$ von $V$, so dass $D = \mathscr{M}_{\mathscr{B}}(\phi)$ eine Diagonalmatrix ist. Die Matrizen $A$ und $D$ sind also die Darstellungsmatrizen von $\phi$ bezüglich der Basen $\mathscr{A}, \mathscr{B}$. Nach (12.5) sind $A$ und $D$ ähnlich, und damit ist $A$ nach Definition diagonalisierbar.

„$\Rightarrow$" Ist $A$ diagonalisierbar, dann gibt es ein $T \in \mathrm{GL}_n(K)$ mit der Eigenschaft, dass $D = TAT^{-1}$ eine Diagonalmatrix ist. Für jedes $v \in V$ gilt nach Definition der Darstellungsmatrix jeweils

$$A\Phi_{\mathscr{A}}(v) = \Phi_{\mathscr{A}}(\phi(v)) \quad \Leftrightarrow \quad T^{-1}DT\Phi_{\mathscr{A}}(v) = \Phi_{\mathscr{A}}(\phi(v)) \quad \Leftrightarrow \quad DT\Phi_{\mathscr{A}}(v) = T\Phi_{\mathscr{A}}(\phi(v)).$$

Wir wählen nun eine Basis $\mathscr{B}$ mit der Eigenschaft, dass $T = \mathscr{T}^{\mathscr{A}}_{\mathscr{B}}$ erfüllt ist (s.u.) und erhalten

$$D\mathscr{T}^{\mathscr{A}}_{\mathscr{B}}\Phi_{\mathscr{A}}(v) = \mathscr{T}^{\mathscr{A}}_{\mathscr{B}}\Phi_{\mathscr{A}}(\phi(v)) \quad \Leftrightarrow \quad D\Phi_{\mathscr{B}}(v) = \Phi_{\mathscr{B}}(\phi(v))$$

Weil $\mathscr{M}_{\mathscr{B}}(\phi)$ die eindeutig bestimmte Matrix mit $\mathscr{M}_{\mathscr{B}}(\phi)\Phi_{\mathscr{B}}(v) = \Phi_{\mathscr{B}}(\phi(v))$ für alle $v \in V$ ist, folgt $\mathscr{M}_{\mathscr{B}}(\phi) = D$. Dies zeigt, dass $\phi$ diagonalsierbar ist. □

Beim Beweis von (12.18) wurde verwendet

**(12.19) Lemma** Sei $V$ ein endlich-dimensionaler $K$-Vektorraum, $\mathscr{A}$ eine geordnete Basis von $V$ und $T \in \mathrm{GL}_n(K)$ eine invertierbare Matrix. Dann gibt es eine geordnete Basis $\mathscr{B}$ von $V$ mit $T = \mathscr{T}^{\mathscr{A}}_{\mathscr{B}}$.

***Beweis:*** Die Gleichung $T = \mathscr{T}^{\mathscr{A}}_{\mathscr{B}}$ ist äquivalent zu $T^{-1} = \mathscr{T}^{\mathscr{B}}_{\mathscr{A}}$, nach (9.15) (ii). Sei $C = (c_{ij}) = T^{-1}$. Sei nun $\mathscr{A} = (v_1, ..., v_n)$, und sei $\mathscr{B} = (w_1, ..., w_n)$ die gesuchte Basis. Die Gleichung $C = \mathscr{T}^{\mathscr{B}}_{\mathscr{A}}$ ist nach Definition der Transformationsformel gerade äquivalent dazu, dass $\Phi_{\mathscr{A}}(w_j)$ für $1 \le j \le n$ die Spalten von $C$ sind. Dies ist offenbar genau dann der Fall, wenn die Elemente von $\mathscr{B}$ durch

$$w_j \quad = \quad \sum_{i=1}^{n} c_{ij} v_i \qquad \text{für} \quad 1 \le j \le n$$

definiert sind, denn $(c_{1j}, ..., c_{nj})$ ist genau der Koordinatenvektor von $\sum_{i=1}^{n} c_{ij} v_i$ bezüglich $\mathscr{A}$.

Wir müssen nun noch zeigen, dass die so definierten Vektoren $w_1, ..., w_n$ tatsächlich eine Basis von $V$ bilden. Für $1 \le k \le n$ gilt jeweils

$$\sum_{j=1}^{n} t_{jk} w_j \quad = \quad \sum_{j=1}^{n} \sum_{i=1}^{n} t_{jk} c_{ij} v_i .$$

Nun ist $\sum_{j=1}^{n} t_{jk} c_{ij}$ genau der Eintrag von $CT$ an der Stelle $(i,k)$. Wegen $CT = E^{(n)}$ ist der Eintrag also gleich $\delta_{ik}$. Es folgt

$$\sum_{j=1}^{n} t_{jk} w_j \quad = \quad \sum_{i=1}^{n} \delta_{ik} v_i \quad = \quad v_k \qquad \text{für} \quad 1 \le k \le n.$$

Dies zeigt, dass $v_1, ..., v_n$ in $\text{lin}\{w_1, ..., w_n\}$ enthalten sind. Weil $\{v_1, ..., v_n\}$ eine Basis von $V$ ist, folgt daraus, dass $\{w_1, ..., w_n\}$ ein Erzeugendensystem von $V$ ist, wegen $n = \dim V$ sogar eine Basis, nach Satz (7.7). Damit ist dann das Tupel $\mathscr{B} = (w_1, ..., w_n)$ eine geordnete Basis von $V$ mit der gewünschten Eigenschaft $\mathscr{T}_{\mathscr{B}}^{\mathscr{A}} = T$. □

Wir können nun ein neues Kriterium für die Diagonalisierbarkeit herleiten.

**(12.20) Proposition** Sei $V \ne \{0_V\}$ ein endlich-dimensionaler $K$-Vektorraum und $\phi \in \text{End}_K(V)$. Dann sind die folgenden Aussagen äquivalent:

(i) Der Endomorphismus $\phi$ ist diagonalisierbar.

(ii) Der Vektorraum $V$ besitzt eine Basis bestehend aus Eigenvektoren von $\phi$.

***Beweis:*** „(i) ⇒ (ii)“ Nach Voraussetzung gibt es eine Basis $\mathscr{A} = (v_1, ..., v_n)$ von $V$ mit der Eigenschaft, dass $D = \mathscr{M}_{\mathscr{A}}(\phi)$ eine Diagonalmatrix ist, $D = \text{diag}(\lambda_1, ..., \lambda_n)$ mit $\lambda_k \in K$ für $1 \le k \le n$. Der $k$-te Spaltenvektor von $D$ ist jeweils das $\lambda_k$-fache des $k$-ten Einheitsvektors $e_k$. Es folgt

$$De_k = \lambda_k e_k \quad \Leftrightarrow \quad \mathscr{M}_{\mathscr{A}}(\phi)\Phi_{\mathscr{A}}(v_k) = \lambda_k \Phi_{\mathscr{A}}(v_k) \quad \Leftrightarrow \quad \Phi_{\mathscr{A}}(\phi(v_k)) = \Phi_{\mathscr{A}}(\lambda_k v_k) \quad \Leftrightarrow \quad \phi(v_k) = \lambda_k v_k$$

für $1 \le k \le n$, wobei im letzten Schritt die Bijektivität von $\Phi_{\mathscr{A}}$ verwendet wurde. Als Element einer Basis ist $v_k \ne 0_V$; zusammen mit der Gleichung $\phi(v_k) = \lambda_k v_k$ zeigt dies, dass $\mathscr{A}$ aus Eigenvektoren von $\phi$ besteht.

„(ii) ⇒ (i)“ Sei $\mathscr{A} = (v_1, ..., v_n)$ eine Basis von $V$, wobei $v_k$ jeweils ein Eigenvektor von $\phi$ zum Eigenwert $\lambda_k$ ist, für $1 \le k \le n$. Außerdem sei $D = \mathscr{M}_{\mathscr{A}}(\phi)$. Dann gilt jeweils $\phi(v_k) = \lambda_k v_k$, und die Rechnung aus dem vorherigen Absatz hat gezeigt, dass dies äquivalent zu $De_k = \lambda_k e_k$ ist. Die $k$-te Spalte von $D$ ist also gleich $\lambda_k e_k$, für $1 \le k \le n$. Daraus folgt $D = \text{diag}(\lambda_1, ..., \lambda_n)$, also ist $D$ eine Diagonalmatrix und $\phi$ damit diagonalisierbar. □

Als nächstes zeigen wir, dass der Vektorraum $V$ bezüglich eines diagonalisierbaren Endomorphismus in Eigenräume zerlegt werden kann. Die beiden folgenden Aussagen dienen zur Vorbereitung.

**(12.21) Proposition** Sei $V$ ein $K$-Vektorraum, $\phi \in \text{End}_K(V)$, und seien $\lambda_1, ..., \lambda_r$ verschiedene Elemente von $K$. Für jedes $k \in \{1, ..., r\}$ sei $v_k \in V$ jeweils ein Eigenvektor zum Eigenwert $\lambda_k$. Dann ist das Tupel $(v_1, ..., v_r)$ linear unabhängig.

***Beweis:*** Wir führen den Beweis durch vollständige Induktion über $r$. Für $r = 1$ ist die Aussage wegen $v_1 \neq 0_V$ klar. Sei nun $r \in \mathbb{N}$, und setzen wir nun die Behauptung für dieses $r$ voraus. Es seien $\lambda_1, ..., \lambda_{r+1} \in K$ verschieden, und sei $v_k \in V$ für $1 \le k \le r+1$ jeweils ein Eigenvektor zum Eigenwert $\lambda_k$. Zum Nachweis der linearen Unabhängigkeit seien $\alpha_1, ..., \alpha_{r+1} \in K$ mit der Eigenschaft

$$\sum_{k=1}^{r+1} \alpha_k v_k \quad = \quad 0_V. \tag{12.1}$$

Dann liefert die Multiplikation von (12.1) mit dem Wert $\lambda_{r+1}$ einerseits

$$\sum_{k=1}^{r+1} \alpha_k \lambda_{r+1} v_k \quad = \quad 0_V \quad , \tag{12.2}$$

andererseits erhält man durch Anwendung von $\phi$ auf (12.1) aber auch

$$\sum_{k=1}^{r+1} \alpha_k \lambda_k v_k \quad = \quad \sum_{k=1}^{r+1} \alpha_k \phi(v_k) \quad = \quad \phi\left(\sum_{k=1}^{r+1} \alpha_k v_k\right) \quad = \quad \phi(0_V) \quad = \quad 0_V. \tag{12.3}$$

Subtrahieren wir die Gleichungen (12.2) und (12.3) voneinander, so erhalten wir

$$\sum_{k=1}^{r+1} \alpha_k \lambda_{r+1} v_k - \sum_{k=1}^{r+1} \alpha_k \lambda_k v_k \quad = \quad \sum_{k=1}^{r+1} \alpha_k (\lambda_{r+1} - \lambda_k) v_k \quad = \quad 0_V.$$

Da das Tupel $(v_1, ..., v_r)$ nach Induktionsvoraussetzung linear unabhängig sind, folgt $\alpha_k(\lambda_{r+1} - \lambda_k) = 0_K$ für $1 \le k \le r$. Wegen $\lambda_{r+1} - \lambda_k \neq 0_K$ folgt $\alpha_k = 0_K$ für $1 \le k \le r$. Setzen wir dies wiederum in (12.1) ein, so erhalten wir $\alpha_{r+1} v_{r+1} = 0_V$, und wegen $v_{r+1} \neq 0_V$ folgt $\alpha_{r+1} = 0_K$. Damit ist die linearen Unabhängigkeit von $(v_1, ..., v_{r+1})$ nachgewiesen. □

**(12.22) Proposition** Sei $\phi$ ein Endomorphismus eines $K$-Vektorraums $V$, und seien $\lambda_1, ..., \lambda_r$ verschiedene Elemente des Körpers $K$. Dann gilt

$$\mathrm{Eig}(\phi, \lambda_k) \cap \left(\sum_{\ell \neq k} \mathrm{Eig}(\phi, \lambda_\ell)\right) \quad = \quad \{0_V\} \qquad \text{für} \quad 1 \le k \le r.$$

***Beweis:*** Nehmen wir an, dass ein $k \in \{1, ..., r\}$ und ein Vektor $v \neq 0_V$ in der angegebenen Schnittmenge. Dann gibt es Vektoren $v_\ell \in \mathrm{Eig}(\phi, \lambda_\ell)$ für $1 \le \ell \le r$ mit

$$v_k = v = \sum_{\ell \neq k} v_\ell \quad \Leftrightarrow \quad \sum_{\ell \neq k} v_\ell + (-1) v_k \quad = \quad 0_V.$$

Aber die Menge $\{v_\ell \mid \ell \in \{1, ..., r\} \text{ mit } v_\ell \neq 0\}$ besteht aus Eigenvektoren zu verschiedenen Eigenwerten, ist also nach (12.21) linear unabhängig, und nichtleer, da sie zumindest $v_k$ enthält. Also kann die Gleichung $\sum_{\ell \neq k} v_\ell + (-1) v_k = 0_V$ nicht gelten. □

Mit diesen Ergebnissen erhalten wir ein neues Kriterium für die Diagonalisierbarkeit eines Endomorphismus.

**(12.23) Proposition** Sei $V \neq \{0_V\}$ endlich-dimensionaler $K$-Vektorraum und $\phi \in \mathrm{End}_K(V)$. Dann sind die folgenden Aussagen äquivalent:

(i) Der Endomorphismus $\phi$ ist diagonalisierbar.

(ii) Es gibt verschiedene Elemente $\lambda_1, ..., \lambda_r \in K$, so dass

$$V \quad = \quad \bigoplus_{\ell=1}^{r} \mathrm{Eig}(\phi, \lambda_\ell) \qquad \text{erfüllt ist.}$$

***Beweis:*** „(i) ⇒ (ii)" Auf Grund der Voraussetzung existiert nach (12.20) eine Basis $\mathscr{A} = \{v_1, ..., v_n\}$ von $V$ bestehend aus Eigenvektoren. Seien $\lambda_1, ..., \lambda_r$ die verschiedenen Eigenwerte von $\phi$. Weil alle Elemente der Basis Eigenvektoren sind, gibt es für jedes $k \in \{1, ..., n\}$ ein $\ell \in \{1, ..., r\}$ mit $\phi(v_k) = \lambda_\ell v_k$. Es gilt dann also $v_k \in \mathrm{Eig}(\phi, \lambda_\ell)$. Setzen wir $U = \sum_{\ell=1}^{r} \mathrm{Eig}(\phi, \lambda_\ell)$, dann gilt insgesamt $\mathscr{A} \subseteq U$. Weil $\mathscr{A}$ eine Basis und $U$ ein Untervektorraum von $V$ ist, stimmt $V$ mit der Summe $U$ ein, und nach (12.22) ist diese Summe direkt.

„(ii) ⇒ (i)" Für jedes $\ell \in \{1, ..., r\}$ sei $\mathscr{A}_\ell$ eine Basis von $\mathrm{Eig}(\phi, \lambda_\ell)$. Auf Grund der direkten Summenzerlegung ist dann $\mathscr{A} = \bigcup_{\ell=1}^{r} \mathscr{A}_\ell$ nach (8.3) eine Basis von $V$. Jedes $\mathscr{A}_\ell$ besteht aus Eigenvektoren von $\phi$, somit auch die Basis $\mathscr{A}$. Nach (12.20) folgt daraus die Diagonalisierbarkeit von $\phi$. □

**(12.24) Definition** Sei $V$ ein endlich-dimensionaler $K$-Vektorraum, $\phi$ ein Endomorphismus von $V$ und $\lambda \in K$ ein Eigenwert von $\phi$.

(i) Die Vielfachheit $\mu(\chi_\phi, \lambda)$ von $\lambda$ als Nullstelle des Polynoms $\chi_\phi$ bezeichnet man als ***algebraische*** Vielfachheit $\mu_a(\phi, \lambda)$ des Eigenwerts $\lambda$.

(ii) Die Eigenraum-Dimension $\mu_g(\phi, \lambda) = \dim \mathrm{Eig}(\phi, \lambda)$ nennt man die ***geometrische*** Vielfachheit von $\lambda$.

Für eine quadratische Matrix $A \in \mathscr{M}_{n,K}$ definiert man algebraische und geometrische Vielfachheit eines Eigenwerts $\lambda$ auf analoge Weise: Man setzt $\mu_a(A, \lambda) = \mu(\chi_A, \lambda)$ und $\mu_g(A, \lambda) = \dim \mathrm{Eig}(A, \lambda)$.

In Beweisen ist es oft günstig, wenn die algebraische und geometrische Vielfachheit auch für Nicht-Eigenwerte eines Endomorphismus $\phi$ definiert sind, weil dadurch einige Fallunterscheidungen unnötig werden. Wenn $\lambda \in K$ kein Eigenwert von $\phi$ ist, dann setzt man $\mu_a(\phi, \lambda) = \mu_g(\phi, \lambda) = 0$.

**(12.25) Proposition** Sei $V$ ein $n$-dimensionaler $K$-Vektorraum, $\phi : V \to V$ ein Endomorphismus und $\lambda \in K$ ein Eigenwert mit algebraischer Vielfachheit $\mu_a$ und geometrischer Vielfachheit $\mu_g$. Dann gilt $1 \leq \mu_g \leq \mu_a$.

**Beweis:** Wir haben bereits gesehen, dass $\lambda$ genau dann ein Eigenwert ist, wenn $\mathrm{Eig}(\phi,\lambda) \neq \{0_V\}$ ist. Deshalb gilt $\mu_g \geq 1$. Sei nun $(v_1,...,v_r)$ eine Basis von $\mathrm{Eig}(\phi,\lambda)$, die wir durch $v_{r+1},...,v_n$ zu einer Basis $\mathscr{A}$ von $V$ ergänzen. Wegen $\phi(v_i) = \lambda v_i$ für $1 \leq i \leq r$ hat $A$ die Form

$$A \quad = \quad \begin{pmatrix} \lambda \mathrm{I}^{(r)} & B \\ 0 & C \end{pmatrix}$$

mit geeignet gewählten Matrizen $B \in \mathscr{M}_{r\times(n-r),K}$ und $C \in \mathscr{M}_{n-r,K}$. Sei nun $\chi \in K[x]$ das charakteristische Polynom von $\phi$ und $\tilde{K}$ ein unendlicher Erweiterungskörper von $K$. Für alle $\alpha \in \tilde{K}$ gilt dann

$$\chi(\alpha) \quad = \quad \det(\alpha E^{(n)} - A) \quad = \quad \det\begin{pmatrix} (\alpha-\lambda)E^{(r)} & -B \\ 0 & \alpha E^{(n-r)} - C \end{pmatrix} \quad = \quad (\alpha-\lambda)^r \det(\alpha E^{(n-r)} - C) \quad ,$$

wobei wir im letzten Schritt (11.4) für Blockmatrizen angewendet haben. Damit erhalten wir für das charakteristische Polynom $\chi = (x-\lambda)^r g = (x-\lambda)^r g$, wobei $g = \det(xE^{(n-r)} - C) \in K[x]$ das charakteristische Polynom der Matrix $C$ bezeichnet. Dies beweist $\mu_g = r \leq \mu_a$. □

**(12.26) Satz** Sei $n \in \mathbb{N}$, $V$ ein $n$-dimensionaler $K$-Vektorraum, $\phi$ ein Endomorphismus von $V$ und $\chi \in K[x]$ sein charakteristisches Polynom. Dann sind die folgenden Aussagen äquivalent:

(i) Der Endomorphismus $\phi$ ist diagonalisierbar

(ii) Der Vektorraum $V$ besitzt eine Basis bestehend aus Eigenvektoren von $\phi$.

(iii) Das Polynom $\chi_\phi$ zerfällt in Linearfaktoren, und für jeden Eigenwert $\lambda$ von $\phi$ stimmen algebraische und geometrische Vielfachheit überein.

(iv) Es gibt $\lambda_1,...,\lambda_r \in K$, so dass $V = \mathrm{Eig}(\phi,\lambda_1) \oplus ... \oplus \mathrm{Eig}(\phi,\lambda_r)$ gilt.

**Beweis:** Die Äquivalenz von (i), (ii) und (iv) wurde bereits bewiesen.

„(i) ⇒ (iii)“ Sei $\mathscr{A} = (v_1,...,v_n)$ eine Basis, und seien $\lambda_1,...,\lambda_n \in K$, so dass $A = M_{\mathscr{A}}(\phi) = \mathrm{diag}(\lambda_1,...,\lambda_n)$ gilt. Sei außerdem $\tilde{K} \supseteq K$ ein unendlicher Erweiterungskörper. Dann gilt für alle $\alpha \in \tilde{K}$ die Gleichung

$$\chi(\alpha) \quad = \quad \det(\alpha E^{(n)} - A) \quad = \quad \det(\mathrm{diag}(\alpha-\lambda_1,...,\alpha-\lambda_n)) \quad = \quad \prod_{i=1}^{n}(\alpha-\lambda_i)$$

und somit $\chi = \prod_{i=1}^{n}(x-\lambda_i)$. Also zerfällt das charakteristische Polynom in Linearfaktoren. Zu zeigen bleibt, dass für jeden Eigenwert $\lambda_i$ jeweils algebraische und geometrische Vielfachheit übereinstimmen. Für jeden Eigenwert $\lambda$ von $\phi$ sei $S(\lambda) = \{i \mid \lambda_i = \lambda\}$. Dann können wir das charakteristische Polynom in der Form $\chi = gh$ mit

$$g \quad = \quad \prod_{i\in S(\lambda)} (x-\lambda_i) \quad = \quad (x-\lambda)^{|S(\lambda)|} \qquad \text{und} \qquad h \quad = \quad \prod_{i\notin S(\lambda)} (x-\lambda_i)$$

zerlegen. Wegen $h(\lambda) \neq 0$ ist $\mu = |S(\lambda)|$ die Vielfachheit der Nullstelle $\lambda$ von $f$, also die algebraische Vielfachheit des Eigenwerts $\lambda$. Für jedes $i \in S(\lambda)$ ist aber auch $v_i$ ein Eigenvektor zum Eigenwert $\lambda$, es gilt also $\dim \mathrm{Eig}(\phi,\lambda) \geq \mu$. Weil die geometrische Vielfachheit nach (12.25) immer durch die algebraische Vielfachheit beschränkt ist, folgt $\mathrm{Eig}(\phi,\lambda) = \mu$.

„(iii) $\Rightarrow$ (iv)" Sei $\chi = a\prod_{i=1}^r (x-\lambda_i)^{\mu_i}$ eine Zerlegung des charakteristischen Polynoms in Linearfaktoren, wobei $a \in K$, die Nullstellen $\lambda_1,...,\lambda_r \in K$ paarweise verschieden und $\mu_i \in \mathbb{N}$ jeweils die Vielfachheiten der Nullstellen $\lambda_i$ sind. Wegen $\mathrm{grad}(f) = n$ gilt $\sum_{i=1}^r \mu_i = n$.

Für jedes $i \in \{1,...,r\}$ definieren wir nun $U_i = \mathrm{Eig}(\phi,\lambda_i)$, außerdem sei $V_r = \sum_{i=1}^r U_i$. Da algebraische und geometrische Vielfachheit übereinstimmen, gilt $\mu_i = \dim U_i$ für $1 \leq i \leq r$. Nach (12.22) ist der Durchschnitt von $U_i$ mit $\sum_{j\neq i} U_j$ jeweils gleich $\{0_V\}$. Dies zeigt, dass $V_r$ die *direkte* Summe der Untervektorräume $U_i$ ist. Nach (8.3) erhalten wir

$$\dim V_r \quad = \quad \sum_{i=1}^r \dim U_i \quad = \quad \sum_{i=1}^r \mu_i \quad = \quad n \quad = \quad \dim V.$$

Aus $V_r \subseteq V$ und $\dim V_r = \dim V$ wiederum folgt $V = V_r$. □

Die Diagonalisierbarkeitskriterien lassen sich ohne großen Aufwand auf Matrizen übertragen. Hier gilt entsprechend

**(12.27) Folgerung** Für jede Matrix $A \in \mathscr{M}_{n,K}$ sind folgende Aussagen äquivalent.

(i) Die Matrix $A$ ist diagonalisierbar.

(ii) Der Vektorraum $K^n$ besitzt eine Basis bestehend aus Eigenvektoren von $A$.

(iii) Das Polynom $\chi_A$ zerfällt in Linearfaktoren, und es gilt $\chi_a(A,\lambda) = \chi_g(A,\lambda)$ für jeden Eigenwert $\lambda$ von $A$.

(iv) Es gibt $\lambda_1,...,\lambda_r \in K$, so dass $K^n = \mathrm{Eig}(A,\lambda_1) \oplus ... \oplus \mathrm{Eig}(A,\lambda_r)$ gilt.

***Beweis:*** Sei $\phi_A$ der Endomorphismus von $K^n$ gegeben durch $\phi_A(v) = Av$. Dann ist $A$ die Darstellungsmatrix von $\phi_A$ bezüglich der Einheitsbasis $\mathscr{E}_n$ von $K^n$, siehe (9.8). Somit ist (i) nach (12.18) äquivalent zur Diagonalisierbarkeit von $\phi_A$, also zur Ausage (i) in Satz (12.26) für den Endomorphismus $\phi_A$. Ebenso leicht überprüft man mit Hilfe der Übereinstimmungen $\mathrm{Eig}(A,\lambda_i) = \mathrm{Eig}(\phi_A,\lambda_i)$ und $\chi_A = \chi_{\phi_A}$, dass die Aussagen (ii),(iii),(iv) äquivalent zu ihrem jeweiligen Pendant in Satz (12.26) sind. Damit ist die Äquivalenz der Aussagen (i) bis (iv) auf Satz (12.26) zurückgeführt. □

Wir demonstrieren die Anwendung des Diagonalisierbarkeitskriteriums an einem etwas größeren Rechenbeispiel und testen die Matrix

$$A \quad = \quad \begin{pmatrix} -2 & -10 & 0 & -40 \\ -24 & -13 & -8 & -68 \\ 15 & 30 & 3 & 120 \\ 6 & 4 & 2 & 20 \end{pmatrix}$$

auf Diagonalisierbarkeit über dem Körper $K = \mathbb{Q}$. (Das ist die Matrix vom Beginn des Kapitels.)

Wir beginnen mit der Bestimmung des charakteristischen Polynoms. Durch Anwendung des Laplaceschen Entwicklungssatzes erhalten wir

$$\chi_A \quad = \quad \det(xE^{(4)} - A) \quad = \quad \det \begin{pmatrix} x+2 & 10 & 0 & 40 \\ 24 & x+13 & 8 & 68 \\ -15 & -30 & x-3 & -120 \\ -6 & -4 & -2 & x-20 \end{pmatrix} \quad =$$

$$(x+2)\cdot\det\begin{pmatrix} x+13 & 8 & 68 \\ -30 & x-3 & -120 \\ -4 & -2 & x-20 \end{pmatrix} - 10\cdot\det\begin{pmatrix} 24 & 8 & 68 \\ -15 & x-3 & -120 \\ -6 & -2 & x-20 \end{pmatrix} - 40\cdot\det\begin{pmatrix} 24 & x+13 & 8 \\ -15 & -30 & x-3 \\ -6 & -4 & -2 \end{pmatrix}$$

$$= \quad (x+2)(x^3 - 10x^2 + 33x - 36) + (-10)(24x^2 - 24x - 144) + (-40)(-6x^2 + 6x + 36)$$

$$= \quad x^4 - 8x^3 + 13x^2 + 30x - 72.$$

Als nächstes bestimmen wir die Eigenwerte von $A$, also die Nullstellen von $\chi_A$. Durch probeweises Einsetzen findet man die Nullstelle $\lambda_1 = -2$, und Polynomdivision liefert die Zerlegung $\chi_A = (x+2)(x^3 - 10x^2 + 33x - 36)$. Erneutes probeweises Einsetzen in den Faktor $g = x^3 - 10x^2 + 33x - 36$ vom Grad 3 liefert die Nullstelle $\lambda_2 = 3$. Durch Polynomdivision erhalten wir $g = (x-3)(x^2 - 7x + 12)$. Die Nullstellen des quadratischen Faktors können wir nun durch Anwendung der $p$-$q$-Formel (auf $p = -7$, $q = 12$) ermitteln: Die Diskriminante des quadratischen Polynoms ist $d = p^2 - 4q = (-7)^2 - 4\cdot 12 = 49 - 48 = 1$, die Nullstellen somit $-\frac{1}{2}p \pm \frac{1}{2}\sqrt{d} = \frac{7}{2} \pm \frac{1}{2}$, also $\lambda_3 = 3$ und $\lambda_4 = 4$. Insgesamt erhalten wir damit die Zerlegung

$$\chi_A \quad = \quad (x+2)(x-3)^2(x-4).$$

Die Eigenwerte von $A$ sind also $-2$, 3 und 4, mit den algebraischen Vielfachheiten $\mu_a(A,-2) = 1$, $\mu_a(A,3) = 2$ und $\mu_a(A,4) = 2$. Die Gleichung zeigt auch, dass $\chi_A$ in Linearfaktoren zerfällt. Damit haben wir bereits einen Teil des Diagonalisierbarkeitskriteriums verifiziert.

Nun müssen wir noch die geometrischen Vielfachheiten der Eigenwerte bestimmen. Nach (12.25) gilt $1 \le \mu_g(A,\lambda) \le \mu_a(A,\lambda)$ für jeden Eigenwert $\lambda$ von $A$. Wegen $\mu_a(A,-2) = 1$ folgt daraus $\mu_g(A,-2) = 1$, und ebenso erhalten wir $\mu_g(A,4) = 1$. Für $\mu_g(A,3)$ sind wegen $\mu_a(A,3) = 2$ noch zwei Möglichkeiten offen, nämlich $\mu_g(A,3) \in \{1,2\}$. Um festzustellen, welcher Fall vorliegt, bestimmen wir die Dimension von $\mathrm{Eig}(A,3) = \ker(A - 3E^{(4)})$ mit dem Gaußverfahren.

$$A - 3E^{(4)} = \begin{pmatrix} -5 & -10 & 0 & -40 \\ -24 & -16 & -8 & -68 \\ 15 & 30 & 0 & 120 \\ 6 & 4 & 2 & 17 \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} 1 & 2 & 0 & 8 \\ -6 & -4 & -2 & -17 \\ 1 & 2 & 0 & 8 \\ 6 & 4 & 2 & 17 \end{pmatrix}$$

$$\mapsto \quad \begin{pmatrix} 1 & 2 & 0 & 8 \\ -6 & -4 & -2 & -17 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} 1 & 2 & 0 & 8 \\ 0 & 8 & -2 & 31 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Die letzte Matrix liegt in ZSF vor, mit Zeilenrang $r = 2$. Nach Satz (8.11) folgt $\mu_g(A,3) = \dim \mathrm{Eig}(A,3) = \ker(A - 3E^{(4)}) = 4 - r = 4 - 2 = 2 = \mu_a(A,3)$. Insgesamt ist damit $\mu_g(A,\lambda) = \mu_a(A,\lambda)$ für alle Eigenwert $\lambda$ der Matrix erfüllt. Dass $\chi_A$ über $\mathbb{Q}$ in Linearfaktoren zerfällt, haben wir bereits oben festgestellt. Nach (12.27) ist $A$ also eine diagonalisierbare Matrix.

Betrachten wir nun noch zwei Beispiele, in denen das Diagonalisierbarkeitskriterium nicht erfüllt ist. Beispielsweise ist die Matrix

$$B \quad = \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

über dem Körper $K = \mathbb{R}$ nicht diagonalisierbar. Zwar zerfällt das charakteristsiche Polynom wegen

$$\chi_B \quad = \quad \det(xE^{(2)} - B) \quad = \quad \det \begin{pmatrix} x-1 & 1 \\ 0 & x-1 \end{pmatrix} \quad = \quad (x-1)^2$$

über $\mathbb{R}$ in Linearfaktoren. Wie man aber leicht überprüft, gilt $\mathrm{Eig}(B,1) = \mathrm{lin}(e_1)$, der Eigenraum von $B$ zum Eigenwert 1 ist also nur eindimensional. Es gilt also $\mu_g(B,1) = 1$, andererseits aber $\mu_a(B,1) = 2$, weil 1 eine doppelte Nullstelle von $B$ ist. Aus $\mu_g(B,1) < \mu_a(B,1)$ folgt nach (12.27), dass $B$ über $\mathbb{R}$ nicht diagonalisierbar ist.

Auch die Matrix

$$C \quad = \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

ist über $\mathbb{R}$ nicht diagonalisierbar. Das charakteristische Polynom $\chi_C = x^2 + 1$ besitzt nämlich in $\mathbb{R}$ keine Nullstellen und zerfällt somit über $\mathbb{R}$ auch nicht in Linearfaktoren. Anders sieht die Sache allerdings aus, wenn man $C$ als Matrix über dem Körper $\mathbb{C}$ betrachtet. Dann besitzt $\chi_C$ eine Produktzerlegung $(x-i)(x+i)$. Aus den Ungleichungen $1 \le \mu_g(C,\lambda) \le \mu_a(C,\lambda)$ für $\lambda \in \{\pm i\}$ folgt, dass auch die Bedingungen $\mu_g(C,i) = \mu_a(C,i)$ und $\mu_g(C,-i) = \mu_a(C,-i)$ erfüllt sind. Also ist die Matrix $C$ über dem Körper $\mathbb{C}$ diagonalisierbar.

# *Literaturverzeichnis*


[Bo] S. Bosch, *Lineare Algebra*. Springer-Lehrbuch, Berlin 2006.

[Fi] G. Fischer, *Lernbuch Lineare Algebra und Geometrie*. Vieweg-Teubner, Wiesbaden 2011.

[Jn] K. Jaenich, *Lineare Algebra*. Springer-Verlag, Berlin 2001.

[dJ] T. de Jong, *Lineare Algebra*. Pearson-Studium, München 2013.